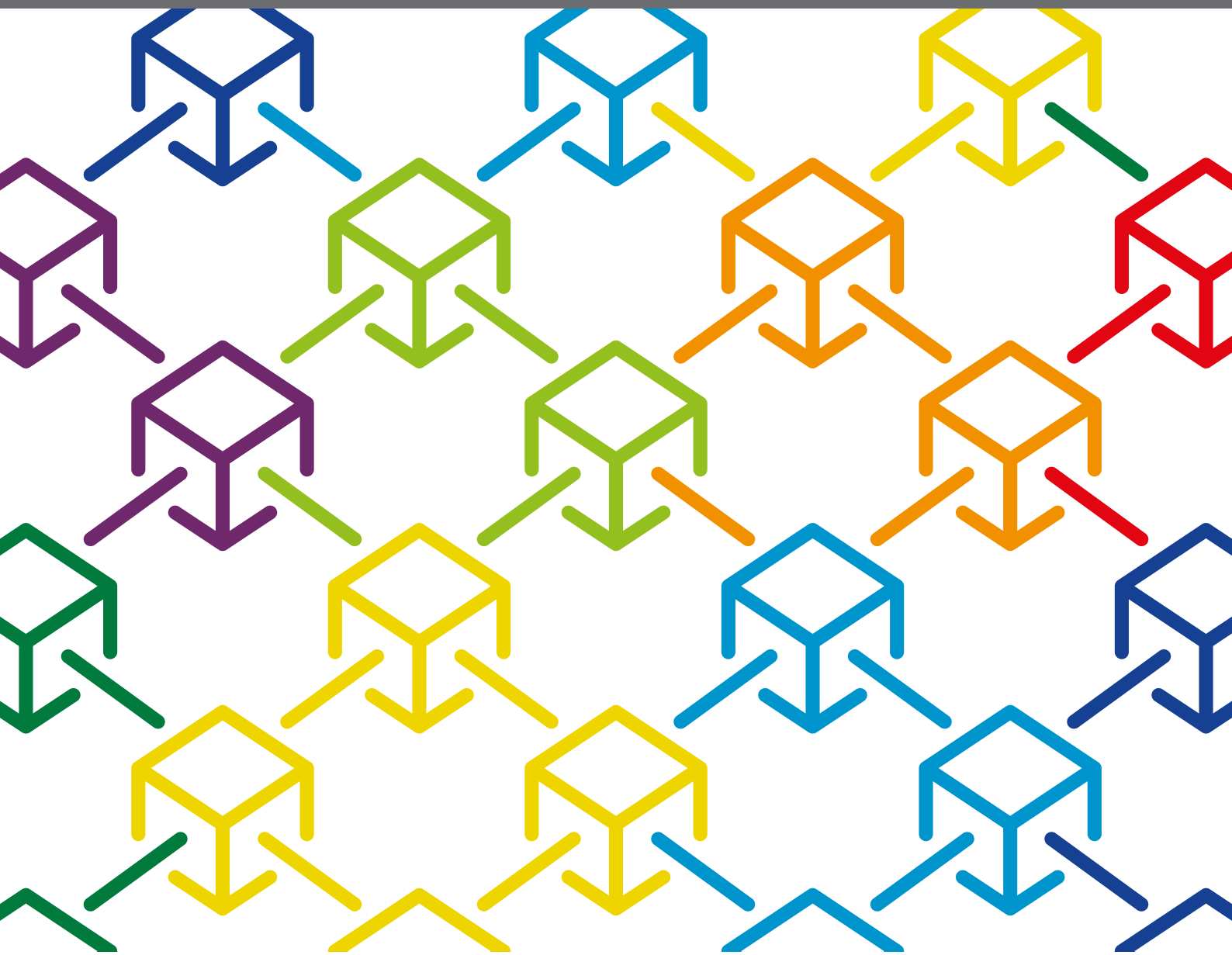
IDENTITY AND PRIVACY GOVERNANCE

EDITED BY: Andrej Zwitter and Oskar Josef Gstrein
PUBLISHED IN: Frontiers in Blockchain



frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88971-413-1
DOI 10.3389/978-2-88971-413-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

IDENTITY AND PRIVACY GOVERNANCE

Topic Editors:
Andrej Zwitter, University of Groningen, Netherlands
Oskar Josef Gstrein, University of Groningen, Netherlands

Citation: Zwitter, A., Gstrein, O. J., eds. (2021). Identity and Privacy Governance. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-413-1

Table of Contents

frontiers in Blockchain

EDITORIAL
published: 06 August 2021
doi: 10.3389/fbloc.2021.738862

Editorial: Identity and Privacy Governance

Andrej Zwitter and Oskar J. Gstrein *

University of Groningen, Campus Fryslân, Leeuwarden, Netherlands

Keywords: digital identity, governance, data protection and privacy, human dignity, blockchain for good

Editorial on the Research Topic

Identity and Privacy Governance

The design and management of digital identity is a complex challenge. On the one hand, it requires a clear understanding of the parameters that are involved in identity management. On the other hand, it requires the cooperation of many stakeholders. In particular, this involves those public authorities and private organisations that need to be aligned to define technical standards, develop identification infrastructures and maintain them. A shared understanding of fundamental concepts that define identity in the digital age is then a prerequisite. Such a complimentary reflection and evaluation of what the emergence of distributed-ledger technologies means from the perspectives of human rights, human dignity, as well as individual and collective autonomy are essential to ensure their use for good purposes. While technical capabilities are important, they are increasingly insufficient without guiding theoretical frameworks. Sound governance mechanisms which respect, protect and promote human rights such as privacy are equally essential. The COVID-19 pandemic has only further increased the desire to use data to understand and manage our societies (Zwitter and Gstrein, 2020), which also increases the degree to which we are defined through data and our access to digital services.

Certainly, we currently witness profound changes in the capabilities to define and manage identity. Established architectures to validate, certify, and manage credentials are usually based on centralized or federated top-down approaches. They rely on territorial sovereignty, trusted authorities and third-party operators which gain considerable power by being able to manage the systems. In recent years, distributed-ledger technologies such as Blockchain have been described as "trust mechanisms", which can operate independently of such trust-mediators and territorial restrictions. One might prefer to rather trust a technical system, as well as the parties that host the software and ensure proper functioning, than traditional institutions such as banks and states. This emerging opportunity to change the practice of identity management raises the questions of 1) how blockchain applications influence trust, and 2) how trust based requirements affect the design of applications based on distributed-ledger technology?

Some identity management architectures presented in this research topic go even further and design full-fledged identity management systems. Their users are not only independent from the gatekeepers mentioned above. They also do not need to maintain a single aggregated identity. This enhances privacy and autonomy, so the authors argue, since aggregated identities can potentially be constrained or reconstructed against the interests of individuals. Such a pattern change could also potentially mitigate information security issues. These security issues are becoming more and more pressing as conventional digital identity management based on passwords and e-mail addresses face enhanced cybersecurity threats, typically associated with identity theft. Nevertheless, private forms of digital identity governance can also create worrying consequences from a security perspective, as the case of "Silk Road"—a historically influential platform for trading on the "dark web"—demonstrates.

OPEN ACCESS

Edited and reviewed by:
Richard Adams,
Cranfield University, United Kingdom

****Correspondence:***
Oskar J. Gstrein
o.j.gstrein@rug.nl

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *09 July 2021*
Accepted: *27 July 2021*
Published: *06 August 2021*

Citation:
Zwitter A and Gstrein OJ (2021) Editorial: Identity and Privacy Governance. Front. Blockchain 4:738862. doi: 10.3389/fbloc.2021.738862

A more hopeful perspective is offered by digital identity management systems that aim at leveraging the potential of "self-sovereign identities" to become a driver for economic inclusion in some regions of the world. These pilots could help to demonstrate the potential of "Blockchain for good", but only if concerns associated with the use of biometric data and autonomy are mitigated. Still, new business models might emerge, such as identity insurance schemes, along with the emergence of value-stable cryptocurrencies ("stablecoins") functioning as local currencies. It remains to be seen how public institutions react to the emergence of these new opportunities. The impact of innovative approaches to digital identity management is not missed by intergovernmental organisations such as the Financial Action Task Force (FATF), which is at the centre of global anti-money laundering and counter-the-financing of terrorism. While FATF is not directly involved in the actual coding of protocols it influences the location and type of centralized modes of control over digital identity governance. In highlighting both the influence of FATF on blockchain governance and blockchain governance on the FATF, it is possible to draw together research areas which have been considered separately. A combination of perspectives might be helpful to understand the future of global digital identity governance more holistically.

With the same objective of developing holistic approaches, some articles of this research topic outline the underlying fundamentals by exploring philosophical conceptualisations of digital identity management. While a naturalist world view establishes identity as a concept that hinges on the concept of uniqueness, it also evokes questions on the dependence and interaction of an individual with its environment and society. Proponents of a constructivist identity emphasize relationality while questions of identity as a complete individual entity remain. At the same time, when considering the legal domain, it can be observed that particularly in the human rights space, identity is determined by several individual rights that states are obliged to grant to individuals. Furthermore, aspects around the ownership of material and immaterial goods (e.g., intellectual property) ultimately highlight the issue of "data ownership" which could be essential to keep rights enforceable on a universal level in the digital domain. These arguments and insights might inspire the design of innovative governance frameworks. Nevertheless, it is also necessary to consider how traditional identity management systems such as citizenship engage and intersect with the emerging technological capabilities. It cannot be overlooked that the development and implementation of digital identity management systems using distributed-ledger technology raise multiple ethical and moral issues.

As editors of this research topic, we can only be grateful for the insights and ideas the authors have shared with us. We hope that the readers of the contributions to this edited volume share our excitement when exploring their content. To us it seems that the development of digital identity systems will continue to remain an important topic in the years to come. Currently, the development and implementation of "vaccine passports" and digital COVID-19 vaccination certificates might eventually morph into general purpose infrastructures that also receive broader tasks in identity management. These and similar developments result in a chorus of ethical, legal and social issues that need to be addressed (Gstrein et al., 2021), and for which the research presented in this research topic provides a rich basis.

AUTHOR CONTRIBUTIONS

OJ provided the first draft which was reviewed and edited by AZ.

REFERENCES

Gstrein, O. J., Kochenov, D. V., and Zwitter, A. (2021). A Terrible Great Idea? COVID-19 'Vaccination Passports' in the Spotlight. Centre on Migration,. *Pol. Soc. Working Pap.* 153, 28.

Zwitter, A., and Gstrein, O. J. (2020). Big Data, Privacy and COVID-19 - Learning from Humanitarian Expertise in Data protection. *Int. J. Humanitarian Action.* 5 (4), 00072–00076. doi:10.1186/s41018-020-00072-6

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

METHODS
published: 05 November 2019
doi: 10.3389/fbloc.2019.00017

A Decentralized Digital Identity Architecture

Geoff Goodell and Tomaso Aste*

Centre for Blockchain Technologies, University College London, London, United Kingdom

Current architectures to validate, certify, and manage identity are based on centralized, top-down approaches that rely on trusted authorities and third-party operators. We approach the problem of digital identity starting from a human rights perspective, with a primary focus on identity systems in the developed world. We assert that individual persons must be allowed to manage their personal information in a multitude of different ways in different contexts and that to do so, each individual must be able to create multiple unrelated identities. Therefore, we first define a set of fundamental constraints that digital identity systems must satisfy to preserve and promote privacy as required for individual autonomy. With these constraints in mind, we then propose a decentralized, standards-based approach, using a combination of distributed ledger technology and thoughtful regulation, to facilitate many-to-many relationships among providers of key services. Our proposal for digital identity differs from others in its approach to trust in that we do not seek to bind credentials to each other or to a mutually trusted authority to achieve strong non-transferability. Because the system does not implicitly encourage its users to maintain a single aggregated identity that can potentially be constrained or reconstructed against their interests, individuals and organizations are free to embrace the system and share in its benefits.

Keywords: identity, privacy, distributed ledgers, authentication/authorization, unlinkability, self-sovereign identity, early binding, tokens

OPEN ACCESS

Edited by:
Andrej Zwitter,
University of Groningen, Netherlands

Reviewed by:
Nichola Cooper,
University of the Sunshine Coast, Australia
Raul Zambrano,
Independent Researcher, New York, NY, United States

****Correspondence:***
Geoff Goodell
g.goodell@ucl.ac.uk

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *13 August 2019*
Accepted: *17 October 2019*
Published: *05 November 2019*

Citation:
Goodell G and Aste T (2019) A Decentralized Digital Identity Architecture. Front. Blockchain 2:17. doi: 10.3389/fbloc.2019.00017

1. INTRODUCTION AND SCOPE

The past decade has seen a proliferation of new initiatives to create digital identities for natural persons. Some of these initiatives, such as the ID4D project sponsored by The World Bank (2019) and the Rohingya Project (2019) involve a particular focus in the humanitarian context, while others, such as Evernym (2019) and ID2020 (2019) have a more general scope that includes identity solutions for the developed world. Some projects are specifically concerned with the rights of children (5rights Foundation, 2019). Some projects use biometrics, which raise certain ethical concerns (Pandya, 2019). Some projects seek strong non-transferability, either by linking all credentials related to a particular natural person to a specific identifier, to biometric data, or to each other, as is the case for the anonymous credentials proposed by Camenisch and Lysyanskaya (2001). Some projects have design objectives that include exceptional access ("backdoors") for authorities, which are widely considered to be problematic (Abelson et al., 1997, 2015; Benaloh et al., 2018).

Although this article shall focus on challenges related to identity systems for adult persons in the developed world, we argue that the considerations around data protection and personal data that are applicable in the humanitarian context, such as those elaborated by the International Committee of the Red Cross (Kuner and Marelli, 2017; Stevens et al., 2018), also apply to the

general case. We specifically consider the increasingly commonplace application of identity systems "to facilitate targeting, profiling and surveillance" by "binding us to our recorded characteristics and behaviors" (Privacy International, 2019). Although we focus primarily upon the application of systems for digital credentials to citizens of relatively wealthy societies, we hope that our proposed architecture might contribute to the identity zeitgeist in contexts such as humanitarian aid, disaster relief, refugee migration, and the special interests of children as well.

We argue that while requiring strong non-transferability might be appropriate for some applications, it is inappropriate and dangerous in others. Specifically, we consider the threat posed by mass surveillance of ordinary persons based on their habits, attributes, and transactions in the world. Although the governments of Western democracies might be responsible for some forms of mass surveillance, for example via the recommendations of Financial Action Task Force (2018) or various efforts to monitor Internet activity (Parliament of the United Kingdom, 2016; Parliament of Australia, 2018), the siren song of surveillance capitalism (Zuboff, 2015), including the practice of "entity resolution" through aggregation and data analysis (Waldman et al., 2018), presents a particular risk to human autonomy.

We suggest that many "everyday" activities such as the use of library resources, public transportation services, and mobile data services are included in a category of activities for which strong non-transferability is not necessary and for which there is a genuine need for technology that explicitly protects the legitimate privacy interests of individual persons. We argue that systems that encourage individual persons to establish a single, *unitary*[1] avatar (or "master key") for use in many contexts can ultimately influence and constrain how such persons behave, and we suggest that if a link between two attributes or transactions can be proven, then it can be forcibly discovered. We therefore argue that support for multiple, unlinkable identities is an essential right and a necessity for the development of a future digital society for humans.

This rest of this article is organized as follows. In the next section section, we offer some background on identity systems; we frame the problem space and provide examples of existing solutions. In section 3, we introduce a set of constraints that serve as properties that a digital identity infrastructure must have to support human rights. In section 4, we describe how a digital identity system with a fixed set of actors might operate and how it might be improved. In section 5, we introduce distributed ledger technology to promote a competitive marketplace for issuers and verifiers of credentials and to constrain the interaction between participants in a way that protects the privacy of individual users. In section 6, we consider how the system should be operated and maintained if it is to satisfy the human rights requirements. In section 7, we suggest some potential use cases, and in section 8 we conclude.

[1]In the context of personal identity, use the term *unitary* to refer to attributes, transactions, or identifiers for which an individual can have at most one and that are, for practical purposes, inseparably bound to their subject.

2. BACKGROUND

Establishing meaningful credentials for individuals and organizations in an environment in which the various authorities are not uniformly trustworthy presents a problem for currently deployed services, which are often based on hierarchical trust networks, all-purpose identity cards, and other artifacts of the surveillance economy. In the context of interactions between natural persons, identities are neither universal nor hierarchical, and a top-down approach to identity generally assumes that it is possible to impose a universal hierarchy. Consider "Zooko's triangle," which states that names can be distributed, secure, or human-readable, but not all three (Wilcox-O'Hearn, 2018). The stage names of artists may be distributed and human-readable but are not really secure since they rely upon trusted authorities to resolve conflicts. The names that an individual assigns to friends or that a small community assigns to its members ("petnames," Stiegler, 2005) are secure and human-readable but not distributed. We extend the reasoning behind the paradox to the problem of identity itself and assert that the search for unitary identities for individual persons is problematic. It is technically problematic because there is no endogenous way to ensure that an individual has only one self-certifying name (Douceur, 2002), there is no way to be sure about the trustworthiness or universality of an assigned name, and there is no way to ensure that an individual exists only within one specific community. More importantly, we assert that the ability to manage one's identities in a multitude of different contexts, including the creation of multiple unrelated identities, is an essential human right.

2.1. Manufacturing Trust

The current state-of-the-art identity systems, from technology platforms to bank cards, impose asymmetric trust relationships and contracts of adhesion on their users, including both the ultimate users as well as local authorities, businesses, cooperatives, and community groups. Such trust relationships, often take the form of a hierarchical trust infrastructure, requiring that users accept either a particular set of trusted certification authorities ("trust anchors") or identity cards with private keys generated by a trusted third party. In such cases, the systems are susceptible to socially destructive business practices, corrupt or unscrupulous operators, poor security practices, or control points that risk coercion by politically or economically powerful actors. Ultimately, the problem lies in the dubious assumption that some particular party or set of parties are universally considered trustworthy.

Often, asymmetric trust relationships set the stage for security breaches. Rogue certification authorities constitute a well-known risk, even to sophisticated government actors (Charette, 2016; Vanderburg, 2018), and forged signatures have been responsible for a range of cyber-attacks including the Stuxnet worm, an alleged cyber-weapon believed to have caused damage to Iran's nuclear programme (Kushner, 2013), as

well as a potential response to Stuxnet by the government of Iran (Eckersley, 2011). Corporations that operate the largest trust anchors have proven to be vulnerable. Forged credentials were responsible for the Symantec data breach (Goodin, 2017a), and other popular trust anchors such as Equifax are not immune to security breaches (Equifax Inc, 2018). Google has published a list of certification authorities that it thinks are untrustworthy (Chirgwin, 2016), and IT administrators have at times undermined the trust model that relies upon root certification authorities (Slashdot, 2014). Finally, even if their systems are secure and their operators are upstanding, trust anchors are only as secure as their ability to resist coercion, and they are sometimes misappropriated by governments (Bright, 2010).

Such problems are global, affecting the developed world and emerging economies alike. Identity systems that rely upon a single technology, a single implementation, or a single set of operators have proven unreliable (Goodin, 2017b,c; Moon, 2017). Widely-acclaimed national identity systems, including but not limited to the Estonian identity card system based on X-Road (Thevoz, 2016) and Aadhaar in India (Tully, 2017), are characterized by centralized control points, security risks, and surveillance.

Recent trends in technology and consumer services suggest that concerns about mobility and scalability will lead to the deployment of systems for identity management that identify consumers across a variety of different services, with a new marketplace for providers of identification services Wagner (2014). In general, the reuse of credentials has important privacy implications as a consumer's activities may be tracked across multiple services or multiple uses of the same service. For this reason, the potential for a system to collect and aggregate transaction data must be evaluated whilst evaluating its impact on the privacy of its users.

While data analytics are becoming increasingly effective in identifying and linking the digital trails of individual persons, it has become correspondingly necessary to defend the privacy of individual users and implement instruments that allow and facilitate anonymous access to services. This reality was recognized by the government of the United Kingdom in the design of its GOV.UK Verify programme (Government Digital Service, 2018), a federated network of identity providers and services. However, the system as deployed has significant technical shortcomings with the potential to jeopardize the privacy of its users (Brandao et al., 2015; Whitley, 2018), including a central hub and vulnerabilities that can be exploited to link individuals with the services they use (O'Hara et al., 2011).

Unfortunately, not only do many of the recently-designed systems furnish or reveal data about their users against their interests, but they have been explicitly designed to do so. For example, consider digital rights management systems that force users to identify themselves *ex ante* and then use digital watermarks to reveal their identities (Thomas, 2009). In some cases, demonstrable privacy has been considered an undesirable feature and designs that protect the user's identity intrinsically are explicitly excluded, for example in the case of vehicular *ad-hoc* networks (Shuhaimi and Juhana, 2012), with the implication

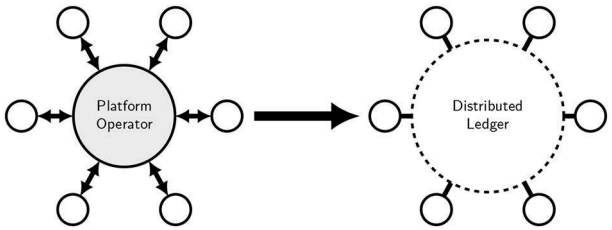


FIGURE 1 | Many network services are centralized in the sense that participants rely upon a specific platform operator to make use of the service **(Left)**, whereas distributed ledgers rely upon network consensus among participants instead of platform operators **(Right)**.

that systems without exceptional access features are dangerous. Finally, of particular concern are systems that rely upon biometrics for identification. By binding identification to a characteristic that users (and in most cases even governments) cannot change, biometrics implicitly prevent a user from transacting within a system without connecting each transaction to each other and potentially to a permanent record. In recent years, a variety of US patents have been filed and granted for general-purpose identity systems that rely upon biometric data to create a "root" identity linking all transactions in this manner (Liu et al., 2008; Thackston, 2018).

2.2. Approaches Using Distributed Ledgers

The prevailing identity systems commonly require users to accept third parties as trustworthy. The alternative to imposing new trust relationships is to work with existing trust relationships by allowing users, businesses, and communities to deploy technology on their own terms, independently of external service providers. In this section we identify various groups that have adopted a system-level approach to allow existing institutions and service providers to retain their relative authority and decision-making power without forcibly requiring them to cooperate with central authorities (such as governments and institutions), service providers (such as system operators), or the implementors of core technology. We suggest that ideally, a solution would not require existing institutions and service providers to operate their own infrastructure without relying upon a platform operator, while concordantly allowing groups such as governments and consultants to act as advisors, regulators, and auditors, not operators. A distributed ledger can serve this purpose by acting as a neutral conduit among its participants, subject to governance limitations to ensuring neutrality and design limitations around services beyond the operation of the ledger that are required by participants. **Figure 1** offers an illustration.

Modern identity systems are used to coordinate three activities: identification, authentication, and authorization. The central problem to address is how to manage those functions in a decentralized context with no universally trusted authorities. Rather than trying to force all participants to use a specific new technology or platform, we suggest using a multi-stakeholder process to develop common standards that define a set of

rules for interaction. Any organization would be able to develop and use their own systems that would interoperate with those developed by any other organization without seeking permission from any particular authority or agreeing to deploy any particular technology.

A variety of practitioners have recently proposed using a distributed ledger to decentralize the administration of an identity system (Dunphy and Petitcolas, 2018), and we agree that the properties of distributed ledger technologies are appropriate for the task. In particular, distributed ledgers allow their participants to share control of the system. They also provide a common view of transactions ensuring that everyone sees the same transaction history.

Various groups have argued that distributed ledgers might be used to mitigate the risk that one powerful, central actor might seize control under the mantle of operational efficiency. However, it is less clear that this lofty goal is achieved in practice. Existing examples of DLT-enabled identity management systems backed by organizations include the following, among others:

- *ShoCard* SITA (2016) is operated by a commercial entity that serves as a trusted intermediary (Dunphy and Petitcolas, 2018).
- *Everest* Everest (2019) is designed as a payment solution backed by biometric identity for its users. The firm behind Everest manages the biometric data and implicitly requires natural persons to have at most one identity within the system (Graglia et al., 2018).
- Evernym (2019) relies on a foundation (Tobin and Reed, 2016) to manage the set of approved certification authorities (Aitken, 2018), and whether the foundation could manage the authorities with equanimity remains to be tested.
- ID2020 (2019) offers portable identity using biometrics to achieve strong non-transferability and persistence (ID2020 Alliance, 2019).
- *uPort* Lundkvist et al. (2016) does not rely upon a central authority, instead allowing for mechanisms such as social recovery. However, its design features an optional central registry that might introduce a means of linking together transactions that users would prefer to keep separate (Dunphy and Petitcolas, 2018). The uPort architecture is linked to phone numbers and implicitly discourages individuals from having multiple identities within the system (Graglia et al., 2018).

Researchers have proposed alternative designs to address some of the concerns. A design suggested by Kaaniche and Laurent does not require a central authority for its blockchain infrastructure but does require a trusted central entity for its key infrastructure (Kaaniche and Laurent, 2017). Coconut, the selective disclosure credential scheme used by Chainspace (2019), is designed to be robust against malicious authorities and may be deployed in a way that addresses such concerns (Sonnino et al., 2018)[2]. We find that many systems such as these require users to bind together their credentials *ex ante*[3] to achieve non-transferability, essentially following a design proposed by Camenisch and Lysyanskaya (2001) that establishes a single "master key" that allows each user to prove that all of her credentials are related to each other. **Figure 2** offers an illustration. Even if users were to have the option to establish multiple independent master keys, service providers or others could undermine that option by requiring proof of the links among their credentials.

The concept of an individual having "multiple identities" is potentially confusing, so let us be clear. In the context of physical documents in the developed world, natural persons generally possess multiple identity documents already, including but not limited to passports, driving licenses, birth certificates, bank cards, insurance cards, and so on. Although individuals might not think of these documents and the attributes they represent as constituting multiple identities, the identity documents generally stand alone for their individual, limited purposes and need not be presented as part of a bundled set with explicit links between the attributes. Service providers might legitimately consider two different identity documents as pertaining to two different individuals, even whilst they might have been issued to the same person. A system that links together multiple attributes via early-binding eliminates this possibility. When we refer to "multiple identities" we refer to records of attributes or transactions that are not linked to each other. Users of identity documents might be willing to sacrifice this aspect of control in favor of convenience, but the potential for blacklisting and surveillance that early-binding introduces is significant. It is for this reason that we take issue with the requirement, advised by various groups including the (International Telecommunications Union, 2018), that individuals must not possess more than one identity. Such a requirement is neither innocuous nor neutral.

Table 1 summarizes the landscape of prevailing digital identity solutions. We imagine that the core technology underpinning these and similar approaches might be adapted to implement a protocol that is broadly compatible with what we describe in this article. However, we suspect that in practice they would need to be modified to encourage users to establish multiple, completely independent identities. In particular, service providers would not be able to assume that users have bound their credentials to each other *ex ante*, and if non-transferability is required, then the system would need to achieve it in a different way.

2.3. Participants in an Identity System

We shall use the following notation to represent the various parties that interact with a typical identity system:

- (1) A *"certification provider"* (**CP**). This would be an entity or organization responsible for establishing a credential based upon foundational data. The credential can be used as a form of identity and generally represents that the organization has

[2]Chainspace was acquired by Facebook in early 2019, and its core technology subsequently became central to the Libra platform (Field, 2019; Heath, 2019).

[3]We use the term *early-binding* to refer to systems that establish provable relationships between transactions, attributes, identifiers, or credentials before they are used. We use the term *late-binding* to refer to systems that allow their users to establish such relationships at the time of use.

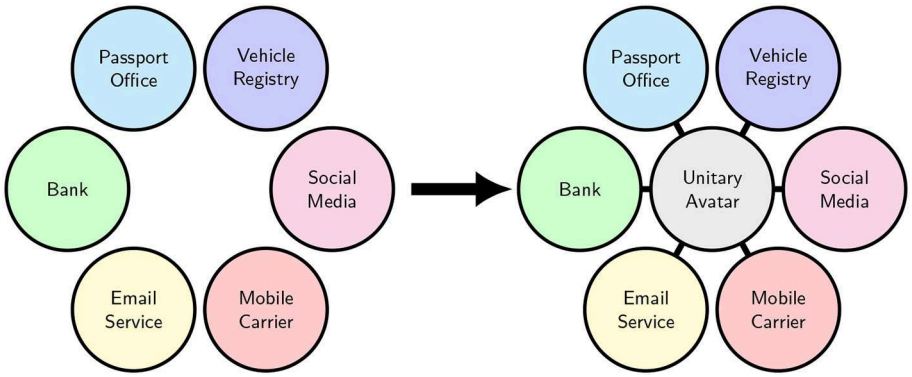


FIGURE 2 | Consider that individual persons possess credentials representing a variety of attributes **(Left)**, and schemes that attempt to achieve strong non-transferability seek to bind these attributes together into a single, unitary "avatar" or "master" identity **(Right)**.

TABLE 1 | A characterization of the landscape of digital identity solutions, with examples.

Name	Objectives	Concerns
Government-operated solutions		
Estonian ID-Card	Interoperability, assurance	Centralized governance, surveillance
Aadhaar (India)	Interoperability, assurance	Centralized governance, surveillance
GOV.UK Verify	Federated management	Central hub, surveillance
Privately-operated solutions		
ShoCard	Strong non-transferability, auditability	Commercial entity is a trusted intermediary, commercial entity stores biometric data
Everest	Strong non-transferability, digital payments	Commercial entity stores biometric data
ID2020	Portability, persistence, strong non-transferability	Identities are unitary through use of biometrics
Evernym	Federated management	Private foundation has an operational role
Kaaniche/Laurent	Hierarchical management	Requires an agreed-upon hierarchy with trusted authority
Decentralized architectures		
uPort	Federated governance and management	Identities become unitary through early-binding or similar mechanisms
Chainspace	Federated governance and management	Identities become unitary through early-binding or similar mechanisms

checked the personal identity of the user in some way. In the context of digital payments, this might be a bank.

- (2) An *"authentication provider"* (**AP**). This would be any entity or organization that might be trusted to verify that a credential is valid and has not been revoked. In current systems, this function is typically performed by a platform or network, for example a payment network such as those associated with credit cards.
- (3) An *"end-user service provider"* (**Service**). This would be a service that requires a user to provide credentials. It might be a merchant selling a product, a government service, or some other kind of gatekeeper, for example a club or online forum.
- (4) A *user* (**user**). This would be a human operator, in most cases aided by a device or a machine, whether acting independently or representing an organization or business.

As an example of how this might work, suppose that a user wants to make an appointment with a local consular office. The consular office wants to know that a user is domiciled in a particular region. The user has a bank account with a bank that is willing to certify that the user is domiciled in that region. In addition, a well-known authentication provider is willing to accept certifications from the bank, and the consular office accepts signed statements from that authentication provider. Thus, the user can first ask the bank to sign a statement certifying that he is domiciled in the region in question. When the consular office asks for proof of domicile, the user can present the signed statement from the bank to the authentication provider and ask the authentication provider to sign a new statement vouching for the user's region of domicile, using information from the bank as a basis for the statement, without providing any information related to the bank to the consular office.

3. DESIGN CONSTRAINTS FOR PRIVACY AS A HUMAN RIGHT

Reflecting on the various identity systems used today, including but not limited to residence permits, bank accounts, payment cards, transit passes, and online platform logins, we observed a plethora of features with weaknesses and vulnerabilities concerning privacy (and in some cases security) that could

potentially infringe upon human rights. Although the 1948 Universal Declaration on Human Rights explicitly recognizes privacy as a human right (United Nations, 1948), the declaration was drafted well before the advent of a broad recognition of the specific dangers posed by the widespread use of computers for data aggregation and analysis (Armer, 1975), to say nothing of surveillance capitalism (Zuboff, 2015). Our argument that privacy in the context of digital identity is a human right, therefore, rests upon a more recent consideration of the human rights impact of the abuse of economic information (European Parliament, 1999). With this in mind, we identified the following eight fundamental constraints to frame our design requirements for technology infrastructure (Goodell and Aste, 2018):

Structural requirements:

1. *Minimize **control points** that can be used to co-opt the system.* A single point of trust is a single point of failure, and both state actors and technology firms have historically been proven to abuse such trust.
2. *Resist establishing potentially abusive **processes and practices**, including legal processes, that rely upon control points.* Infrastructure that can be used to abuse and control individual persons is problematic even if those who oversee its establishment are genuinely benign. Once infrastructure is created, it may in the future be used for other purposes that benefit its operators.

Human requirements:

3. *Mitigate architectural characteristics that lead to **mass surveillance** of individual persons.* Mass surveillance is about control as much as it is about discovery: people behave differently when they believe that their activities are being monitored or evaluated (Mayo, 1945). Powerful actors sometimes employ monitoring to create incentives for individual persons, for example to conduct marketing promotions or credit scoring operations. Such incentives may prevent individuals from acting autonomously, and the chance to discover irregularities, patterns, or even misbehavior often does not justify such mechanisms of control.
4. *Do not impose **non-consensual** trust relationships upon beneficiaries.* It is an act of coercion for a service provider to require a client to maintain a direct trust relationship with a specific third-party platform provider or certification authority. Infrastructure providers must not explicitly or implicitly engage in such coercion, which should be recognized for what it is and not tolerated in the name of convenience.
5. *Empower individual users to manage the **linkages** among their activities.* To be truly free and autonomous, individuals must be able to manage the cross sections of their activities, attributes, and transactions that are seen or might be discovered by various institutions, businesses, and state actors.

Economic requirements:

6. *Prevent solution providers from establishing a **monopoly position**.* Some business models are justified by the opportunity to achieve status as monopoly infrastructure. Monopoly infrastructure is problematic not only because it deprives its users of consumer surplus but also because it empowers the operator to dictate the terms by which the infrastructure can be used.
7. *Empower local businesses and cooperatives to establish their own **trust relationships**.* The opportunity to establish trust relationships on their own terms is important both for businesses to compete in a free marketplace and for businesses to act in a manner that reflects the interests of their communities.
8. *Empower service providers to establish their own **business practices and methods**.* Providers of key services must adopt practices that work within the values and context of their communities.

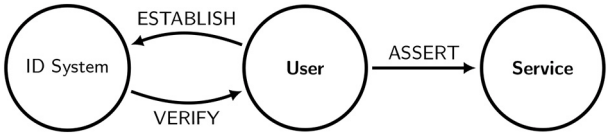


FIGURE 3 | A schematic representation of a generalized identity system. Users first establish a credential with the system, then use the system to verify the credential, and then use the verified identity to assert that they are authorized to receive a service.

These constraints constitute a set of *system-level* requirements, involving human actors, technology, and their interaction, not to be confused with the *technical* requirements that have been characterized as essential to self-sovereign identity (SSI) (Stevens et al., 2018). Although our design objectives may overlap with the design objectives for SSI systems, we seek to focus on system-level outcomes. While policy changes at the government level might be needed to fully achieve the vision suggested by some of the requirements, we would hope that a digital identity system would not contain features that intrinsically facilitate their violation.

Experience shows that control points will eventually be co-opted by powerful parties, irrespective of the intentions of those who build, own, or operate the control points. Consider, for example, how Cambridge Analytica allegedly abused the data assets of Facebook Inc to manipulate voters in Britain and the US (Koslowska et al., 2018) and how the Russian government asserted its influence on global businesses that engaged in domain-fronting (Lunden, 2018; Savov, 2018). The inherent risk that centrally aggregated datasets may be abused, not only by the parties doing the aggregating but also by third parties, implies value in system design that avoids control points and trusted infrastructure operators, particularly when personal data and livelihoods are involved.

4. A DIGITAL IDENTITY SYSTEM

Various digital identity architectures and deployments exist today to perform the three distinct functions we mentioned earlier: identification, authentication, and authorization (Riley, 2006). We introduce a fourth function, *auditing*, by which the basis for

judgements made by the system can be explained and evaluated. We characterize the four functions as follows:

- IDENTIFICATION. A user first *establishes* some kind of credential or identifier. The credential might be a simple registration, for example with an authority, institution, or other organization. In other cases, it might imply a particular attribute. The implication might be implicit, as a passport might imply citizenship of a particular country or a credential issued by a bank might imply a banking relationship, or it might be explicit, as in the style of attribute-backed credentials (Camenisch and Lysyanskaya, 2003; IBM Research Zurich, 2018)[4].
- AUTHENTICATION. Next, when the provider of a service seeks to authenticate a user, the user must be able to *verify* that a credential in question is valid.
- AUTHORIZATION. Finally, the user can use the authenticated credential to *assert* to the service provider that she is entitled to a particular service.
- AUDITING. The identity system would maintain record of the establishment, expiration, and revocation of credentials such that the success or failure of any given authentication request can be explained.

Ultimately, it is the governance of a digital identity system, including its intrinsic policies and mechanisms as well as the accountability of the individuals and groups who control its operation, that determines whether it empowers or enslaves its users. We suggest that proper governance, specifically including a unified approach to the technologies and policies that the system comprises, is essential to avoiding unintended consequences to its implementation. We address some of these issues further in section 6.

Figure 3 gives a pictorial representation of the functions. **Table 2** defines the notation that we shall use in our figures. With the constraints enumerated in section 3 taken as design requirements, we propose a generalized architecture that achieves our objectives for an identity system. The candidate systems identified in section 2 can be evaluated by comparing their features to our architecture. Since we intend to argue for a practical solution, we start with a system currently enjoying widespread deployment.

4.1. SecureKey Concierge

As a baseline example of an identity framework, we consider a system that uses banks as certification providers whilst circumventing the global payment networks. *SecureKey Concierge* (SKC) (SecureKey Technologies, Inc, 2015) is a solution used by the government of Canada to provide users with access to its various systems in a standard way. The SKC architecture seeks the following benefits:

1. Leverage existing "certification providers" such as banks and other financial institutions with well-established, institutional procedures for ascertaining the identities of their customers. Often such procedures are buttressed by legal frameworks such as Anti-Money Laundering (AML) and "Know Your Customer" (KYC) regulations that broadly deputize banks and substantially all financial institutions (GOV.UK, 2014) to collect identifying information on the various parties that make use of their services, establish the expected pattern for the transactions that will take place over time, and monitor the transactions for anomalous activity inconsistent with the expectations (Better Business Finance, 2017).
2. Isolate service providers from personally identifying bank details and eliminate the need to share specific service-related details with the certification provider, whilst avoiding traditional authentication service providers such as payment networks.

Figure 4 offers a stylized representation of the SKC architecture, as interpreted from its online documentation (SecureKey Technologies, Inc, 2015). When a user wants to access a

[4]We do not describe how to use attribute-backed credentials here.

TABLE 2 | Notation used in the subsequent figures.

Request	A request for a credential.
Request *x*	A request for a credential with a parameter, *x*.
A(m)	A message *m* signed by party A.
Identify	The foundational identifying elements that a user presents to a certification provider, encrypted so that other parties (including AP) cannot read them.
Revoke-one	A message invalidating an earlier signature on a specific user credential.
Revoke-all	A message invalidating all signatures by a certain key.
[m]	Blinded version of message *m*.
A([m])	Blind signature of *m* by A.
Prove-owner *x**	Proof of ownership of some private key *x**, for example via a challenge-response (which would imply two extra messages not shown) or by using the key to sign a pre-existing secret created by the recipient and shared with the sender.
Request-certs A	A request for all of the certificates on the ledger signed by A, followed by a response containing all of the matching certificates.
Receipt	Response from the distributed ledger system indicating that a transaction completed successfully.
Object	Physical, tamper-resistant object containing a unique receipt for a transaction.

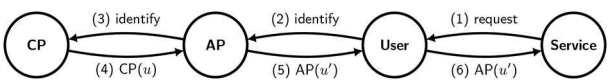


FIGURE 4 | A stylized schematic representation of the SecureKey Concierge (SKC) system. The parties are represented by the symbols "CP," "AP," "User," and "Service," and the arrows represent messages between the parties. The numbers associated with each arrow show the sequence, and the symbol following the number represents the contents of a message. First, the service provider (Service) requests authorization from the user, who in turn sends identifying information to the authentication provider (AP) to share with the certification provider (CP). If the CP accepts the identifying information, it sends a signed credential *u* to the AP, which in turn issues a new credential *u′* for consumption by the Service, which can now authorize the user.

service, the service provider sends a request to the user (1)[5] for credentials. The user then sends encrypted identifying information (for example, bank account login details) to the authentication provider (2), which in this case is SKC, which then forwards it to the certification provider (3). Next, the certification provider responds affirmatively with a "meaningless but unique" identifier u representing the user, and sends it to the authentication provider (4). The authentication provider then responds by signing its own identifier u' representing the user and sending the message to the user (5), which in turn passes it along to the service provider (6). At this point the service provider can accept the user's credentials as valid. The SKC documentation indicates that SKC uses different, unlinked values of u' for each service provider.

4.2. A Two-Phase Approach

We might consider modifying the SKC architecture so that the user does not need to log in to the CP each time it requests a service. To achieve this, we divide the protocol into two phases, as shown in **Figure 5**: a *setup phase* (**Figure 5A**) in which a user establishes credentials with an "certification provider" (CP) for use with the service, and an *operating phase* (**Figure 5B**) in which a user uses the credentials in an authentication process with a service provider. So, the setup phase is done once, and the operating phase is done once per service request. In the setup phase, the user first sends authentication credentials, such as those used to withdraw money from a bank account, to an authentication provider (1). The authentication provider then uses the credentials to authenticate to the certification provider (2), which generates a unique identifier u that can be used for subsequent interactions with service providers and sends it to the authentication provider (3), which forwards it to the user (4). Then, in the operating phase, a service provider requests credentials from the user (5), which in turn uses the previously established unique identifier u to request credentials from the authentication provider (6). This means that the user would implicitly maintain a relationship with the authentication provider, including a way to log in. The authentication provider then verifies that the credentials have not been revoked by the certification provider. The process for verifying that the CP credential is still valid may be offline, via a periodic check, or online, either requiring the AP to reach out to the AP when it intends to revoke a credential or requiring the AP to send a request to the CP in real-time. In the latter case, the AP is looking only for updates on the set of users who have been through the setup phase, and it does not need to identify which user has made a request. Once the AP is satisfied, it sends a signed certification of its identifier u' to the user (7), which forwards it to the service provider as before (8).

Unfortunately, even if we can avoid the need for users to log in to the CP every time they want to use a service, the authentication provider itself serves as a trusted third party. Although the SKC architecture may eliminate the need to trust the existing payment networks, the authentication provider maintains the mapping between service providers and the certification providers used by individuals to request services. It is also implicitly trusted to manage all of the certification tokens, and there is no way to ensure that it does not choose them in a way that discloses information to service providers or certification providers. In particular, users need to trust the authentication provider to use identifiers that do not allow service providers to correlate their activities, and users may also also want to use different identifiers from time to time to communicate with the *same* service provider. As a monopoly platform, it also has the ability to tax or deny service to certification providers, users, or service providers according to its own interests, and it serves as a single point of control vulnerable to exploitation. For all of these reasons, we maintain that the SKC architecture remains problematic from a public interest perspective.

4.3. A User-Oriented Identity Architecture

In the architecture presented in section 4.2, the authentication provider occupies a position of control. In networked systems, control points confer economic advantages on those who occupy them (Value Chain Dynamics Working Group (VCDWG), 2005), and the business incentives associated with the opportunity to build platform businesses around control points have been used to justify their continued proliferation (Ramakrishnan and Selvarajan, 2017).

However, control points also expose consumers to risk, not only because the occupier of the control point may abuse its position but also because the control point itself creates a vector for attack by third parties. For both of these reasons, we seek to prevent an authentication provider from holding too much information about users. In particular, we do not want an authentication provider to maintain a mapping between a user and the particular services that a user requests, and we do not want a single authentication provider to establish a monopoly position in which it can dictate the terms by which users and service providers interact. For this reason, we put the user, and not the authentication provider, in the center of the architecture.

4.4. Isolation Objectives

For a user to be certain that she is not providing a channel by which authentication providers can leak her identity or by which service providers can trace her activity, then she must *isolate* the different participants in the system. The constraints allow us to define three *isolation objectives* as follows:

1. *Have users generate unlinked identifiers on devices that they own and trust*. Unless they generate the identifiers themselves, users have no way of knowing for sure whether identifiers assigned to them do not contain personally identifying information. For users to verify that the identifiers will not disclose information that might identify them later, they would need to generate random identifiers using devices and software that they control and trust. We suggest that for a user to trust a device, its hardware and software must be of an open-source, auditable design with auditable provenance. Although we would not expect that most users would be able to judge the security properties of the devices they use,

[5]The numbers in italics correspond to messages in the figure indicated, in this case (**Figure 4**).

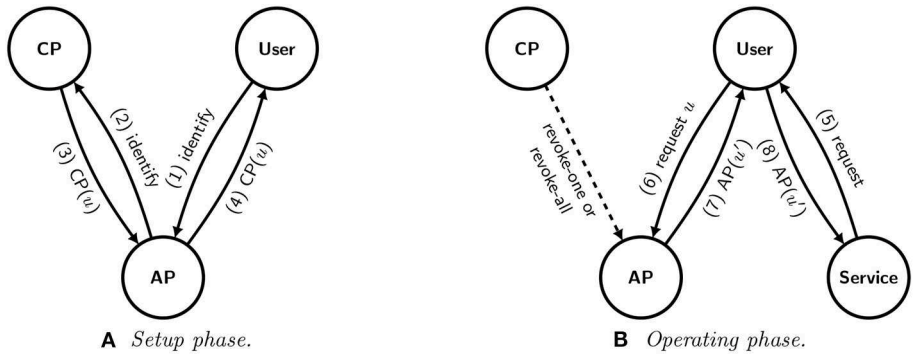


FIGURE 5 | A schematic representation of a modified version of the SKC system with a stateful authentication provider and one-time identifiers for services. The user first establishes credentials in the setup phase **(A)**. Then, when a service provider requests credentials from the user in the operating phase **(B)**, the user reaches out to the authentication provider for verification, which assigns a different identifier u' each time.

open-source communities routinely provide mechanisms by which users without specialized knowledge can legitimately conclude, before using new hardware or software, that a diverse community of experts have considered and approved the security aspects of the technology. Examples of such communities include Debian (Software in the Public Interest, Inc, 2019) for software and (Arduino, 2019) for hardware, and trustworthy access to these communities might be offered by local organizations such as libraries or universities.

2. *Ensure that authentication providers do not learn about the user's identity or use of services*. Authentication providers that require foundational information about a user, or are able to associate different requests over time with the same user, are in a position to collect information beyond what is strictly needed for the purpose of the operation. The role of the authentication provider is to act as a neutral channel that confers authority on certification providers, time-shifts requests for credentials, and separates the certification providers from providers of services. Performing this function does not require it to collect information about individual users at any point.
3. *Ensure that information given to service providers is not shared with authentication providers*. The user must be able to credibly trust that his or her interaction with service providers must remain private.

The communication among the four parties that we propose can be done via simple, synchronous (e.g., HTTP) protocols that are easily performed by smartphones and other mobile devices. The cryptography handling public keys can be done using standard public-key-based technologies.

4.5. Repositioning the User to Be in the Center

Figure 6 shows how we can modify the system shown in **Figure 5** to achieve the three isolation objectives defined above. Here, we introduce blind signatures (Chaum, 1983) to allow the user to present a verified signature without allowing the signer and the relying party to link the identity of the subject to the subject's legitimate use of a service.

Figure 6A depicts the new setup phase. First, on her own trusted hardware (see section 4.4), the user generates her own set of identifiers $x_1, ..., x_n$ that she intends to use, at most once each, in future correspondence with the authentication provider. Generating the identifiers is not computationally demanding and can be done with an ordinary smartphone. By generating her own identifiers, the user has better control that nothing encoded in the identifiers that might reduce her anonymity. The user then sends both its identifying information and the identifiers $x_1, ..., x_n$ to the certification provider (1). The certification provider then responds with a set of signatures corresponding to each of the identifiers (2). The user then sends the set of signatures to the authentication provider for future use (3).

Figure 6B depicts the new operating phase. First, the service sends a request to the user along with a new *nonce* (one-time identifier) y corresponding to the request (4). The user then applies a blinding function to the nonce y, creating a blinded nonce $[y]$. The user chooses one of the identifiers x_i that she had generated during the setup phase and sends that identifier along with the blinded nonce $[y]$ to the authentication provider (5). Provided that the signature on x_i has not been revoked, the authentication provider confirms that it is valid by signing $[y]$ and sending the signature to the user (6). The user in turn "unblinds" the signature on y and sends the unblinded signature to the service provider (7). The use of blind signatures ensures that the authentication provider cannot link what it sees to specific interactions between the user and the service provider.

4.6. Architectural Considerations

To satisfy the constraints listed in section 3, all three process steps (identification, authentication, and authorization) must be isolated from each other. Although our proposed architecture introduces additional interaction and computation, we assert that the complexity of the proposed architecture is both parsimonious and justified:

1. If the certification provider were the same as the service provider, then the user would be subject to direct control and surveillance by that organization, violating Constraints 1, 3, and 5.

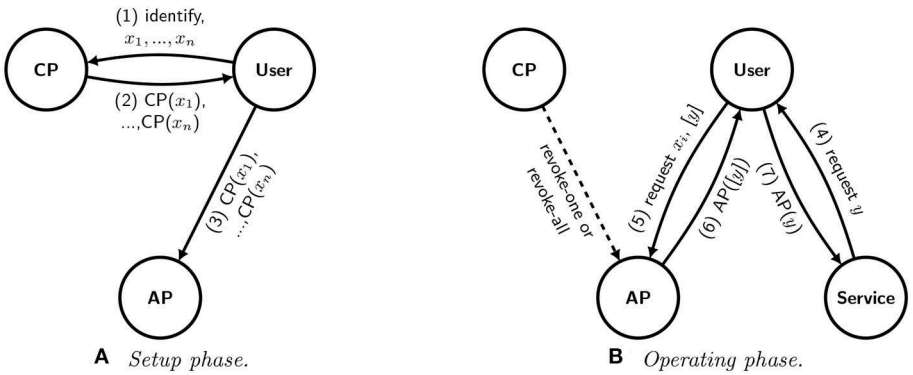


FIGURE 6 | A schematic representation of a digital identity system with a user-oriented approach. The new protocol uses user-generated identifiers and blind signatures to isolate the authentication provider. The authentication provider cannot inject identifying information into the identifiers, nor can it associate the user with the services that she requests. **(A)** Setup phase. **(B)** Operating phase.

2. If the authentication provider were the same as the certification provider, then the user would have no choice but to return to the same organization each time it requests a service, violating Constraints 1 and 4. That organization would then be positioned to discern patterns in its activity, violating Constraints 3 and 5. There would be no separate authentication provider to face competition for its services as distinct from the certification services, violating Constraint 6.
3. If the authentication provider were the same as the service provider, then the service provider would be positioned to compel the user to use a particular certification provider, violating Constraints 1 and 4. The service provider could also impose constraints upon what a certification provider might reveal about an individual, violating Constraint 3, or how the certification provider establishes the identity of individuals, violating Constraint 8.
4. If the user could not generate her own identifiers, then the certification provider could generate identifiers that reveal information about the user, violating Constraint 3.
5. If the user were not to use blind signatures to protect the requests from service providers, then service providers and authentication providers could compare notes to discern patterns of a user's activity, violating Constraint 5.

The proposed architecture does not achieve its objectives if either the certification provider or the service provider colludes with the authentication provider; we assume that effective institutional policy will complement appropriate technology to ensure that sensitive data are not shared in a manner that would compromise the interests of the user.

5. A DECENTRALIZED IDENTITY ARCHITECTURE

A significant problem remains with the design described in section 4.5 in that it requires $O(n^2)$ relationships among authentication providers and certification providers (i.e., with each authentication provider connected directly to each certification provider that it considers valid) to be truly decentralized. Recall that the system relies critically upon the ability of an certification provider to *revoke* credentials issued to users, and authentication providers need a way to learn from the certification provider whether a credential in question has been revoked. Online registries such as OCSP (Juniper Networks, 2018), which are operated by a certification provider or trusted authority, are a common way to address this problem, although the need for third-party trust violates Constraint 1. The issue associated with requiring each authentication provider to establish its own judgment of each candidate certification provider is a *business* problem rather than a technical one. Hierarchical trust relationships emerge because relationships are expensive to maintain and introduce risks; all else being equal, business owners prefer to have fewer of them. Considered in this context, concentration and lack of competition among authentication providers makes sense. If one or a small number of authentication providers have already established relationships with a broad set of certification providers, just as payment networks such as Visa and Mastercard have done with a broad set of banks, then the cost to a certification provider of a relationship with a new authentication provider would become a barrier of entry to new authentication providers. The market for authentication could fall under the control of a monopoly or cartel.

5.1. Introducing Distributed Ledger Technology

We propose using *distributed ledger technology* (DLT) to allow both certification providers and authentication providers to proliferate whilst avoiding industry concentration. The distributed ledger would serve as a standard way for certification providers to establish relationships with any or all of the authentication providers at once, or vice-versa. The ledger itself would be a mechanism for distributing signatures and revocations; it would be shared by participants and not controlled by any single party. **Figure 7** shows that users would not interact with the distributed ledger directly but via their chosen certification providers and authentication providers.

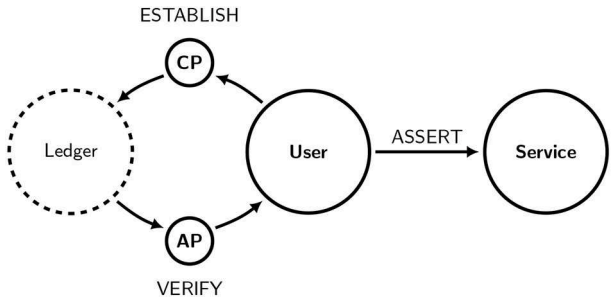


FIGURE 7 | A schematic representation of a decentralized identity system with a distributed ledger. The user is not required to interact directly with the distributed ledger (represented by a dashed circle) and can rely upon the services offered by certification providers and authentication providers.

Additionally, users would not be bound to use any particular authentication provider when verifying a particular credential and could even use a different authentication provider each time. Provided that the community of participants in the distributed ledger remains sufficiently diverse, the locus of control would not be concentrated within any particular group or context, and the market for authentication can remain competitive.

Because the distributed ledger architecture inherently does not require each new certification provider to establish relationships with all relevant authentication providers, or vice-versa, it facilitates the entry of new authentication providers and certification providers, thus allowing the possibility of decentralization.

We argue that a distributed ledger is an appropriate technology to maintain the authoritative record of which credentials have been issued (or revoked) and which transactions have taken place. We do not trust any specific third party to manage the list of official records, and we would need the system to be robust in the event that a substantial fraction of its constituent parts are compromised. The distributed ledger can potentially take many forms, including but not limited to blockchain, and, although a variety of fault-tolerant consensus algorithms may be appropriate, we assume that the set of node operators is well-known, a characteristic that we believe might be needed to ensure appropriate governance.

If implemented correctly at the system level, the use of a distributed ledger can ensure that the communication between the certification provider and the authentication provider is limited to that which is written on the ledger. If all blind signatures are done without including any accompanying metadata, and as long as the individual user does not reveal which blind signature on the ledger corresponds to the unblinded signature that he or she is presenting to the authentication provider for approval, then nothing on the ledger will reveal any information about the individual persons who are the subjects of the certificates. We assume that the certification authorities would have a limited and well-known set of public keys that they would use to sign credentials, with each key corresponding to the category of individual persons who have a specific attribute. The size of the anonymity set for aa credential, and therefore the effectiveness of the system in protecting the privacy of an individual user with that credential, depends upon the generality of the category. We would encourage certification authorities to assign categories that are as large as possible. We would also assume that the official set of signing keys used by certification providers and authentication providers is also maintained on the ledger, as doing so would ensure that all users of the system have the same view of which keys the various certification providers and authentication providers are using.

5.2. Achieving Decentralization With a Distributed Ledger

Figure 8 shows how the modified architecture with the distributed ledger technology would work. **Figure 8A** shows the setup phase. The first two messages from the user to the certification provider are similar to their counterparts in the protocol shown in **Figure 6A**. However, now the user also generates n asymmetric key pairs (x_i, x_i^*), where x_i is the public key and x_i^* is the private key of pair i, and it sends each public key $x_1, ..., x_n$ to the certification provider (1). Then, rather than sending the signed messages to the authentication provider via the user, the certification provider then instead writes the signed certificates directly to the distributed ledger (2). Importantly, the certificates would not contain any metadata but only the public key x_i and its bare signature; eliminating metadata is necessary to ensure that there is no channel by which a certification provider might inject information that might be later used to identify a user. **Figure 8B** shows the operating phase, which begins when a service provider asks a user to authenticate and provides some nonce y as part of the request.

The certification provider can revoke certificates simply by transacting on the distributed ledger and without interacting with the authentication provider at all. Because the user and the authentication provider are no longer assumed to mutually trust one another, the user must now prove to the authentication provider that the user holds the private key x_i^* when the user asks the authentication provider to sign the blinded nonce $[y]$ (4). At this point we assume that the authentication provider maintains its own copy of the distributed ledger and has been receiving updates. The authentication provider then refers to its copy of the distributed ledger to determine whether a credential has been revoked, either because the certification provider revoked a single credential or because the certification provider revoked its own signing key. Provided that the credential has not been revoked, the authentication provider signs the blinded nonce $[y]$ (5), which the user then unblinds and sends to the service provider (6). The following messages are carried out as they are done in **Figure 6B**.

We assume that each certification provider and authentication provider has a distinct signing key for credentials representing each possible policy attribute, and we further assume that each possible policy attribute admits for a sufficiently large anonymity set to not identify the user, as described in section 5.1. A policy might consist of the union of a set of attributes, and because users could prevent arbitrary subsets of the attributes to authentication providers and service providers, we believe that in most cases it would not be practical to structure policy

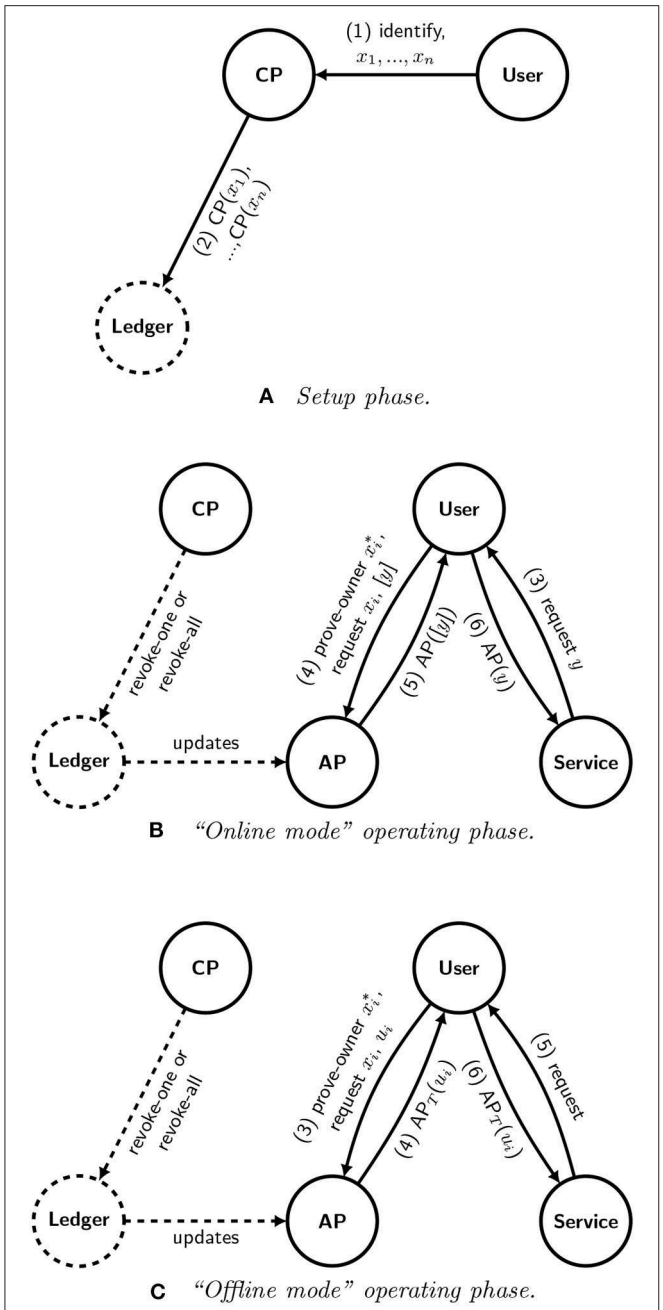


FIGURE 8 | A schematic representation of one possible decentralized digital identity system using a distributed ledger. Diagrams **(A,B)** show the setup and operating phases for an initial sketch of our design, which uses a distributed ledger to promote a scalable marketplace that allows users to choose certification providers and authentication providers that suit their needs. Diagram **(C)** shows a variation of the operating phase that can be used in an offline context, in which the user might not be able to communicate with an up-to-date authentication provider and the service provider at the same time.

attributes in such a manner that one attribute represents a qualification or restriction of another. Additionally, at a system level, authentication providers and service providers must not require a set of attributes, either from the same issuer or from different issuers, whose combination would restrict the anonymity set to an extent that would potentially reveal the identity of the user.

5.3. Operating the System Offline

The proposed approach can also be adapted to work *offline*, specifically when a user does not have access to an Internet-connected authentication provider at the time that it requests a service from a service provider. This situation applies to two cases: first, in which the authentication provider has only intermittent access to its distributed ledger peers (perhaps because the authentication provider has only intermittent access to the Internet), and second, in which the user does not have access to the authentication provider (perhaps because the user does not have access to the Internet) at the time that it requests a service.

In the first case, note that the use of a distributed ledger allows the authentication provider to avoid the need to send a query in real-time[6]. If the authentication provider is disconnected from the network, then it can use its most recent version of the distributed ledger to check for revocation. If the authentication provider is satisfied that the record is sufficiently recent, then it can sign the record with a key that is frequently rotated to indicate its timeliness, which we shall denote by AP_T. We presume that AP_T is irrevocable but valid for a limited time only. If the authentication provider is disconnected from its distributed ledger peers but still connected to the network with the service provider, then it can still sign a nonce from the service provider as usual.

In the second case, however, although the user is disconnected from the network, the service provider still requires an indication of the timeliness of the authentication provider's signature. The generalized solution is to adapt the operating phase of the protocol as illustrated by **Figure 8C**. Here, we assume that the user knows in advance that she intends to request a service at some point in the near future, so she sends the request to the authentication provider pre-emptively, along with a one-time identifier u_i (3). Then, the authentication provider verifies the identifier via the ledger and signs the one-time identifier u_i with the time-specific key AP_T (4). Later, when the service provider requests authorization (5), the user responds with the signed one-time identifier that it had obtained from the authentication provider (6). In this protocol, the service provider also has a new responsibility, which is to keep track of one-time identifiers to ensure that there is no duplication.

5.4. Achieving Unlinkability With Blinded Credentials

Unfortunately, the architecture described in sections 5.2 and 5.3 has an important weakness as a result of its reliance on the revocation of user credentials. Because the credential that an certification provider posts to the ledger is specifically identified by the user at the time that the user asks the authentication provider to verify the credential, the certification provider may

[6]Not sending the query over the network may also improve the privacy of the transaction.

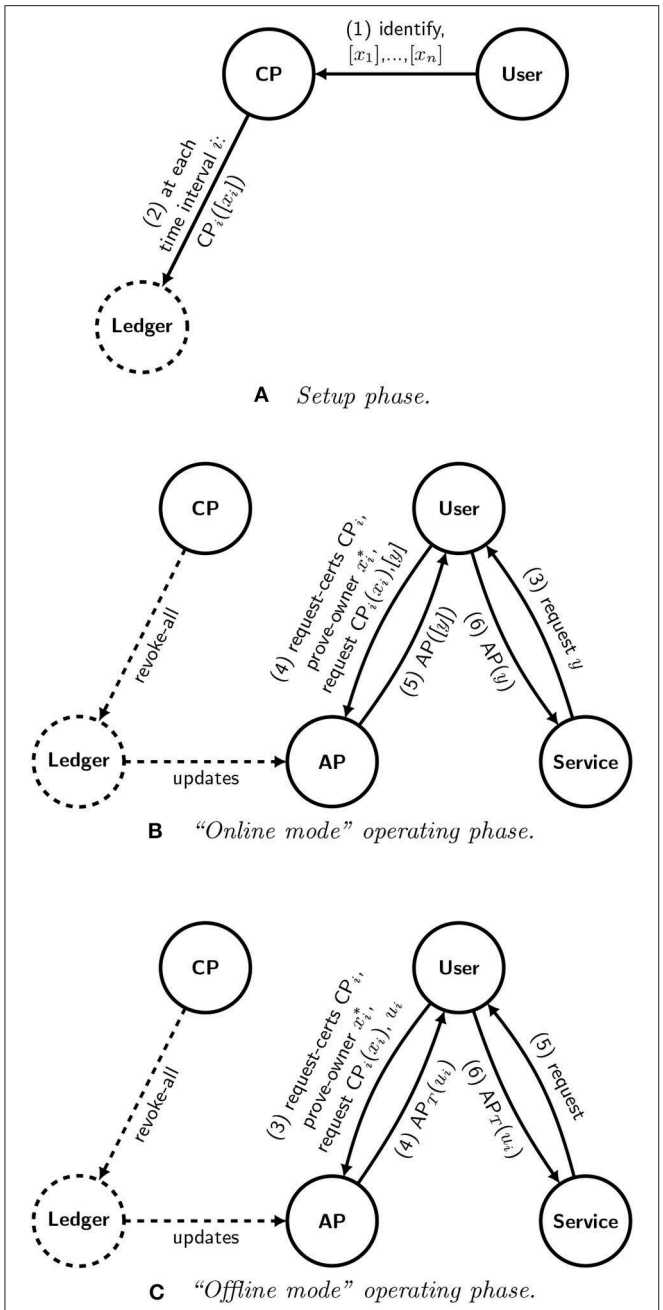


FIGURE 9 | A schematic representation of a metadata-resistant decentralized identity architecture. This version of the design represents our recommendation for a generalized identity architecture. By writing only blinded credentials to the ledger, this version extends the design shown in **Figure 8** to resist an attack in which the certification provider can expose linkages between different credentials associated with the same user. Diagrams **(A,B)** show the setup and operating phases, analogously to the online example shown in **Figure 8**; Diagram **(C)** shows the corresponding offline variant.

collude with individual authentication providers to determine when a user makes such requests. Even within the context of the protocol, an unscrupulous (or compromised, or coerced) certification provider may post revocation messages for all of the credentials associated with a particular user, hence linking them to each other.

For this reason, we recommend modifying to the protocol to prevent this attack by using blinded credentials to improve its metadata-resistance. **Figure 9** shows how this would work. Rather than sending the public keys x_i directly to the certification provider, the user sends blinded public keys $[x_i]$, one for each of a series of specific, agreed-upon time intervals (1). which in turn would be signed by the certification provider using a blind signature scheme that does not allow revocation (1). The certification provider would not sign all of the public keys and publish the certificates to the ledger immediately; instead, it would sign them and post the certificates to the ledger at the start of each time interval *i*, in each instance signing the user keys with a key of its own specific to that time interval, CP_i (2). If a user expects to make multiple transactions per time interval and desires those transactions to remain unlinked from each other, the user may send multiple keys for each interval.

When the time comes for the user to request a service, the user must demonstrate that it is the owner of the private key corresponding to a (blinded) public key that had been signed by the certification provider. So the user must first obtain the set of all certificates signed by CP_i, which it can obtain from the authentication provider via a specific request, request-certs. Then it can find the blind signature on $[x_i]$ from the list and unblind the signature to reveal $\mathrm{CP}_i(x_i)$. It can then send this signature to the authentication provider along with its proof of ownership of x_i^* as before.

This version of the protocol is the one that we recommend for most purposes. Although the request-certs exchange might require the user to download a potentially large number of certificates, such a requirement would hopefully indicate a large anonymity set. In addition, there may be ways to mitigate the burden associated by the volume of certificates loaded by the client. For example, we might assume that the service provider offers a high-bandwidth internet connection that allows the user to request the certificates anonymously from an authentication provider. Alternatively, we might consider having the certification provider subdivide the anonymity set into smaller sets using multiple well-known public keys rather than a single CP_i, or we might consider allowing an interactive protocol between the user and the authentication provider in which the user voluntarily opts to reduce her anonymity set, for example by specifying a small fraction of the bits in $[x_i]$ as a way to request only a subset of the certificates.

5.5. Adapting the Design for Spending Tokens

The architecture defined in section 5.4 can also be adapted to allow users to spend tokens on a one-time basis. This option may be of particular interest for social and humanitarian services, in which tokens to be used to purchase food, medicine, or essential services may be issued by a government authority or aid organization to a community at large. In such cases, the human rights constraints are particularly important. **Figure 10** shows how a certification provider might work with an issuer of tokens

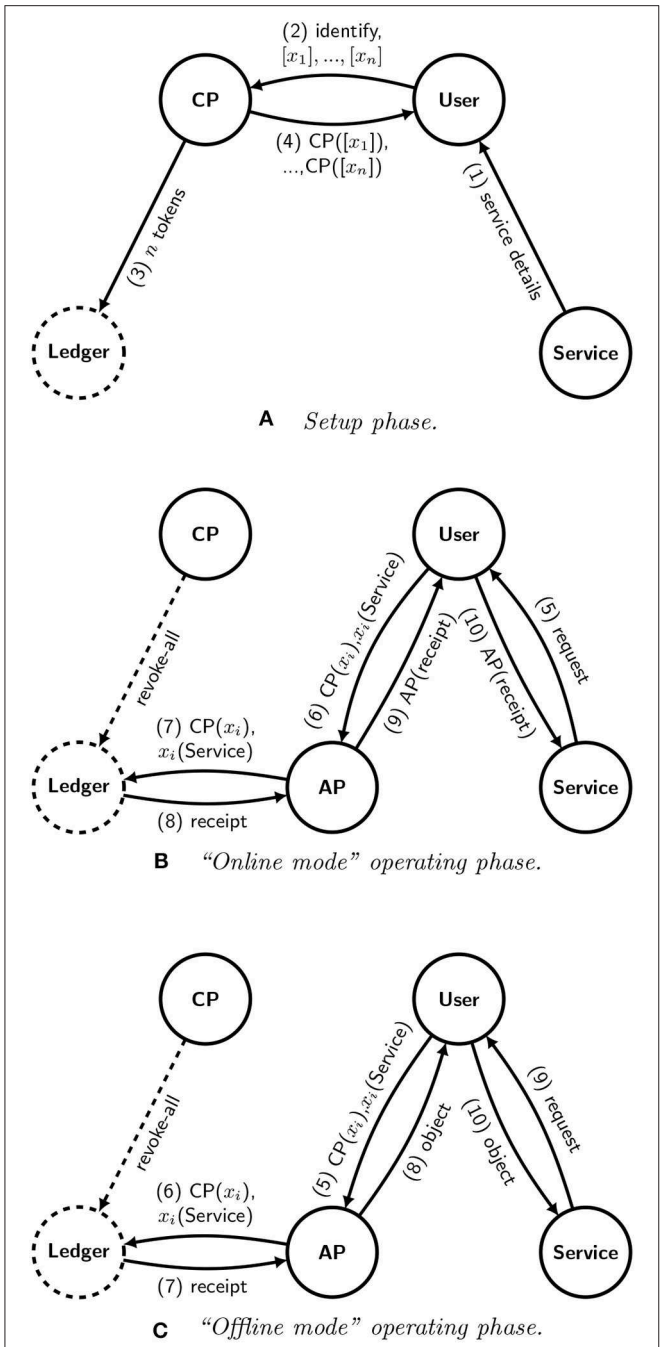


FIGURE 10 | A schematic representation of a decentralized identity architecture for exchanging tokens. The protocol represented in this figure can be used to allow users to pay for services privately. Diagrams **(A,B)** show the setup and operating phases, analogously to the online example shown in **Figures 8**, **9**; Diagram **(C)** shows the corresponding offline variant.

to issue spendable tokens to users, who may in turn represent themselves or organizations.

Figure 10A shows the setup phase. We assume that the service provider first tells the user that it will accept tokens issued by the certification provider (1). The user then sends a set of *n* newly-generated public keys, along with any required identity or credential information, to the certification provider (2). We assume that the tokens are intended to be fungible, so the certification provider issues *n* new, fungible tokens on the ledger (3). We need not specify the details of how the second message works in practice; depending upon the trust model for the ledger it may be as simple as signing a statement incrementing the value associated with the certification provider's account, or it may be a request to move tokens explicitly from one account to another. Then, the certification provider signs a set of *n* messages, each containing one of the blinded public keys, and sends them to the user (4). The messages will function as promissory notes, redeemable by the user that generated the keys, for control over the tokens.

Figure 10B shows the operating phase. When a service provider requests a token (5), the user sends a message to an authentication provider demonstrating both that it has the right to control the token issued by the certification provider and that it wishes to sign the token over to the service provider (6). The authentication provider, who never learns identifying information about the user, lodges a transaction on the ledger that assigns the rights associated with the token to the service provider (7), generating a receipt (8). Once the transaction is complete, the authentication provider shares a receipt with the user (9), which the user may then share with the service provider (10), who may now accept that the payment as complete.

Like the "main" architecture described in section 5.4, the "token" architecture can also be configured to work in an offline context by modifying the operating phase. **Figure 10C** shows how this would work. The user requests from the authentication provider one or more physical "objects", which may take the form of non-transferable electronic receipts or physical tokens, that can be redeemed for services from the service provider (5). The authentication provider sends the objects to the user (8), who then redeems them with the service provider in a future interaction (9, 10).

6. GOVERNANCE CONSIDERATIONS

An important challenge that remains with the distributed ledger system described in section 5 is the management of the organizations that participate in the consensus mechanism of the distributed ledger. We believe that this will require the careful coordination of local businesses and cooperatives to ensure that the system itself does not impose any non-consensual trust relationships (Constraint 4), that no single market participant would gain dominance (Constraint 6), and that participating businesses and cooperatives will be able to continue to establish their own business practices and trust relationships on their own terms (Constraint 7), even while consenting to the decisions of the community of organizations participating in the shared ledger. We believe that our approach will be enhanced by the establishment of a multi-stakeholder process to develop the protocols by which the various parties can interact, including but not limited to those needed to participate in the distributed ledger, and ultimately facilitate

a multiplicity of different implementations of the technology needed to participate. Industry groups and regulators will still need to address the important questions of determining the rules and participants. We surmize that various organizations, ranging from consulting firms to aid organizations, would be positioned to offer guidance to community participants in the network, without imposing new constraints.

6.1. Open Standards

The case for a common, industry-wide platform for exchanging critical data via a distributed ledger is strong. Analogous mechanisms have been successfully deployed in the financial industry. Established mechanisms take the form of centrally-managed cooperative platforms such as SWIFT (Society for Worldwide Interbank Financial Telecommunication, 2018), which securely carries messages on behalf of financial markets participants, while others take the form of consensus-based industry standards, such as the Electronic Data Interchange (EDI) standards promulgated by X12 (The Accredited Standards Committee, 2018) and EDIFACT (United Nations Economic Commission for Europe, 2018). Distributed ledgers such as Ripple (2019) and Hyperledger (IBM, 2019) have been proposed to complement or replace the existing mechanisms.

For the digital identity infrastructure, we suggest that the most appropriate application for the distributed ledger system described in section 5 would be a technical standard for business transactions promulgated by a self-regulatory organization working concordantly with government regulators. A prime example from the financial industry is *best execution*, exemplified by Regulation NMS (Securities and Exchange Commission, 2005), which led to the dismantling of a structural monopoly in electronic equities trading in the United States[7]. Although the US Securities and Exchange Commission had the authority to compel exchanges to participate in a national market system since 1975, it was not until 30 years later that the SEC moved to explicitly address the monopoly enjoyed by the New York Stock Exchange (NYSE). The Order Protection Rule imposed by the 2005 regulation (Rule 611) "requir[ed] market participants to honor the best prices displayed in the national market system by automated trading centers, thus establishing a critical linkage framework" (Securities and Exchange Commission Historical Society, 2018). The monopoly was broken, NYSE member firms became less profitable, and NYSE was ultimately bought by Intercontinental Exchange in 2013 (Reuters, 2018).

We believe that distributed ledgers offer a useful mechanism by which self-regulatory organizations satisfy regulations precisely intended to prevent the emergence of control points, market concentration, or systems whose design reflects a conflict of interest between their operators and their users.

6.2. No Master Keys, No Early-Binding

An important expectation implicit to the design of our system is that users can establish and use as many identities as they want, without restriction. This means not only that a user can choose which credentials to show to relying parties, but also that a user would not be expected to bind the credentials to each other in any way prior to their use. Such a binding would violate Constraint 5 from section 3. In particular, given two credentials, there should be no way to know or prove that they were issued to the same individual or device. This property is not shared by some schemes for non-transferable anonymous credentials that encourage users to bind together credentials to each other via a master key or similar mechanism (Camenisch and Lysyanskaya, 2001) as described in section 2.2.

If it were possible to prove that two or more credentials were associated with the same identity, then an individual could be forced to associate a set of credentials with each other inextricably, and even if an individual might be given an option to reveal only a subset of his or her credentials to a service provider at any given time, the possibility remains that an individual might be compelled to reveal the linkages between two or more credentials. For example, the device that an individual uses might be compromised, revealing the master key directly, which would be problematic if the same master key were used for many or all of the individual's credentials. Alternatively, the individual might be coerced to prove that the same master key had been associated with two or more credentials.

The system we describe explicitly does not seek to rely upon the *ex ante* binding together of credentials to achieve non-transferability or for any other purpose. We suggest that the specific desiderata and requirements for non-transferability might vary across use cases and can be addressed accordingly. Exogenous approaches to achieve non-transferability might have authentication providers require users to store credentials using trusted escrow services or physical hardware with strong counterfeiting resistance such as two-factor authentication devices. Endogenous approaches might have authentication providers record the unblinded signatures on the ledger once they are presented for inspection, such that multiple uses of the same credential become explicitly bound to each other *ex post* or are disallowed entirely. Recall that the system assumes that credentials are used once only and that certification providers would generate new credentials for individuals in limited batches, for example at a certain rate over time.

TABLE 3 | A categorization matrix for use cases.

	Assert entitlements	Spend tokens
Private sector	Membership programmes	Mobile communication services
Public sector	Library access programme eligibility	Public transportation

We divide the universe of use cases into four categories based upon whether the purpose is to assert entitlements or to spend tokens and upon whether the services in question are operated by the public sector or the private sector, and we include some examples.

7. USE CASES

We anticipate that there might be many potential use cases for a decentralized digital identity infrastructure that affords users the ability to manage the linkages among their credentials. **Table 3** offers one view of how the use cases might be divided

[7]See also MiFID, its European Union counterpart (European Parliament, 2004).

into four categories on the basis of whether the purpose is to assert entitlements or spend tokens and upon whether the services in question are operated by the public sector or the private sector. Use cases that involve asserting entitlements might include asserting membership of a club for the purpose of gaining access to facilities, accessing restricted resources, or demonstrating eligibility for a discount, perhaps on the basis of age, disability, or financial hardship, at the point of sale. Use cases that involve spending tokens can potentially be disruptive, particularly in areas that generate personally identifiable information. We imagine that a decentralized digital identity infrastructure that achieves the privacy requirements would be deployed incrementally, whether general purpose or not. We suggest the following three use cases might be particularly appropriate because of their everyday nature, and might be a fine place to start:

1. ***Access to libraries.*** Public libraries are particularly sensitive to risks associated with surveillance (Zimmer and Tijerina, 2018). The resources of a public library are the property of its constituency, and the users have a particular set of entitlements that have specific limitations in terms of time and quantity. Entitlement levels could be managed by having the issuer use a different signing key for each entitlement. User limits could be enforced in several ways. One method involves requiring a user to make a security deposit that is released at the time that a resource has been returned and determined to be suitable for recirculation. Another method involves requiring the library to check the ledger to verify that a one-time credential has not already been used as a precondition for providing the resource and requiring the user to purchase the right to a one-time credential that can only be re-issued upon return of the resource.
2. ***Public transportation.*** It is possible to learn the habits, activities, and relationships of individuals by monitoring their trips in an urban transportation system, and the need for a system-level solution has been recognized (Heydt-Benjamin, 2006). Tokens for public transportation (for example, pay-as-you-go or monthly bus tickets) could be purchased with cash in one instance and then spent over time with each trip unlinkable to each of the others. This can be achieved by having an issuer produce a set of one-time use blinded tokens in exchange for cash and having a user produce one token for each subsequent trip. Time-limited services such as monthly travel passes could be issued in bulk, including a signature with a fixed expiration date providing a sufficiently large anonymity set. An issuer could also create tokens that might be used multiple times, subject to the proviso that trips for which the same token is used could be linked.
3. ***Wireless data service plans.*** Currently, many mobile devices such as phones contain unique identifiers that are linked to accounts by service providers and can be identified when devices connect to cellular towers (GSM Association, 2019). However, it is not actually technically necessary for service providers to know the particular towers to which a specific customer is connecting. For the data service business to be tenable, we suggest that what service providers really need is a way to know that their customers have paid. Mobile phone service subscribers could have their devices present blinded tokens, obtained from issuers following purchases at sales offices or via subscription plans, to cellular towers without revealing their specific identities, thus allowing them to avoid tracking over extended periods of time. Tokens might be valid for a limited amount of time such as an hour, and a customer would present a token to receive service for a limited time. System design considerations would presumably include tradeoffs between the degree of privacy and the efficiency of mobile handoff between towers or time periods.

We do not anticipate or claim that our system will be suitable for all purposes for which an individual might be required to present electronic credentials. We would imagine that obtaining security clearances or performing certain duties associated with public office might explicitly require unitary identity. Certain activities related to national security undertaken by ordinary persons, such as crossing international borders, might also fall into this category, although we argue that such use cases must be narrowly circumscribed to offer limited surveillance value through record linkage. In particular, linking any strongly non-transferable identifiers or credentials to the identities that individuals use for routine activities (such as social media, for example, or the use cases described above) would specifically compromise the privacy rights of their subjects. Other application domains, such as those involving public health or access to medical records, present specific complications that might require a different design. Certain financial activities would require interacting with regulated financial intermediaries who are subject to AML and KYC regulations, as mentioned in section 4.1. For this reason, achieving privacy for financial transactions might require a different approach that operates with existing financial regulations (Goodell and Aste, 2019).

8. CONCLUSIONS AND FUTURE WORK

We argue that the ability of individuals to create and maintain multiple unrelated identities is a fundamental, inalienable human right. For the digital economy to function while supporting this human right, individuals must be able to control and limit what others can learn about them over the course of their many interactions with services, including government and institutional services.

We have introduced a framework for an open digital identity architecture that promotes the implementation of identity architectures that satisfy constraints that we consider essential to the protection of human rights, and we believe that a combination of strong technology and thoughtful policy will be necessary to promote and ensure the implementation, deployment, and use of technology that satisfies them. We have elaborated eight requirements for technology infrastructure and demonstrated that they can be achieved by means of a decentralized architecture. Our framework does not seek strong non-transferability via an early-binding approach, and we argue that distributed ledgers can be used not only

to achieve the privacy objectives but also to deliver an alternative to strong non-transferability. We have identified challenges associated with scalability and governance, and we have also demonstrated how tokens can be spent via such a system as well as how the system might be used in an offline context.

Future work may include formal analysis of the information security properties of a system designed according to this framework, as well as the development of a proof-of-concept implementation and a corresponding evaluation of the various implementation tradeoffs relevant to different use cases. We suggest that different use cases would entail significantly different design choices.

The specific mechanism for fostering a community of participating organizations will depend upon the relationship between those organizations and the group that ultimately assumes the role of ensuring that the system does not impose non-consensual trust relationships on its users. It must be noted that any system that puts control in the hands of end-users carries the burden of education, both for the well-functioning of the system as well as for safeguarding its role in protecting the public interest. Future research, therefore, must include case studies of how similar systems have been developed, deployed, and maintained over time, in a variety of different social and business contexts.

AUTHOR CONTRIBUTIONS

GG is the primary author he performed the research and wrote the paper. TA directed the research, edited the paper for citations, audience framing, language, and contextualization with existing literature.

ACKNOWLEDGMENTS

We thank Valerie Khan, Edgar Whitley, Paul Makin, and Oscar King for their thoughtful insights. GG is also an associate of the Centre for Technology and Global Affairs at the University of Oxford. We acknowledge the Engineering and Physical Sciences Research Council (EPSRC) for the BARAC project (EP/P031730/1) and the European Commission for the FinTech project (H2020-ICT-2018-2 825215). TA acknowledges the Economic and Social Research Council (ESRC) for funding the Systemic Risk Centre (ES/K0 02309/1). This manuscript has been released as a Pre-Print at http://export.arxiv.org/pdf/1902.08769.

REFERENCES

5Rights Foundation (2019). Available online at: https://5rightsfoundation.com/ (accessed October 09, 2019).

Abelson, H., Anderson, R., Bellovin, S., Benaloh, J., Blaze, M., Diffie, W., et al. (1997). *The Risks of Key Recovery, Key Escrow, and Trusted Third-Party Encryption.* Available online at: https://academiccommons.columbia.edu (accessed March 11, 2019).

Abelson, H., Anderson, R., Bellovin, S., Benaloh, J., Blaze, M., Diffie, W., et al. (2015). Keys under doormats: mandating insecurity by requiring government access to all data and communications. *J. Cybersecur.* 1, 69–79. doi: 10.1093/cybsec/tyv009

Aitken, R. (2018). IBM Blockchain Joins Sovrin's "Decentralized" Digital Identity Network To Stem Fraud. *Forbes*. Available online at: https://www.forbes.com/sites/rogeraitken/2018/04/05/ibm-blockchain-joins-sovrins-decentralized-digital-identity-network-to-stem-fraud/ (accessed January 11, 2019).

Arduino (2019). Available online at: https://www.arduino.cc/ (accessed October 11, 2019).

Armer, P. (1975). Computer technology and surveillance. *Comput. People* 24, 8–11.

Benaloh, J. (2018). What if responsible encryption back-doors were possible? *Lawfare Blog*. Available online at: https://www.lawfareblog.com/what-if-responsible-encryption-back-doors-were-possible (accessed December 11, 2018).

Better Business Finance (2017). *What Are the AML and KYC Obligations of a Bank in the UK?* Available online at: https://www.betterbusinessfinance.co.uk/aml-and-kyc/what-are-the-aml-and-kyc-obligations-of-a-bank-in-the-uk (accessed May 28, 2017).

Brandao, L, Christin, N, Danezis, G, and Anonymous. (2015). Toward mending two nation-scale brokered identification systems. *Proc. Privacy Enhanc. Technol.* 2015, 135–155. doi: 10.1515/popets-2015-0022

Bright, P. (2010). Gov't, certificate authorities conspire to spy on SSL users? *Ars Technica*. Available online at: https://web.archive.org/web/20171004131406 (accessed May 11, 2018).

Camenisch, J., and Lysyanskaya, A. (2001). "An efficient system for non-transferable anonymous credentials with optional anonymity revocation," in *Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques (EUROCRYPT 2001: Advances in Cryptology)*, 93–118. Available online at: https://eprint.iacr.org/2001/019.pdf (accessed July 25, 2019).

Camenisch, J., and Lysyanskaya, A. (2003). "A signature scheme with efficient protocols." in *Lecture Notes in Computer Science*, Vol. 2576 (Springer). Available online at: http://rd.springer.com (accessed May 10, 2019).

Chainspace (2019). Available online at: https://chainspace.io/ (accessed October 11, 2019).

Charette, R. (2016). "DigiNotar Certificate Authority Breach Crashes e-Government in the Netherlands," in *IEEE Spectrum*. Available online at: https://spectrum.ieee.org/riskfactor/telecom/security/diginotar-certificate-authority-breach-crashes-egovernment-in-the-netherlands (accessed May 11, 2018).

Chaum, D. (1983). Blind signatures for untraceable payments. *Adv. Cryptol. Proc. Crypto* 82, 199–203.

Chirgwin, R. (2016). Google publishes list of Certificate Authorities it doesn't trust. *The Register* Available online at: https://www.theregister.co.uk/2016/03/23/google_now_publishing_a_list_of_cas_it_doesnt_trust/ (accessed May 11, 2018).

Douceur, J. (2002). "The Sybil Attack," *IPTPS '01 Revised Papers from the First International Workshop on Peer-to-Peer Systems*, 251–260. Available online at: https://www.freehaven.net/anonbib/cache/sybil.pdf (accessed May 11, 2018).

Dunphy, P., and Petitcolas, F. (2018). A first look at identity management schemes on the blockchain. *arXiv preprint* arXiv:1801.03294v1.

Eckersley, P. (2011). *Electronic Frontier Foundation Technical Analysis*. Available online at: https://www.eff.org/deeplinks/2011/03/iranian-hackers-obtain-fraudulent-https (accessed May 11, 2018).

Equifax Inc (2018). *2017 Cybersecurity Incident & Important Consumer Information.* Available online at: https://www.equifaxsecurity2017.com/ (accessed May 11, 2018).

European Parliament (1999). *Development of Surveillance Technology and Risk of Abuse of Economic Information.* Luxembourg: Scientific and Technological Options Assessment (STOA), PE 168.184/Vol 1/5/EN. Available online at: https://www.europarl.europa.eu/RegData/etudes/etudes/join/1999/168184/DG-4-JOIN_ET%281999%29168184_EN.pdf (accessed October 11, 2019).

European Parliament (2004). *Directive 2004/39/EC*. Official Journal of the European Union. Available online at: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:02004L0039-20060428 (accessed October 01, 2019).

Everest (2019). Everest blockchain software for verified value exchange. Available online at: https://everest.org/ (accessed October 01, 2019).

Evernym (2019). *The Self-Sovereign Identity Company* Available online at: https://www.evernym.com/ (accessed October 09, 2019).

Field, M. (2019). The tiny UK start-up founded by UCL scientists now at the heart of Facebook's Libra currency. *The Telegraph*. Available online at: https://www.telegraph.co.uk/technology/2019/06/26/inside-tiny-london-start-up-heart-facebooks-push-reinvent-world/ (accessed October 11, 2019).

Financial Action Task Force (FATF) (2018). *The FATF Recommendations*. Available online at: http://www.fatf-gafi.org/media/fatf/documents/recommendations/pdfs/FATF%20Recommendations%202012.pdf (accessed September 16, 2018).

Goodell, G., and Aste, T. (2018). *Blockchain Technology for the Public Good: Design Constraints in a Human Rights Context*. Open Access Government. Available online at: https://www.openaccessgovernment.org/blockchain-technology-for-the-public-good-design-constraints-in-a-human-rights-context/44595/

Goodell, G., and Aste, T. (2019). Can cryptocurrencies preserve privacy and comply with regulations? *Front. Blockchain* 2:4. doi: 10.3389/fbloc.2019.00004

Goodin, D. (2017a). Already on probation, Symantec issues more illegit HTTPS certificates. *Ars Technica*, Available online at: https://arstechnica.com/information-technology/2017/01/already-on-probation-symantec-issues-more-illegit-https-certificates/ (accessed May 11, 2018).

Goodin, D. (2017b). Flaw crippling millions of crypto keys is worse than first disclosed. *Ars Technica*. Available online at: https://arstechnica.com/information-technology/2017/11/flaw-crippling-millions-of-crypto-keys-is-worse-than-first-disclosed/ (accessed April 21, 2018).

Goodin, D. (2017c). Stuxnet-style code signing is more widespread than anyone thought. *Ars Technica*. Available online at: https://arstechnica.com/information-technology/2017/11/evasive-code-signed-malware-flourished-before-stuxnet-and-still-does/ (accessed April 21, 2018).

GOV.UK (2014). *Money Laundering Regulations: Who Needs to Register*. Available online at: https://www.gov.uk/guidance/money-laundering-regulations-who-needs-to-register (accessed May 28, 2017).

Government Digital Service (UK) (2018). *GOV.UK Verify: Guidance*. Available online at: https://www.gov.uk/government/publications/introducing-govuk-verify/introducing-govuk-verify (accessed Febuary 15, 2019).

Graglia, M., Mellon, C., and Robustelli, T. (2018). The nail finds a hammer: self-sovereign identity, design principles, and property rights in the developing world. *New America* Available online at: https://www.newamerica.org/future-property-rights/reports/nail-finds-hammer/exploring-three-platforms-through-the-principles/ (accessed October 10, 2019).

GSM Association (2019). *IMEI Database*. Available online at: https://imeidb.gsma.com/imei/index (accessed October 11, 2019).

Heath, A. (2019). Facebook makes first blockchain acquisition with chainspace: sources. *Cheddar, Inc*. Available online at: https://cheddar.com/media/facebook-blockchain-acquisition-chainspace (accessed Febuary 05, 2019).

Heydt-Benjamin, T., Chae, H., Defend, B., and Fu, K. (2006). "Privacy for public transportation." in *International Workshop on Privacy Enhancing Technologies (PETS)*, 1–19. [online] Available online at: https://petsymposium.org/2006/preproc/preproc_01.pdf (accessed October 05, 2019).

IBM (2019). *Hyperledger: Blockchain Collaboration Changing the Business World*. Available online at: https://www.ibm.com/blockchain/hyperledger.html (accessed January 11, 2019).

IBM Research Zurich (2018). *IBM Identity Mixer*. [online] Available online at: https://www.zurich.ibm.com/identity_mixer/ (accessed April 21, 2018).

ID2020 (2019). *Digital Identity Alliance*. Available online at: https://id2020.org/ (accessed October 09, 2019).

ID2020 Alliance (2019). ID2020 Alliance launches digital ID program with Government of Bangladesh and Gavi, announces new partners at annual summit. *PR Newswire*. Available online at: https://www.prnewswire.com/news-releases/id2020-alliance-launches-digital-id-program-with-government-of-bangladesh-and-gavi-announces-new-partners-at-annual-summit-300921926.html (accessed October 11, 2019).

International Telecommunications Union (2018). *Digital Identity Roadmap Guide*, Available online at: http://handle.itu.int/11.1002/pub/81215cb9-en (accessed October 01, 2019).

Juniper Networks (2018). *Understanding Online Certificate Status Protocol*. Available online at: https://www.juniper.net/documentation/en_US/junos/topics/concept/certificate-ocsp-understanding.html (accessed May 11, 2018).

Kaaniche, N., and Laurent, M. (2017). "A blockchain-based data usage auditing architecture with enhanced privacy and availability," in *IEEE 16th International Symposium on Network Computing and Applications (NCA)*. Available online at: https://ieeexplore.ieee.org/abstract/document/8171384/ (accessed January 11, 2019).

Kozlowska, H., Gershgorn, D., and Todd, S. (2018). *The Cambridge Analytica Scandal is Wildly Confusing. This Timeline Will Help*. Quartz. Available online at: https://qz.com/1240039/the-cambridge-analytica-scandal-is-confusing-this-timeline-will-help/ (accessed April 20, 2018).

Kuner, C., and Marelli, M. (2017). *Handbook on Data Protection in Humanitarian Action*. Geneva: International Committee of the Red Cross. Available online at: https://www.icrc.org/en/publication/handbook-data-protection-humanitarian-action (accessed October 01, 2019).

Kushner, D. (2013). "The Real Story of Stuxnet," in *IEEE Spectrum* Available online at: https://spectrum.ieee.org/telecom/security/the-real-story-of-stuxnet (accessed May 11, 2018).

Liu, S., Wei, J., and Li, C. (2008). *Method and System for Implementing Authentication on Information Security*. United States Patent Application US20080065895A1. Available online at: https://patents.google.com/patent/US20080065895A1/en (accessed January 11, 2019).

Lunden, I. (2018). Russia's Telegram ban that knocked out 15M Google, Amazon IP addresses had a precedent in Zello. *TechCrunch*. Available online at: https://techcrunch.com/2018/04/17/russias-telegram-ban-that-knocked-out-15m-google-amazon-ip-addresses-had-a-precedent-in-zello/ (accessed April 20, 2018).

Lundkvist, C., Heck, R., Torstensson, J., Mitton, Z., and Sena, M. (2016). *uPort: A Platform for Self-Sovereign Identity*. Available online at: http://blockchainlab.com/pdf/uPort_whitepaper_DRAFT20161020.pdf (accessed January 11, 2019).

Mayo, E. (1945). "Hawthorne and the Western Electric Company" in *The Social Problems of an Industrial Civilization*. Boston, MA: Division of Research, Harvard Business School. Available online at: http://www.practicesurvival.com/wa_files/Hawthorne_20Studies_201924_20Elton_20Mayo.pdf (accessed January 06, 2019).

Moon, M. (2017). Estonia freezes resident ID cards due to security flaw. *Engadget*. Available online at: https://www.engadget.com/2017/11/04/estonia-freezes-resident-id-cards-security-flaw/ (accessed April 21, 2018).

O'Hara, K., Whitley, E., and Whittall, P. (2011). Avoiding the jigsaw effect: experiences with ministry of justice reoffending data. *Monograph*. Available online at: https://eprints.soton.ac.uk/273072/1/AVOIDING%2520THE%2520JIGSAW%2520EFFECT.pdf (accessed July 26, 2018).

Pandya, J. (2019). Hacking Our identity: the emerging threats from biometric technology. *Forbes*. Available online at: https://www.forbes.com/sites/cognitiveworld/2019/03/09/hacking-our-identity-the-emerging-threats-from-biometric-technology/ (accessed October 09, 2019).

Parliament of Australia (2018). *Telecommunications and Other Legislation Amendment (Assistance and Access) Bill 2018*. Available online at: https://www.aph.gov.au/Parliamentary_Business/Bills_Legislation/Bills_Search_Results/Result?bId=r6195 (accessed October 01, 2019).

Parliament of the United Kingdom (2016). *Investigatory Powers Bill: Committee Stage Report*. House of Commons Library (Commons Briefing papers CBP-7578). Available online at: https://researchbriefings.parliament.uk/ResearchBriefing/Summary/CBP-7578 (accessed October 09, 2019).

Privacy International (2019). *Identity*. Available online at: https://privacyinternational.org/topics/identity (accessed October 01, 2019)

Ramakrishnan, K., and Selvarajan, R. (2017). *Transforming the Telecom Value Chain with Platformization*. TATA Consultancy Services White Paper. Available online at: https://www.tcs.com/content/dam/tcs/pdf/Industries/communication-media-and-technology/Abstract/Transforming%20the%20Telecom%20Value%20Chain%20with%20Platformization.pdf (accessed January 11, 2019).

Reuters (2018). *ICE completes takeover of NYSE*. Available online at: https://www.reuters.com/article/us-ice-nyse-sprecher/ice-completes-takeover-of-nyse-idUSBRE9AB16V20131112 (accessed April 19, 2018).

Riley, S. (2006). *It's Me, and Here's My Proof: Why Identity and Authentication Must Remain Distinct*. Microsoft TechNet Security Viewpoint. Available online at:

https://technet.microsoft.com/en-us/library/cc512578.aspx (accessed May 11, 2018).

Ripple (2019). *One Frictionless Experience to Send Money Globally*. Available online at: https://ripple.com/ (accessed January 11, 2019).

Rohingya Project (2019). *Unlocking Potential*. Available online at: http://rohingyaproject.com/ (accessed June 05, 2019).

Savov, V. (2018). Russia's Telegram ban is a big, convoluted mess. *The Verge*. Available online at: https://www.theverge.com/2018/4/17/17246150/telegram-russia-ban (accessed April 20, 2018).

SecureKey Technologies, Inc (2015). *Trust Framework – SecureKey Concierge in Canada, SKUN-117*. Available online at: http://securekey.com/wp-content/uploads/2015/09/SK-UN117-Trust-Framework-SecureKey-Concierge-Canada.pdf (accessed May 11, 2018).

Securities and Exchange Commission (US) (2005). *REGULATION NMS*. 17 CFR PARTS 200, 201, 230, 240, 242, 249, and 270; Release No. 34-51808; File No. S7-10-04, RIN 3235-AJ18. Available online at: https://www.sec.gov/rules/final/34-51808.pdf (accessed January 30, 2019).

Securities and Exchange Commission Historical Society (2018). *2000s: Timeline*. [online] Available online at: http://www.sechistorical.org/museum/timeline/2000-timeline.php (accessed April 19, 2018).

Shuhaimi, N., and Juhana, T. (2012). "Security in vehicular *ad-hoc* network with Identity-Based Cryptography approach: a survey," in *7th International Conference on Telecommunication Systems, Services, and Applications (TSSA)*. Available online at: http://ieeexplore.ieee.org/abstract/document/6366067/ (accessed January 11, 2019).

SITA (2016). *Travel Identity of the Future – White Paper*. ShoCard. Available online at: https://shocard.com/wp-content/uploads/2016/11/travel-identity-of-the-future.pdf (accessed January 11, 2019).

Slashdot (2014). *School Tricks Pupils Into Installing a Root CA*. Available online at: https://news.slashdot.org/story/14/03/09/0225224/school-tricks-pupils-into-installing-a-root-ca (accessed May 11, 2018).

Society for Worldwide Interbank Financial Telecommunication (2018). *Discover SWIFT*. Available online at: https://www.swift.com/about-us/discover-swift (accessed April 21, 2018).

Software in the Public Interest, Inc (2019). *Debian: The Universal Operating System*. Available online at: https://www.debian.org/ (accessed October 11, 2019).

Sonnino, A., et al. (2018). Coconut: threshold issuance selective disclosure credentials with applications to distributed ledgers. *arXiv preprint* arXiv:1802.07344v3.

Stevens, L. (2018). "Self-sovereign identity systems for humanitarian interventions," in *Working Paper*. Available online at: https://pdfs.semanticscholar.org/f821/4975160857f1f020ff8dbc2db65f88fcac03.pdf (accessed October 01, 2019).

Stiegler, M. (2005). *An Introduction to Petname Systems*. Available online at: http://www.skyhunter.com/marcs/petnames/IntroPetNames.html (accessed May 11, 2018).

Thackston, J. (2018). *System and Method for Verifying User Identity in a Virtual Environment*. US Patent Grant US10153901B2. Available online at: https://patents.google.com/patent/US10153901B2/en (accessed January 11, 2019).

The Accredited Standards Committee (2018). *ASC X12*. Available online at: https://web.archive.org/web/20140927153741/http://www.x12.org/x12org/about/index.cfm (accessed April 21, 2018).

The World Bank (2019). *Identification for Development*. Available online at: https://id4d.worldbank.org/ (accessed October 09, 2019).

Thevoz, P. (2016). Diving into a 'Digital Country': e-Estonia. *Medium*. Available online at: https://medium.com/@PhilippeThevoz/diving-into-a-digital-country-e-estonia-af561925c95e (accessed April 21, 2018).

Thomas, T. (2009). Joint watermarking scheme for multiparty multilevel DRM architecture. *IEEE Trans. Inform. Forensics Secur.* 4, 758–767. doi: 10.1109/TIFS.2009.2033229

Tobin, A., and Reed, D. (2016). The inevitable rise of self-sovereign identity. *The Sovrin Foundation*. Available online at: https://sovrin.org/wp-content/uploads/2017/06/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf (accessed January 11, 2019).

Tully, M. (2017). The problem with Aadhaar cards is the way they are being pushed by the State. *Hindustan Times*. Available online at: https://www.hindustantimes.com/analysis/the-problem-with-aadhaar-cards-is-the-way-they-are-being-pushed-by-the-state/story-RTlWUXgF3ck4rsoN1zKXUI.html (accessed April 21, 2018).

United Nations (1948). *Universal Declaration of Human Rights*. General Assembly Resolution 217 A, Paris. Available online at: http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf (accessed October 11, 2019).

United Nations Economic Commission for Europe (2018). *Trade Programme – Trade – UNECE*. Available online at: http://www.unece.org/tradewelcome/trade-programme.html (accessed April 21, 2018).

Value Chain Dynamics Working Group (VCDWG) (2005). *Value Chain Dynamics in the Communication Industry*. MIT Communications Futures Program and Cambridge University Communications Research Network. Available online at: http://cfp.mit.edu/docs/core-edge-dec2005.pdf (accessed January 11, 2019).

Vanderburg, E. (2018). A certified lack of confidence: the threat of Rogue Certificate authorities. *TCDI Blog*. Available online at: https://www.tcdi.com/the-threat-of-rogue-certificate-authorities/ (accessed May 11, 2018).

Wagner, R. (2014). Identity and access management 2020. *ISSA J.* 12, 26–30.

Waldman, P., Chapman, L., and Robertson, J. (2018). Palantir knows everything about you. *Bloomberg*. Available online at: https://www.bloomberg.com/features/2018-palantir-peter-thiel/ (accessed April 19, 2018).

Whitley, E. (2018). *Trusted Digital Identity Provision: GOV.UK Verify's Federated Approach*. Center for Global Development Policy Paper 131. Available online at: https://www.cgdev.org/sites/default/files/Trusted-Digital-ID-Provision-govuk.pdf (accessed Febuary 15, 2019).

Wilcox-O'Hearn, Z. (2018). *Names: Decentralized, Secure, Human-Meaningful: Choose Two*. Available online at: https://web.archive.org/web/20011020191610/http://zooko.com/distnames.html (accessed April 21, 2018).

Zimmer, M., and Tijerina, B. (2018). *Library Values & Privacy in Our National Digital Strategies: Field Guides, Convenings, and Conversations*. National Leadership Grant for Libraries Award Report. Milwaukee, WI. Available online at: https://www.michaelzimmer.org/2018/08/02/project-report-library-values-privacy/ (accessed October 05, 2019).

Zuboff, S. (2015). Big Other: surveillance capitalism and the prospects of an information civilization. *J. Inform. Technol.* 30, 75–89. doi: 10.1057/jit.2015.5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

HYPOTHESIS AND THEORY
published: 23 January 2020
doi: 10.3389/fbloc.2019.00028

Self-Sovereign Identity in a Globalized World: Credentials-Based Identity Systems as a Driver for Economic Inclusion

Fennie Wang[1] *and Primavera De Filippi*[2,3]*

[1] *Independent Researcher, New York, NY, United States,* [2] *Berkman-Klein Center for Internet & Society, Harvard University, Cambridge, MA, United States,* [3] *CERSA/CNRS, Paris, France*

After introducing key concepts and definitions in the field of digital identity, this paper will investigate the benefits and drawbacks of existing identity systems on the road toward achieving self-sovereign identity. It will explore, in particular, the use of blockchain technology and biometrics as a means to ensure the "unicity" and "singularity" of identities, and the associated challenges pertaining to the security and confidentiality of personal information. The paper will then describe an alternative approach to self-sovereign identity based on a system of blockchain-based attestations, claims, credentials, and permissions, which are globally portable across the life of an individual. While not dependent on any particular government or organization for administration or legitimacy, credentials and attestations might nonetheless include government-issued identification and biometrics as one of many indicia of identity. Such a solution—based on a recorded and signed digital history of attributes and activities—best approximates the fluidity and granularity of identity, enabling individuals to express only specific facets of their identity, depending on the parties with whom they wish to interact. To illustrate the difficulties inherent in the implementation of a self-sovereign identity system in the real world, the paper will focus on two blockchain-based identity solutions as case studies: (1) Kiva's identity protocol for building credit history in Sierra Leone, and (2) World Food Programme's Building Blocks program for delivering cash aid to refugees in Jordan. Finally, the paper will explore how the combination of blockchain-based cryptocurrencies and self-sovereign identity may contribute to promoting greater economic inclusion. With digital transactions functioning as identity claims within an ecosystem based on self-sovereign identity, new business models might emerge, such as identity insurance schemes, along with the emergence of value-stable cryptocurrencies ("stablecoins") functioning as local currencies.

Keywords: blockchain, self-sovereign identity, migrants, credentials, digital identity

OPEN ACCESS

Edited by:
Oskar Josef Gstrein, University of Groningen, Netherlands

Reviewed by:
Michael Cooper, Emergence, United States
Nichola Cooper, University of the Sunshine Coast, Australia

***Correspondence:**
Primavera De Filippi
pdefilippi@gmail.com

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *19 September 2019*
Accepted: *20 December 2019*
Published: *23 January 2020*

Citation:
Wang F and De Filippi P (2020) Self-Sovereign Identity in a Globalized World: Credentials-Based Identity Systems as a Driver for Economic Inclusion. Front. Blockchain 2:28. doi: 10.3389/fbloc.2019.00028

INTRODUCTION TO IDENTITY MANAGEMENT SYSTEMS

In this section, we will introduce a set of principles and terminology relevant in the identity space, particularly as applied to technologies used to implement identity management systems such as web standards, cryptography, blockchain ledgers, and cryptocurrency applications.

Preliminary Definitions

There is currently much confusion in the identity space with regard to specific core terms such as "identity" and "identifier," "attributes" and "persona," which are often used interchangeably and ambiguously, without properly defining the meaning and scope of each term. We provide here a preliminary distinction between these terms, along with a tentative definition that will be used in the remainder of this paper.

An "**identity**" has been defined in different manners, depending on the field of endeavor. In psychology, it is generally used to refer to all the psychological traits of a person, inclusive of the personality, beliefs and other personal attributes (Strohminger et al., 2017). In sociology, it includes the culture, history, religion and tradition that an individual is part of it (Côté, 1996). From a legal standpoint, an identity can be associated to the concept of a "natural person" (i.e., an actual human being), or a "legal person" (which might refer to a company, a trust, a partnership, or another collective of people identified as a single person under the law).

For the purpose of this paper, we use the terminology of "identity" to describe all attributes of a person that uniquely defines the person over the course of a lifetime, providing sameness and continuity despite varying aspects and conditions. As such, we distinguish between the notion of "*numerical identity*" which describes the relationship that holds exclusively between a thing and itself[1], and the notion of "*qualitative identity*" which merely describes the properties that different things have in common (Garrett, 2002): only when there is total qualitative identity between two things, can these two things be regarded as being numerically identical.

Yet, even in the context of numerical identity, it is important to note that the attributes of an identity can evolve over time. Identity formation is an ongoing process, whereby a person's identity is developed over the course of the years, and constantly evolves as a result of the interactions with the person's environment (Eakin, 1999). Accordingly, identity is dynamic and multifaceted, and every identity management system must therefore be designed in such a way as to be sufficiently flexible, resilient, and dynamic to accommodate the variable and complex nature of human identity. However, regardless of the sophistication of these systems, no identity management system will ever be able to categorically capture all aspects of one's identity. Indeed, insofar as we attempt to design a system to manage and categorize a variety of different identities, it is important to understand from the outset that such categorization will necessarily be a reduction of the specific facet or use case of each identity it comprises[2].

A "**persona**" is a specific facet of an identity that is expressed in a particular context. While the identity uniquely defines a person, the same person can hold multiple personas, depending on the social context that is taken in consideration (Suler, 2002). For instance, Alice might be a dedicated mom for her daughter, and a loving wife for her husband. She might be a trusted friend to some of her peers, and strict manager to her employees. All these personas are part of the same identity but might display slightly modified features or psychological traits. From a technical standpoint, they can be described as pseudonyms or practical identities (Christman, 2013). While an identity is an abstract concept that relates to the individual as a whole, a persona is a crucial component of any identity management system, because it relates to the way in which individuals "authenticate" themselves to the system (Toth and Subramanium, 2003).

An "**attribute**" describes an essential, definitional property of a person that qualifies it as a member of a given set (or class) of persons. As such, an attribute is generally not unique to that person. Each person can have an indefinite number of attributes: elements like gender, height, weight, handicaps or capabilities which are inherent to the person, or elements like nationality and citizenship, which have been assigned (and could potentially be revoked) by a third-party, with a view to distinguish or organize people into specific categories (e.g., U.S. vs. French citizens). Of course, most of these categories are abstract classes that can be arbitrarily defined, even if they refer to an inherent property. Consider the attribute of having "red hair" that qualifies a person as part of the red-hair people set. Clearly, it is a natural, non-revocable attribute, yet the class of red-hair people is somewhat arbitrarily defined (what is the exact shade of red that qualifies someone as such?). Similarly, the "gender" category which had been for a long time limited to "male" or "female" is recently being expanded with the advent of people who identify as "non-binary." Finally, one of the key characteristic of attributes is that, because they are intended to classify an entity into a particular category, they are not unique to it: multiple entities may share the exact same attributes.

An "**identifier,**" conversely, is not intended to describe or qualify a person, but rather to be used as a "reference" to a real-world identity (or a specific persona). As such, identifiers are often assigned (arbitrarily) by a third-party, with regard to a particular use case or domain (e.g., the legal name of a person, a social security number, or a simple username). In other cases, they can be a particular representation of an observable property of an entity (like fingerprints or other biometric data). It is important to note that both attributes and identifiers are, from a strictly technical perspective, mere data strings that can be used as a means to authenticate a particular individual (or persona). Depending on the domain at hand, the same data string can be used to qualify an entity as a member of a set,

[1] As its name indicates, numerical identity describes the relation through which things can be counted: x and y can be counted as one only if they are numerically identical (Geach, 1973).

[2] This is mostly due to the gap that exists between a first person knowledge of self, and a third party knowledge of a person by description (Burge, 1988).

distinguish from members of different sets, or uniquely identify them within a set. Yet, attributes and identifiers differ with regard to their purpose: an attribute (as a "qualifier") is aimed at classifying people within a particular category, whereas an identifier (as a "reference") is intended to identify someone within a particular domain. Accordingly, even though some identity management systems allow for multiple individuals to share the same identifier (e.g., many individuals share an identical name), or for one individual to have more than one identifier (e.g., in the case of pseudonyms), in order to facilitate the process of identification and authentication, it is often desirable that an identifier be able to identify a person in a *unique* and *unambiguous* way (Jøsang and Pope, 2005). This requires an identity management system to fulfill at least two basic criteria: (1) no two people should have the same identifier (*unicity*), and (2) no one person should have more than one identifier (*singularity*) in the same domain.

In light of this, most identifiers are comprised of a random string of characters that are unique in a particular domain. These are generally issued by a centralized entity, such as a government agency or administrative body, as in the case of a passport number or social security number; or by a company or organization, as in the case of a bank account or an email address. Centralization, in this context, helps ensure a degree of confidence that the identifier is *unique* (i.e., that the same social security number has not been assigned to two different persons) and *singular* to one identity (i.e., that no one may have more than one social security number).

Alternatively, an identifier can be generated directly by the person, as in the case of a pair of cryptographic keys used to access a cryptocurrency wallet. In this case, *unicity* is guaranteed by mathematics—at least at a very high degree of probability (Schartner and Schaffer, 2005), but *singularity* cannot be guaranteed (i.e., the same person can generate more than one identifier). Similarly, decentralized identifiers (DIDs) are an open source web-based standard, which uses a web address (URL) as the unique identifier that contains or points to public identifying information about the identity subject. The public identifying information linked to a DID may include publicly viewable credentials or attestations, or the public key/address of a cryptocurrency wallet. In this way, DIDs may be used in conjunction with blockchain technology and public-private key pairs (Mühle et al., 2018).

Finally, recent technological advances made it possible to develop biometric identifiers that are directly related to the physicality of a person, as in the case of a fingerprint, iris scan or face recognition. If we discount possible errors and inaccuracies related to the technology (Proença and Alexandre, 2010; Canham, 2018), biometric identifiers are often touted as being both *unique* and *singular* to one identity. However, biometric templates are limited to the extent that even the most sophisticated scanning tools only provide approximate representations (Nagar et al., 2010). This is somewhat mitigated by multimodal biometrics (iris scan, combined with fingerprints, face recognition, etc.) that provide higher degree of rarity (Ross and Jain, 2004). Ultimately, it all depends on the size of a population set (Duta, 2009): given a small population, such identifiers can be said to be unique—although this creates serious privacy problems (see below for more details on the matter).

The Interplay of Identifiers, Personas, and Key Pairs on the Web

With respect to the Internet, the most fundamental identifier, at the network layer, is the IP address, which makes it possible to route packets from one machine to another, until it reaches the right machine. The IP address does not communicate any information about the machine it refers to (i.e., it is not an attribute of it), however, in some cases, it is possible to link an IP address back to a particular individual or organization, whose identity can be ascertained by the relevant Internet Service Provider (ISP).[3]

At the application layer, user accounts and passwords are used to identify specific personas (which may be persons, companies, machines or other entities) interacting on an online service. While these also do not provide, as such, any personal information about the persona, many online service providers require users to communicate additional attributes or identifiers (e.g., real name, age, etc.) in order to ensure that only legitimate individuals can access the service.

Yet, it is worth mentioning that both in the case of an IP address and a user account, only a subset of these identifiers may actually resolve to a natural person. De facto, these identifiers merely refer to a particular endpoint interacting with an online service, but there is no guarantee that this endpoint can be uniquely associated with an individual identity. For instance, an IP address might be used by a multiplicity of persons, and many user accounts are nowadays controlled by bots, rather than persons.

In the context of a blockchain-based system, identifiers are generally managed with public/private key pairs, which uniquely identify the wallet holder (De Filippi and Wright, 2018). Yet, these also do not communicate any personal identifying information about the person, unless additional information is associated with them (Androulaki et al., 2013). Therefore, the same entity (a person, a computer or bot) may own or control multiple key pairs, as key pairs do not necessarily refer to an individual identity. For example, Mary owns a key pair to her Bitcoin wallet, and a different key pair to her Ether wallet.

From a technical perspective, the public-private key pairs are proof of both custody and ownership to any cryptocurrency or tokenized asset held in a particular digital address, or wallet. The private key is necessary to execute transactions to and from the blockchain address identified by the public key. A transaction is not limited to the transfer of a crypto-asset such as a Bitcoin or Ether, but may also represent the transfer or issuance of a cryptographic token through a smart contract transaction (Wright and De Filippi, 2015). An example would be a data access

[3]The European General Data Protection Regulation 2016/679 (GDPR) states that IP addresses should be considered personal data, to the extent that the ISP has a record of the IP address and knows to whom it has been assigned. See recital 30 of the GDPR, which clarifies "online identifier" as mentioned in the Article 4 definition of personal data: "*Natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers*".

token, which the owner of a dataset (such as a health record or credit history) issues to a third party wishing to access some of the data. The token functions like a key to the datastore, and transactions of that token are recorded on a blockchain ledger to keep track of who has been granted permission and access (Maesa et al., 2017).

In a public and permissionless[4] blockchain like Bitcoin or Ethereum, which operates without any centralized authority or intermediary operator (De Filippi and Loveluck, 2016), the nodes maintaining the network (e.g., the "miners") operate without association to a particular given identity (El Haddouti and El Kettani, 2019). In a permissioned blockchain, where a centralized entity or consortium is in charge of identifying or policing the nodes that maintain the blockchain ledger, the key-pairs controlled by each miner are generally associated with real world identities (Hardjono and Pentland, 2019). Reliance on real-world identities provides the additional ability to police (and punish), thereby enabling permissioned blockchains to dispense with some of the security measures that anonymous (or pseudonymous) permissionless public chains must employ, e.g., Proof of Work or Proof of Stake (Shrier et al., 2016). The caveat is that users must trust the governance practices of the central entity or consortium policing the permissioned blockchain (Davidson et al., 2016).

Centralized Identification System Based on Unique Identifiers vs. Multifaceted Web of Trust Claims and Credentials System

As previously discussed, the key tenets of any properly functioning identity system are the properties of "unicity" and "singularity." Unicity refers to the fact that each identifier is used to uniquely identify one (and only one) individual, i.e., no two persons should have the same identifier. Singularity refers to the fact that each individual possesses one (and only one) identifier in a particular domain, i.e., no two identifiers should refer to the same individual.

Unicity can be achieved without a centralized authority, because mathematical primitives can ensure that no two people get the same identifier, even if there is no central authority to coordinate the identifiers. Each identity provider can issue an identifier using very large random numbers, and even though there is a theoretical possibility that two actors issue the same identifier to different beneficiaries, the probability is so low to be negligible.

In order to fulfill these the singularity requirements, however, most of the existing identity systems rely on a central authority to ensure that each unique and unambiguous identifier is linked to a singular identity (Kulkarni et al., 2012). The centralized authority must collect personal information to ensure the singularity of any given identifier issued into the system. Such a system is generally expensive and bureaucratic, likely politically impractical for the use case of migrants (especially for vulnerable populations on the move), and subject to high privacy, data abuse, and cybersecurity risks (Whitley and Hosein, 2010). For instance, in 2012, India has launched the Aadhaar identity management system, using biometric data to identify its 1.3 billion inhabitants—many of whom do not have any formal identification (Sarkar, 2014). Participation into the Aadhaar system has become a requirement for Indians to receive welfare benefits, sign up for mobile phones or register at school. However, such a system has raised concerns from civil liberties groups (Jain and Nandakumar, 2012), with multiple lawsuits before India's Supreme Court whether such a system violates India's constitutional right to privacy[5].

Ideally, an identity system should respect the multifaceted nature of identity and look at the different attributes or personas depending on the use cases. Only a small handful of use cases actually require a unique and singular link between an individual and its identifier (i.e., that an individual be identified by a single and unique identifier in a particular domain). This might be the case of voting, whereby a single person should be excluded from voting multiple times under multiple identifiers (Cap and Maibaum, 2001; Alvarez et al., 2009).

An alternative to an identity system based on unique and singular identifiers is a claims and credentials based system (Rannenberg et al., 2015). In such a system, identity is not reduced to an *authoritative* identifier, such as biometric or government issued identification numbers; rather, identity is defined through a network of claims and credentials based on a *web of trust*[6] authentication (Khare and Rifkin, 1997). Such a system better mirrors the multifaceted nature of human identity, allowing for different *profiles* and *personas* to emerge through a combination of different claims and credentials depending on the use cases. A profile that is appropriate for a loan application may be different than the one used in public forums. While such a system would not necessarily guarantee the singularity of individuals using the system, it would suit a large majority of day-to-day use cases.

THE ROLE OF IDENTITY FOR SOCIO-ECONOMIC INCLUSION

For many years, the World Bank has stressed the need for every citizen to be endowed with a valid proof of identity, as identification has become a necessity for financial inclusion and access to essential services and rights. Specifically, from a

[4]A "permissionless" blockchain is a blockchain that anyone can join, and where every node is entitled to both read the current state of the blockchain, and add new blocks to the blockchain. A "public" blockchain, conversely, refers only to the ability to read the blockchain, which can be either permissioned or permissionless based on the rights for who may add information to the blockchain.

[5]Since 2012, Aadhaar was the object of more than 30 petitions and its constitutionality has been repeatedly challenged in courts. In September 2018, the Indian Supreme Court held that, in spite of these claims, Adhaar was legitimate, although with a limited scope and restrictions on data storage. For more information, see https://www.sci.gov.in/supremecourt/2012/35071/35071_2012_Judgement_26-Sep-2018.pdf

[6]The "web of trust" concept was first put forth by PGP creator Phil Zimmermann in 1992 in the manual for PGP version 2.0: "As time goes on, you will accumulate keys from other people that you may want to designate as trusted introducers. Everyone else will each choose their own trusted introducers. And everyone will gradually accumulate and distribute with their key a collection of certifying signatures from other people, with the expectation that anyone receiving it will trust at least one or two of the signatures. This will cause the emergence of a decentralized fault-tolerant web of confidence for all public keys."

development perspective, a recent report of the World Bank[7] identifies three overarching goals for any identification system:

- ***Inclusion and access to essential services*** *such as health care and education, electoral rights, financial services, and social safety net programs;*
- ***Effective and efficient administration of public services****, transparent policy decisions and improved governance—particularly to reduce duplication and waste;*
- ***More accurate measure of development progress*** *in areas such as reduction in maternal and infant mortality.*

Yet, still today, more than 1.5 billion people are excluded from accessing basic services due to their inability to prove their identity[8]. A large majority of these people are located in Asia and Africa, in areas that lack the proper infrastructure to register births and other life events (e.g., in South Asia and Sub-Saharan Africa, respectively, only 39 and 44% of children have births registered[9]) and generally belong to some of the poorest segments of the population.

At the same time, according to the UNHCR[10], there are currently over 70 million forcibly displaced people as a result of conflict or persecution, 25 million of which are refugees—mostly from Syria, Afghanistan, and South Sudan. There are also approximately four million stateless people, who have been denied a nationality, and therefore have been cut off access to basic services and rights. These numbers are expected to grow in the years to come, especially in light of the growing impact of climate change—which has been recognized as a key contributing factor to political conflicts[11], and as a significant driver to both internal and international migration[12].

In light of this, the UN has recently launched the ID2020 Alliance[13], a multi-stakeholder partnership that brings together multinational organizations, non-profits, businesses and government, all geared toward the objective of ensuring that digital identity is responsibly implemented and widely accessible. The goals of the Alliance are twofold: on the one hand, it is in charge of defining the parameters for good and ethical digital identity systems, and, on the other hand, it is responsible for funding and implementing digital identity projects with a social good mindset. Among other things, the ID2020 Alliance has also created a Certification Mark[14], used to label technological solutions that meet the technical standards and requirements established by the Alliance and that satisfy the principles of portability, persistence, privacy, and user-control.

Many proof-of-concepts are currently being developed by public and private institutions to provide digital identity to those currently lacking formal means of identification[15]. Yet, when devising these identity solutions, it is important to ensure that one single actor does not hold and control the personal identity records of every identified individual, which may raise significant privacy concerns. In the case of refugees lacking proper identification, in particular, digital identity could be used as a means to identify specific individuals or families which are eligible for cash aid or other type of benefits. Yet, because of the fragility of these populations, it is particularly important to find ways to identify these individuals in a unique and unambiguous way, while simultaneously ensuring that their privacy is protected. This requires devising an identity management system that minimizes the control of one single actor over the personal information of a refugee's population.

Hence, while it remains technology-neutral, the ID2020 Alliance has shown particular interest in blockchain technology, as a possible solution to provide digital identities in a way that is both traceable and immutable, and potentially not under the control of one single company or organization. One of the fundamental requirements defined by ID2020 for digital identities is, in fact, that identities remain portable, and that people retain control over their personal data by choosing with whom it can be shared and for what purposes.

Several non-profit organizations in the humanitarian sector are also involved in the definition of best practices and guidelines to ensure that people dealing with migrants and refugees respect their fundamental right of privacy and data protection. Core documentation has been developed in that regard, including the "Handbook on Data Management" (Blazewicz et al., 2012), the Privacy International's report (2018) on the "Humanitarian Metadata Problem,"[16] and the International Committee of the Red Cross' Handbook on "Data Protection in Humanitarian Action" (ICRC, 2017), which specifically addresses the additional privacy requirements that must be put in place when interacting with vulnerable persons. All these guidelines invite organizations providing humanitarian assistance to take all the necessary measures to protect the personal data of all concerned individuals, while focusing on the core humanitarian principles of "do no harm" and the promotion of human dignity.

[7] In 2016, the World Bank's Identification for Development (ID4D) Initiative issued a Strategic framework, recognizing the transformational potential of modern identification systems for the delivery of basic services and rights for the poor. The report is available at the following address: http://pubdocs.worldbank.org/en/215714605674816 55/April-2016-ID4D-Strategic-RoadmapID4D.pdf

[8] World Bank's 2016 ID for Development (ID4D) report showed that ∼1.5 billion people around the world (over 21% of the world's population) cannot prove their identity. See *Ibid*.

[9] *Ibid*.

[10] UNHCR, Statistical Yearbook, available at https://www.unhcr.org/en-us/figures-at-a-glance.html

[11] See e.g., Gleick (2014), describing the extreme drought in Syria as a driving factor for the 2011 civil war, and Werz and Conley (2012), associating the success of al-Qaida's recruiting strategies with the overall decline of agricultural and pastoral livelihoods.

[12] The UN's Global Compact on Refugees recognized that "climate, environmental degradation, and natural disasters increasingly interact with the drivers of refugee movements." According to the Internal Displacement Monitoring Centre, there were 18.8 million new disaster-related internal displacements recorded in 2017. While most disaster displacement linked to natural hazards and the impacts of climate change is internal, displacement across borders also occurs, and may be interrelated with situations of conflict or violence.

[13] https://id2020.org/

[14] https://id2020.org/technical-certification-mark

[15] See e.g., McMullen et al. (2019) analyzing the various blockchain-based initiatives for digital identity, and their various degrees of decentralization and privacy compliance.

[16] https://privacyinternational.org/report/2509/humanitarian-metadata-problem-doing-no-harm-digital-era

Yet, even if the organization collecting the data respects all of these privacy guidelines, any centralized institution holding such a large amount of personal data inevitably constitutes a single point of failure, which might inadvertently lead to significant data leaks. A true decentralized solution would enable people to maintain full control over their personal data (with a real self-sovereign identity solution), but the lack of a centralized database of identities would make it difficult to guarantee the "unicity" and "singularity" of these identities.

One identified solution to offer a persistent identity from birth, without the need for a centralized authority in charge of assigning a particular identifier to each person, is to rely on biometric data to generate a unique identifier (a biometric hash) associated to every individual. Indeed, in the absence of a centralized authority capable of ensuring that no same person registers twice for an identity, the only way to ensure the singularity of identifiers, without publicly disclosing any sensitive data about the individual concerned, is for these identifiers to be linked to cryptographically-hashed biometric information. This biometric hash can be used as a means of authentication, as it can be verified easily by comparing it with another biometric hash, but it cannot be used to retrieve the biometric information of the individual concerned.

Yet, while such a model is likely to provide important privacy benefits, it comes with the caveat that the singularity of an identifier is inversely correlated with the reliability of the system[17]. Indeed, unicity and singularity are a matter of degree: different identifiers with different characteristics may situate themselves on different points on that continuum. While biometric data could be used to create unique and unambiguous identifiers, whether or not they pass a sufficient threshold of singularity will ultimately depend on the degree of technological sophistication and the size of the population (Bhargav-Spantzel et al., 2010; Unar et al., 2014). We analyze below the benefits and the risks of these systems, in order to assess the extent to which they can be legitimately used for the purpose of refugee's identification and aid disbursement.

BENEFITS AND RISKS OF BIOMETRIC IDENTITY SYSTEMS

Using biometrics as part of an identity management system comes with a few advantages. If people can identify themselves through their biometrics, they no longer need to use passwords (often weak passwords which are easier to remember but very easy to breach). Insofar as a biometric is difficult to forge (or more expensive to forge compared to breaking weak passwords), biometrics may be relatively more secure than existing authentication systems. However, the use of biometrics within an identity management system may raise significant security and privacy risks, depending on how biometrics are used, stored, and permissioned (Prabhakar et al., 2003). For example, biometrics stored in centralized systems, without mitigating data access policies or security design measures, may be subject to greater security risk than if the data were stored locally on the user's device (Muller, 2010).

Hence, in recent years, there has been an increasing amount of research and initiatives exploring the use of decentralized infrastructures, mostly based on blockchain technology, to bootstrap new types of self-sovereign identity management systems (Baars, 2016; Jacobovitz, 2016; Tobin and Reed, 2016; Dunphy and Petitcolas, 2018) and combining them with biometrics as a means to ensure the singularity of identities within these systems (Hammudoglu et al., 2017; Garcia, 2018; Othman and Callahan, 2018).

Without going into the merits of these solutions, we describe below the basic operations and procedural aspects of these identity management systems, focusing on the key issues that must be taken into account when designing an identity system that relies on a blockchain-based infrastructure and on biometric information as part of the identification and authentication process.

Decentralized Infrastructure vs. Centralized Custody of Keys

All blockchain-based systems rely on a public-private key pair to record information (including, but not limited to, financial transactions) on a shared and decentralized ledger. Hence, one important aspect of any blockchain-based identity system is who ultimately possesses or controls the private keys necessary to execute a transaction. On that point, an important distinction needs to be made between the decentralized blockchain-based infrastructure, and the mechanism by which the blockchain-based identity system manages the keys associated with each individual entity.

A blockchain is decentralized insofar as its transaction history is immutably recorded and maintained by a distributed network of computer nodes, in order to prevent systemic theft (i.e., rewriting the transaction history to enable double spending). The decentralized nature of a blockchain network does not, however, apply to the custody and secure storage of the keys that control the individual wallets on that network (Hileman and Rauchs, 2017). Centralized control and storage of these keys is a major security hole that explains numerous high-profile cryptocurrency exchange heists. From a purely technical perspective (notwithstanding legal and contractual obligations), ownership of assets on the blockchain is equated with control of the assets, which is managed through the private keys associated with a wallet that contains the assets.

[17]Biometric information is normally stored in its raw form, rather than hashed, as hash functions require the exact same input each time. While hashing works well for inputs such as passwords that are exact in nature, biometric inputs are variable by nature; as such exact inputs cannot be guaranteed. For example, an iris photographed under slightly different lighting conditions will produce a different input such that the hashed results do not match exactly. Biometric inputs are compared against templates through comparing the number of stable bits extractable from each biometric scan. While it may be possible to hash a biometric input by reducing the number of stable bits required to the minimum, it would make the biometric authentication less reliable. If the number of stable bits required for a match is increased, reliability is improved; however, it will be more difficult to authenticate given the increased difficulty of achieving the required number of stable bits.

To the extent that cryptocurrency exchanges control the private keys associated with the wallets (or accounts internal to the exchange) containing customer funds, they also effectively control these funds, because custody of these keys ultimately implies full control of the funds stored in that account—much like physical paper cash (De Filippi, 2014). Hence, because the customer's private keys were not properly stored and secured in a decentralized fashion, these centralized exchanges rapidly became valuable "honey pots" attracting attackers (Gerard, 2017).

When marrying biometrics with cryptocurrency, it is important not to use biometric data as the seed of the private key unlocking access to cryptocurrency funds. Otherwise, anyone who can acquire access to an individual's biometric data would be able to derive that individual's private key, and therefore unlock the cryptocurrency funds. From a security and privacy perspective, such a system is more dangerous than an ordinary centralized cryptocurrency exchange, as biometric data contain the most sensitive and immutable personal identifying information (van der Ploeg, 2003). In short, even if a decentralized blockchain infrastructure like Bitcoin or Ethereum is used as the backbone of an identity system (De Filippi and Mauro, 2014), the security benefits of decentralization do not transfer insofar as custody of keys remain centralized without mitigating security design factors.

Identification vs. Verification

Next, when assessing an identity system, it is important to identify the types of information that must be provided at the different steps of the process, as individuals enroll into a particular identity system, and as they authenticate themselves within that system. We analyze below the various steps with regard to a biometrics-based identity system based.

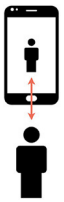
Enrollment

Enrollment is the process of creating a new user identity on the biometric system. Each user must provide relevant biometric samples (e.g., fingerprint, iris, or face) that will be captured by a biometric scanner or similar device. The collected biometric data will be used to generate a biometric template and biometric identifier, associated with personal information (such as demographic data) for subsequent authentication purposes (Araújo et al., 2005).

Authentication

Authentication is the process by which, after individuals have enrolled into the system, the system checks whether these individuals have the proper permissions to access a particular service or to benefit from a particular type of aid, by matching a new biometric sample against the biometric template created during enrollment (O'Gorman, 2003). The authentication stage can be subdivided into two different steps: *identification* and *verification*.

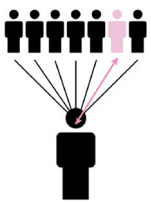
Verification



Verification is the process of verifying one's identity. It provides an answer to the question: are you who you say you are? It is a one-to-one matching process, whereby a new biometric sample is matched against an authenticated record. This is the case of using a fingerprint or face scan to access devices like computers or mobile phones. Currently, the standard practice is that the biometric record is stored locally and in encrypted format on the actual device (Schneier, 1999). Thus, neither mobile app developers nor device manufacturers have access to the template. The original scan used to create the template for matching purposes is destroyed, and so are the new scans made upon each new login, once the matching process is complete (Uludag et al., 2004).

Local storage of the biometric template on the device (rather than a central server) is a form of decentralized data storage, which can be further decentralized by breaking the biometric template into multiple pieces that must come together in order to be readable. This method protects privacy and improves security (Zibran, 2012).

Identification



Identification is the process of retrieving the identity of a particular individual, based on an identifier. It provides an answer to the question: who are you? It is a one-to-many matching process, whereby a new biometric sample is matched against many templates in an identity database in order to retrieve the specific identity it has been associated with (Jain et al., 2007).

Ideally, the sample scan should be destroyed once the transaction is complete. However, the original biometric template must necessarily be stored on a server, or be otherwise accessible to the operator of the identity system, for matching purposes. Therefore, as opposed to the verification process which can be done locally on a user's device, in the identification process, biometric templates need to be accessible online. In order to minimize security risks, it is thus important to identify mechanisms for secure decentralized storage and processing of data (Ganapathy et al., 2011), such as secure multi-party computation (Goldreich, 1998) or emerging solutions based on homomorphic encryption (Gentry and Boneh, 2009).

Individual Control vs. Organizational Control of Personal Data

Except for the case where the biometric template is stored locally on the user's device (mostly for verification purposes), in all other cases described above, the biometric and personal identifying information is not under the possession of the data subject, but rather that of the organizations that collect, store and administer the data for a particular identity system. While data protection regulations—especially in Europe—enable the data subjects to restrict the collection and processing of personal data (Tikkinen-Piri et al., 2018), once collected, such data might remain under the control of whoever owns the hardware (servers, devices) where the data is stored. The same is true for behavioral and social data that corporations collect about their users, which are statistically compiled as identity profiles that may be used for purposes of advertisements, alternative credit scoring, identity verification, and so on (Bygrave, 2012).

Privacy laws and data protection regulations provide some protection in terms of how information may be stored, used or collected. However, data protection regulations merely impose an obligation for data collectors and processors to obtain *informed and explicit consent* from the data subjects before they can engage in the collection or use of personal data for a particular purpose (Kosta, 2013). Some jurisdictions—such as Europe with the newly enacted General Data Protection Regulation[18]—have introduced additional rights, including the right to data portability[19] and the right to erasure[20] (better known as the right to be forgotten). Yet, where such protections do not exist, there is a risk that personal data (including biometric templates or samples) will remain siloed by the organizations that control them, with no real possibility for the data subject to request the deletion or the portability of such data—unless such organizations implement their own privacy policies that enforce these requirements.

Biometrics vs. Other Types of Identifiers

While biometrics provide interesting benefits to an identity management system, they are not devoid of any drawback. First of all, using biometric data to create a singular and unique identifier obliges individuals to identify themselves as one and only one persona—even when it is not necessary for a particular use case (Jain et al., 2004)—which may present significant privacy issues, especially in the case of political refugees.

Biometrics can also be significantly more problematic than traditional forms of authentication (e.g., passwords and other identifiers such PIN codes, hardware devices, etc.) because one cannot change his or her biometric data (Prabhakar et al., 2003). Importantly, biological information is effectively public information: we are leaving biological information everywhere, e.g., fingerprints, DNA, recordings of our gait, photographs of our faces or irises—from which advanced computer algorithms can extract a biometric template (Mordini and Massari, 2008). Fingerprints are easily stolen, copied, or lifted. Facial recognition can be easily spoofed through photographs or videos. Iris scans or behavioral biometrics such as gait may be more difficult or expensive to spoof or copy, but are not foolproof (e.g., contact lenses can fool iris scans). Accordingly, because of their inherently public nature, biometrics should only be used as the username (i.e., public key) rather than the password (i.e., private key). Whenever biometrics are used, some form of second factor authentication should be required, such as a PIN or a physical token, verification of photo ID or a physically present person (Rane et al., 2013).

Moreover, our bodies are subject to physical change. Iris scans become clouded due to cataracts. Fingerprints may disappear due to hard labor or burns. Gait may change due to aging, accidents, or illness. According to a study[21] by the National Institute of Standards and Technology (NIST), even in healthy people, the error rate for single iris scans can range from 2.5% to up to 20% in some cases—a significant percentage given the world's population of 7.5 billion people. As identity practitioners like Vinay Gupta have argued, because of the complex variation and nuance of biological forms, it is fundamentally impossible to rely on biometric measures as singular and unique identifiers for human beings.[22] Indeed, if biometrics are used as a universal identifier of one's identity and rights, the consequences for those in the three percent baseline error rate may be paralyzing and dire. For instance, in the case of India's biometric ID system, one study showed that 20 percent of the households in Jharkand state had failed to get their food rations due to biometrics errors—which is five times higher than the failure rate of ordinary ration cards[23].

Finally, because of the public perception of biometrics as being more "scientific" and therefore more authoritative, the downside errors of biometrics is often overlooked. Yet, if a biometric identifier is used as the backbone of an identity management system used for the protection of fundamental rights and privileges, it cannot fail disastrously, even if the probability of a failure is very small. Ideally, a properly functioning identity system must be resilient against low probability but highly consequential negative events and gains value from increased input and interactions with the world. However, an identity system that relies on biometrics as the only authoritative identifier is not only a brittle and fragile system (Friedman et al.,

[18] The General Data Protection Regulation (GDPR) 2016/679 is a regulation in EU law on data protection and privacy for all individuals within the European Union and the European Economic Area.

[19] Article 20 of the GDPR stipulates that: "The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided"

[20] Article 17 of the GDPR stipulates that: "The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies: the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed."

[21] https://ws680.nist.gov/publication/get_pdf.cfm?pub_id=910385

[22] https://medium.com/humanizing-the-singularity/a-blockchain-solution-for-identity-51fbcae94caa

[23] https://www.independent.co.uk/news/world/asia/india-tech-fingerprint-eye-scan-id-food-benefits-bank-accounts-a8297391.html

2011), it is also highly problematic from a cybersecurity and privacy perspective (Prabhakar et al., 2003; Campisi, 2013)—which is particularly relevant for vulnerable populations such as migrants and refugees.

SELF-SOVEREIGN IDENTITY AND CREDENTIAL MANAGEMENT SYSTEMS

The notion of self-sovereign identity has emerged in the past few years, although there is no agreed upon definition yet on what the terminology really means (van Wingerde, 2017). On a general level, self-sovereign identity is intended to preserve the right for the selective disclosure of different aspects of one's identity and the various components thereof, in different domains and contextual settings. This right should apply irrespectively of whether these aspects and components have been issued by a particular government, company, or organization. More specifically, self-sovereign identity also refers to the idea that individuals shall retain control over their personal data and, to a certain degree, over the representations of their identities (or personas) within a particular identity management system. This requires giving them the ability to establish (and control) who has the right to access specific pieces of information about them, with a high degree of granularity (Der et al., 2017).

From a technical perspective, self-sovereign identity is generally regarded as a new paradigm of online identity management, whereby individuals and entities can manage their identity-related information (i.e., identifiers, attributes and credentials, or other personal data) by storing them locally on their own devices (or remotely on a distributed network) and selectively grant access to this information to authorized third parties, without the need to refer to any trusted authority or intermediary operator to provide or validate these claims (Mühle et al., 2018). This enables greater control over personal identifying information, or other relevant data about an individual or entity. Because digital identifiers can be in a variety of formats, an important requirement for a global identity system is the establishment of technical standards for interoperability. We describe below the most prevalent standard, the Decentralized Identifier (DID), that we mentioned earlier in the paper.

Open Source Digital Identity and Verifiable Claims Web Standards

The World Wide Web Consortium (W3C) is a technical standards body for the open internet, working on a decentralized identifier (DID) standard.[24] DIDs are a new type of identifier for verifiable, self-sovereign digital identity that is universally discoverable and interoperable across a range of systems.[25] The DID standard is supported by the Decentralized Identity Foundation, a consortium of companies that are developing and building applications using the DID standard, including Microsoft, IBM, Hyperledger, Accenture, Mastercard, RSA, and all the major blockchain identity and data companies such as Civic, uPort, BigChainDB, Sovrin, and many others[26].

DIDs are URLs (i.e., unique web addresses) that resolve to a DID Document, which provides information on how to use that specific DID[27]. For example, a DID Document can specify that a particular verification method (such as a cryptographic public key or pseudonymous biometric protocol) can be used for the purpose of authentication. The DID document might also reference a series of service endpoints, enabling further interactions with the DID controller. For instance, a DID can reference the location of associated personal data, which a requester would need to ask the DID controller for permission to access (McMullen et al., 2019).

A DID by itself is only useful for the purpose of authentication. It becomes particularly useful when used in combination with verifiable claims or credentials—another W3C standard that can be used to make any number of attestations about a DID subject (Dunphy and Petitcolas, 2018). These attestations include credentials and certifications that grant the DID subject access rights or privileges. For example, a verifiable claim can attest that an individual has been Know-Your-Customer (KYC) approved and therefore eligible to open a bank account, that the same individual has been certified as eligible to drive, or authorized to access certain programs as a system administrator (Aydar and Ayvaz, 2019).

A verifiable claim contains the DID of its subject (e.g., a bank customer), the attestation (e.g., KYC approval), and must be signed by the person or entity making the claim using the private keys associated with the claim issuer's DID (e.g., the bank). Verifiable claims are thus methods for trusted authorities, such as banks, to provably issue a certified credential associated to a particular DID. DID claims remain under the control of the DID subject and can be used to prove a particular attribute of the DID subject, independently from a certificate authority, an identity provider or a centralized registry (Baars, 2016). Proving to be the actual subject of that DID (through a specified authentication method) will enable an individual or entity to benefit from access privileges associated with these credentials.

While DIDs are independent of and do not require blockchain technology, they are designed to be compatible with any distributed ledger or blockchain network. Since a DID may be associated with a particular private/public key pair used to sign identity claims, it is possible to associate that key pair (i.e., the key pair linked to the DID) with key pairs used to sign financial transactions on a blockchain. Most importantly, the DID specification also makes it possible to associate particular methods to a DID, which specifies the procedures for key registration, replacement, rotation, recovery, and expiration. Several method schemes have been implemented so far that leverage the resilience and tamper-resistance of blockchain

[24]W3C is led by internet industry pioneer Tim Berners-Lee, who invented the World Wide Web. W3C has 479 members including all the major internet and technology companies such as Amazon, Apple, Boeing, Cisco, Microsoft, Google, Facebook, Alibaba, Tencent, Baidu, along with research universities and governments. See https://www.w3.org/

[25]See W3C DID primer for introduction: https://github.com/w3c-ccg/did-primer

[26]https://identity.foundation/

[27]https://w3c-ccg.github.io/did-spec/

technology to manage DIDs (e.g., BTCR DID, Blockstack DID, Ethereum ERC725 DID)[28]. The W3C group is working to ensure technical interoperability between different DID methods.

It is important to note, however, that given the transparency and immutability of a blockchain, personal information should never be stored on the blockchain itself (De Filippi, 2016). Yet, a blockchain can be used to track permissioning and access of personally identifying data that is stored off-chain, thereby creating an auditable trail of information access. Therefore, in addition to the standardized DID methods, a blockchain can also be used for the recording and eventual revocation of claims or attestations, for the granting and revocation of access to personal data stores[29], and other functions that may be specific to particular identity system (e.g., claims filed and resolved as part of a dispute resolution system regarding false attestations).

A Road-Map Toward Self-Sovereign Identity

The road toward true self-sovereign identity is still long, as we are only at the early stages of understanding how to implement a digital identity system that provides full control and autonomy to the individuals. Yet, in light of the refugee crisis in Europe, and the increasing number of displaced people who lack a formalized form of identification, today—perhaps more than ever—the quest toward self-sovereign identity has become of crucial importance.

As described earlier, self-sovereign identity solutions are designed to give individuals control over their own identity—that is, people should have the possibility to decide precisely what information to disclose about themselves, to whom, and under what circumstances. Under a self-sovereign identity model, identity providers should not have the possibility to prevent individuals from exercising basic human rights, such as the right to be oneself, the right to freedom of expression and the right to privacy. While this does not necessarily require individuals to be the sole holders of any information regarding themselves, an important precondition for self-sovereign identity is that digital identities are not locked into any given platform, nor controlled by a given operator, but rather remain portable and interoperable across multiple platforms, so that individuals are free to choose the identity operator that they trust the most, and to move from one operator to another, if so desired.

While a precise definition of what constitutes a self-sovereign identity does not currently exist, a series of criteria have been identified as the underpinning principles of self-sovereign identity[30]. These principles can be regarded as a preliminary benchmark to assess existing self-sovereign identity solutions:

1. **Existence**: individuals must have an independent existence, independently of the digital identifiers that merely serve as a reference to them.
2. **Control**: individuals must control their identities, they should always be able to refer to it, update it, or even hide it—even if others can make claims about these identities.
3. **Access**: individuals must have access to all the data related to their identities, and should be able to retrieve their claims whenever needed.
4. **Transparency**: systems and algorithms used to administer and operate digital identities must be open and transparent, with regard to both their operations and maintenance.
5. **Persistence**: identities must be long-lived, preferably they should last forever, or at least for as long as the user wishes to maintain them.
6. **Portability**: information and services about identity must be transportable, and not be held by a single third-party entity, even if it's a trusted entity.
7. **Interoperability**: identities should be as widely usable as possible, as opposed to being framed only to work in siloed environments.
8. **Consent**: individuals must agree to the use of their identities, sharing user data must only occur with the consent of the data subject.
9. **Minimization**: disclosure of claims must be limited to the minimum necessary to accomplish the task at hand
10. **Protection**: the rights of users must be protected at any cost, even if doing so would go counter to the interests of the identity providers.

Most digital identity projects will not meet all of these criteria—and many do not even purport to qualify as "self-sovereign" identity projects—we will discuss in this paper two case studies that make use of biometrics in combination with blockchain technology to provide users with a certain degree of sovereignty over their digital identities. The first case study is the *Kiva Protocol*, which focuses on identity for credit scoring and secure sharing of credit history amongst microfinance institutions. The second case study is the World Food Programme's *Building Blocks* and its biometric identity solution for delivering services to beneficiaries in need—particularly in providing better delivery of services to beneficiaries served by multiple UN agencies.

These two initiatives were chosen because of their higher degree of technological readiness with respect to other alternatives, their credibility and their potential impact in terms of future large-scale deployment, and, finally, because of the previous experimentations they have undertaken, which enabled us to collect valuable data points concerning the extent to which their current implementation fulfills the criteria of a self-sovereign identity system.

As the following sections describe, these two projects have prioritized specific principles of self-sovereign identity that are most relevant to their use cases. In both of the cases, it appears that the identity solutions focus, first and foremost, on principles relating to interoperability and the secure sharing of identity claims between parties. The principles of minimization, consent, portability, and persistence are also given significant importance. The use of a blockchain ledger is useful because it enables data to be shared securely across multiple parties, and parties must be

[28] A list of currently available DID method schemes is available at: https://w3c-ccg.github.io/did-method-registry/

[29] For a general overview of the different blockchain-based self-sovereign identity solutions and their characteristics, see (McMullen et al., 2019).

[30] The Path to Self-Sovereign Identity, written by Chris Allen and the Rebooting Web of Trust community: https://github.com/WebOfTrustInfo/self-sovereign-identity/blob/master/ThePathToSelf-SovereignIdentity.md

granted permission in order to access and append information to the blockchain. From an identity perspective, a persistent and portable digital identity and digital history is highly valuable to vulnerable populations who are often on the move. The validity of the attestations, especially from trusted organizations such as Kiva and UN agencies, are important for the identity subject to establish or re-establish credibility and access to resources.

However, the principles of control and access remain difficult to achieve from a technical perspective in developing economies, as smartphone penetration and technical knowledge necessary for self-custody is still nascent. The lack of proper connectivity and hardware infrastructure (e.g., while most refugees do have a mobile phone, they do not always have a smartphone) is a key obstacle to overcome in the roadmap toward self-sovereign identity. Both Kiva and the Building Block initiatives therefore had to implement custodial models for their identity solutions, significantly reducing the degree of control that individuals can exercise over their digital identities. However, that may change over time as smartphones become cheaper and users become more technically knowledgeable. In any event, both case studies provide valuable lessons concerning the multiple obstacles associated with the implementation of self-sovereign identity solutions in the humanitarian context, and the different approaches adopted by each of these initiatives, as an attempt to overcome these obstacles in the short term while focusing on the immediate user needs.

KIVA CASE STUDY: SOLVING FOR CREDIT HISTORY[31]

Kiva[32] is building an identity protocol that is expected to be rolled out across the whole country of Sierra Leone—this is a testament to the strength of the programme and the significance of provisioning vulnerable persons with a digital identity system.

Kiva is based on the DID and credentials model described above, using Hyperledger Indy as the underlying blockchain layer. It relies on a credential-based identity system, wherein the basic identifier is a public/private key pair, to which multiple claims and attestations can be associated. In the Kiva protocol, issuers of verifiable credentials are called "trust anchors" who have real world reputations at stake. The Kiva identity protocol is currently designed as a private permissioned system, whereby all trust anchors must be approved by Kiva and/or the Sierra Leone government in order to issue credentials, sign attestations, and read identity claims. In the future, trust anchors may be broadened to include NGOs, technology companies such as Facebook and Google, and other organizations that can provide information relevant to a particular identity[33].

[31] Most of the information in this section has been drafted as a result of several calls and interviews with kevin o'brien and aaron goldsmid from kiva.

[32] https://www.kiva.org/protocol

[33] A future identity protocol may enable permissionless trust anchors that do not need to be centrally approved *ex ante*; or else, trust anchors may be automatically approved according to a set of programmable rules e.g., number of credentials or types of credentials associated with a particular trust anchor to establish their reputation.

Currently, trust anchors are limited to the Sierra Leone government bodies and microfinance institutions, because of the immediate goal of solving the problem facing the microlending industry—whereby many constituents are ineligible for loans due to lack of any formal identity and history (data for underwriting loans). In fact, the government of Sierra Leone, through the influence of the Central Bank which issues bank licenses, will *require* that all microfinance institutions, banks and other financial institutions participate as credential issuers for Kiva's identity system. This is particularly relevant for microlending in developing economies that do not have national credit bureaus, making it difficult for lenders to check cross indebtedness. Without the ability to check the total indebtedness of a borrower, it is difficult to properly price default risk and underwrite these loans.

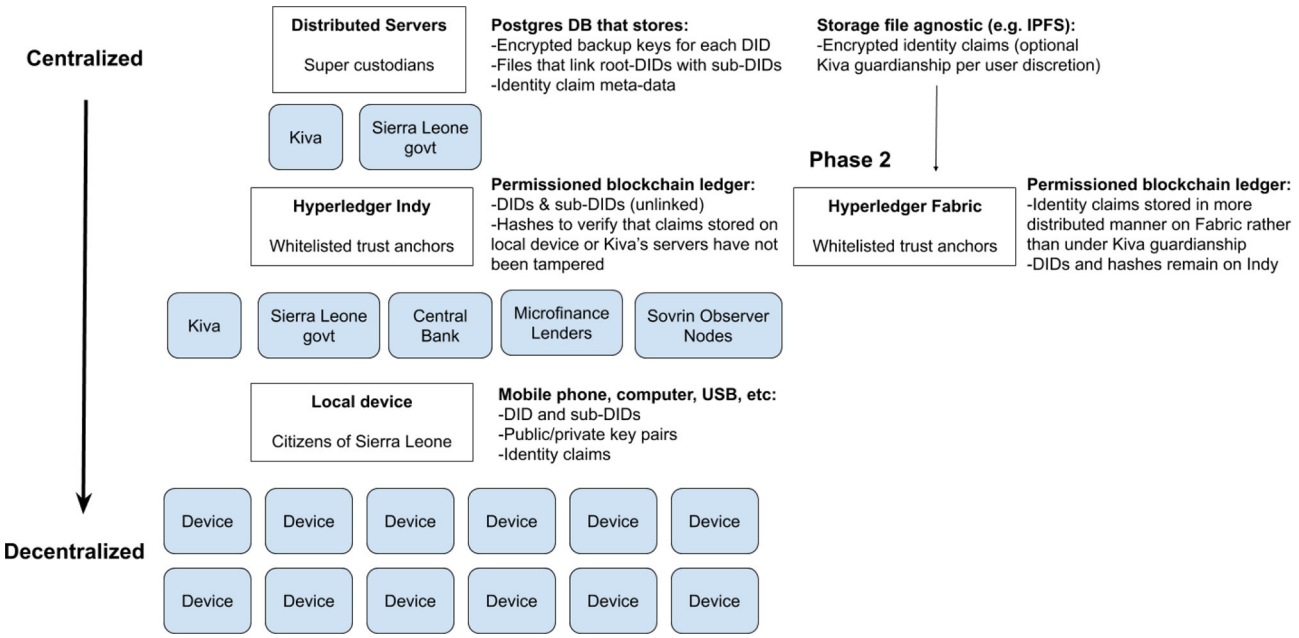
The Kiva protocol functions like a credit bureau with greater privacy and control by the individual compared to traditional credit bureaus. The credit profile, comprised of linked credit attestations and claims, is portable by the individual, rather than locked within a centralized credit bureau. Importantly, individuals have control over who can access their profiles, whereas currently anyone can perform a credit check without permission. Under the Kiva model, an individual may decide whether to provide a lender access to her credit history.

Kiva will provide every Sierra Leone citizen eligible for a government issued ID with a DID and associated public/private key pair to sign identity claims, along with a first attestation from the Sierra Leone government (in the form of a verifiable credential containing hashes of the citizen's biometrics and other government-issued identifiers). In the Kiva framework, biometrics are simply another attribute that is attached to the DID, similar to a date of birth, place of birth, or any other piece of identity information. This reduces much of the systemic risk of using biometrics as an exclusive form of identification, as described above. In this system, because biometric data do not serve as an identifier, biometrics will be used primarily for verification rather than for identification purposes.

Anyone seeking to access the information linked to that profile must request the DID subject (i.e., the Sierra Leone citizen) to grant permission. The Hyperledger blockchain is used to record that a third party (as identified by its public key linked to a DID) requested, was granted, and was eventually revoked access to the relevant verifiable claims or credentials, according to timestamps. Each citizen will have a root-DID that maps to an unlimited number of sub-DIDs that are generated for each new loan transaction or relationship with a lender. Sub-DIDs can also be created for different purposes i.e., each sub-DID represents a different persona or profile. The use of sub-DIDs enables a degree of privacy (see below section on privacy).

Kiva Protocol Architecture

The following is a high-level architecture of the Kiva identity protocol:



In an ideal model, all sensitive information, such as private keys, the files that link root-DIDs with sub-DIDs, identity claims and other information are stored only locally on devices controlled by the identity subject, such as mobile phones and computers. Thus, control and storage of personal information is structurally decentralized. In developing countries, however, this will take time as smartphone penetration is still low (though growing rapidly in many markets) and many people may not necessarily own individual devices i.e., a phone may be shared amongst a family. Currently, it is not possible to securely store private keys on feature phones. Therefore, it is likely that third parties such as non-profits or commercial businesses may serve as proxies that help manage private keys or shared devices. Ideally, private keys are never shared and remained locked in wallets on shared devices, whereby users can unlock their individual private keys using biometrics, PIN or password when they access the shared device.

Even if non-profits and other community organizations serve as trustees or proxies to help users manage their private keys, backups of identity claims and private keys will be necessary. In light of the practical difficulties of managing the public/private key pairs associated with a particular DID, the Kiva identity protocol deploys a guardianship model, whereby Kiva and the Sierra Leone government serve as the super custodians in the system. Kiva will escrow the key pairs on behalf of the identity subject, who may take the key pairs out of escrow at any time. Under Kiva's guardianship model, backup keys in custody are encrypted and can only be restored through a multi-factor process e.g., biometrics and/or PIN.

Kiva's servers also store the data files that map the links between root-DIDs and related sub-DID, as well as backup copies of encrypted identity claims (with accompanying meta-data) on a separate data storage format such as IPFS. In the next phase of the protocol, the encrypted identity claims may be stored in a more distributed manner on a permissioned ledger such as Hyperledger Fabric, which is better designed to store data, whereas Hyperledger Indy is fit-for-purpose for validating DIDs.

The most private and sensitive data is held in guardianship on Kiva's distributed servers in a Postgres database. A local copy of the database (or parallel database) may be maintained by the Sierra Leone government, pursuant to Sierra Leone data localization regulations that require sensitive citizen data to be stored in-country.

Beneath the Kiva guardianship layer is the private permissioned blockchain ledger running on Hyperledger Indy. The Central Bank of Sierra Leone would be a permissioned node, along with Kiva and the Sierra Leone government. Because the Central Bank is requiring all lending institutions to report loan transactions on the protocol, the microfinance lenders and other financial institutions that fall under the Central Bank's mandate will be required to register as nodes. In addition, other parties such as non-profits, may apply to be Trust Anchors or Stewards (Sovrin observer nodes), which helps increase the security and resiliency of the ledger by diversifying nodes away from entities domiciled in Sierra Leone.

The nodes store copies of the unlinked DID and sub-DIDs, as well as hashes of the associated identity claims. As noted above, Hyperledger Indy is not designed to store actual claims data, which identity subjects will have the choice to store in Kiva's guardianship, and later those claims can be migrated to Hyperledger Fabric, which is built to support claims data, as described above.

Interacting With the Kiva Protocol: Step-by-Step

We describe here the intended step-by-step operations of the Kiva protocol, and how a Sierra Leone citizen might interact with the Kiva protocol, once fully deployed. The Sierra Leone government will deploy campaigns to enroll citizens into the identity protocol. Citizens will register at polling stations, where they will receive both a physical ID card with biometrics and a digital ID, in the form of a DID and associated private keys held in a wallet, ideally on the individual's device. In many cases, as described above, the individual may not own a phone or have a phone with the capability to hold private keys in a wallet. In this case, the keys and future identity claims will be held in guardianship by Kiva. The government of Sierra Leone will make the first attestation by signing an identity claim that the individual is a citizen of Sierra Leone with official identity information such as biometric string, date of birth and other data.

When the individual, whom we will call Mary, goes to the local microfinance lender to ask for a loan, the bank will first ask for Mary's identity claim signed by the government of Sierra Leone (the official state ID). Mary will access an application (either on her phone or on a device at the bank) that grants the bank permission to validate the government's signed claim. Ideally, to preserve privacy, a bank does not actually read the contents of the claim (e.g., the biometric, the date of birth) if such information is not actually relevant for purposes of KYC or credit underwriting. All the bank needs to know is that the government has signed a valid claim attesting to Mary's identity, which fulfills the bank's minimum KYC obligations.

Next, the bank will ask Mary for permission to disclose her credit history. If Mary says yes, Mary will then unlock her identity claims using her private key. The bank will then validate the identity claims against the hashes in Hyperledger Indy to confirm that the identity claims are both complete and authentic. If there is an error, the bank will receive a failure message.

If Mary is unable to use her own device to manage identity claims and keys, the bank will ask for permission to retrieve the identity claims from Kiva's servers directly. In order to sign this permission using her keys in Kiva's custody, Mary would need to provide a second factor authentication such as her biometrics or PIN.

Once a loan is approved, the bank would sign identity claims relating to the loan disbursement and repayment. Mary would receive messages to her mobile phone application informing her that the bank is writing a claim e.g., regarding repayment, and Mary could accept this action[34]. The claim would be sent to Mary's device, if she chooses to only keep data on her local device; or else the claim would be encrypted and stored in Mary's wallet in guardianship on Kiva's servers (Kiva may also store a backup copy if Mary so chooses even if she manages her data on her own device).

Mary may also initiate a dispute resolution action if she believes the bank has written an incorrect claim or failed to provide a claim for a repayment. The dispute resolution process will likely be off-chain, whereby Mary would file a ticket with the facts to be decided by an arbitral body. If the arbitral body decided in Mary's favor that she did indeed pay the bank in cash for her monthly installment, the arbitral body would then require the bank to sign such a claim, or else the arbitral body could sign such a claim with its own keys.

Where loans are made and repaid in cash, Mary would need to trust her bank to make the repayment claim. She would likely receive a physical receipt for her cash repayment, which she could present to the bank to request a repayment claim (or to an arbitrator if her bank fails to do so). In a future model, if the loans were disbursed as digital currency, disbursements, and repayments could be automatically recorded as identity claims, with the blockchain transactions appended as proof of payment.

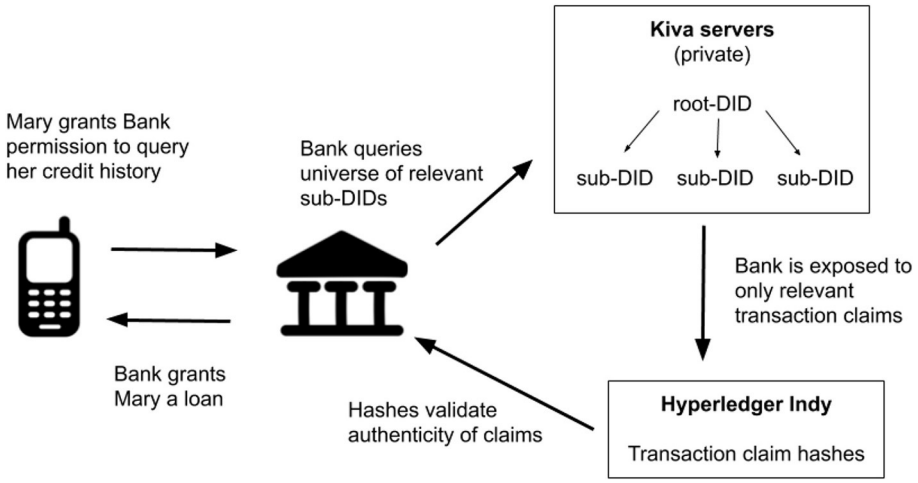
Privacy Considerations vs. the Problem of Selective Disclosure

In order to maintain privacy and reduce fallout from security breaches, the Kiva protocol strives to operate under the principles of zero knowledge proofs, whereby only the absolute necessary information is exposed and measures are taken to ensure that no one can seek information in the system without permission. Accordingly, each loan that Mary takes will be associated with a new sub-DID, rather than directly tied to her root-DID. This prevents banks from being able to monitor future credit activity tied to a root-DID without asking the identity subject for permission, as future credit transactions will be associated with newly generated sub-DIDs, and banks do not have access to the file that maps sub-DIDs to the root-DID.

Privacy is countered with the problem of selective disclosure, whereby lenders must check for cross-leverage. During the underwriting process, Mary's bank can see the full credit history across her sub-DIDs because during the validation process, the bank will first query Kiva's servers to get the universe of sub-DIDs tied to Mary's root-DID. As described above, the file that maps sub-DIDs to a root-DID is only available on Kiva's servers. However, at no time is the bank exposed to the actual sub-DIDs or root-DID; the bank is only exposed to the transaction claims associated with the sub-DIDs. The bank will then proceed to match the transaction claims against the hashes in Hyperledger Indy to authenticate the claims, as described above.

Even in the case of a fully self-sovereign identity system, not all data will be owned and controlled by the individual, as some of the data may be produced and maintained by third parties making attestations. For example, a bank will retain control over its own records regarding an individual's lending history with that bank. However, compared to a centralized credit bureau, information will not be centrally aggregated and communicated to a single operator. Data can remain stored by third parties, while the associated attestation (in the form of a verifiable claim) is assigned and controlled by the individual and stored on the blockchain. Hence, even though the citizens of Sierra Leone may not control all the information regarding them, they nonetheless

[34]Initially, Mary will give permission at the outset for her bank to write *all* claims related to her loan for the duration the loan remains outstanding. In the future, Kiva hopes to provide even greater control to users (especially as technology penetration improves), such that Mary would be able to grant permission for *each* claim that the bank wishes to append to her profile.



control the set of verifiable credentials that represent their attributes, which they can freely combine into a useful identity or set of profiles and personas.

Finally, it is important to note that Kiva's identity system is, at its root, a repository of verifiable claims data, which does not discriminate against politically sensitive identity claims. While it has been designed for Sierra Leone, the same identity system may be applied, for example, to Syrian refugees, allowing the Syrian government to issue attestations concerning the identity of a particular refugee, with a signature and time stamp. If the Syrian government that issued the identity no longer exists, the refugee will nonetheless be able to prove his or her identity at that particular point in time.

In returning to our list of self-sovereign identity principles, the Kiva identity system focuses first on consent, interoperability, and minimization. This serves the primary use case of enabling microfinance institutions to share information and create a persistent record of credit history, in a way that still preserves the privacy of the borrower by revealing only the necessary information for a microfinance institution to make a decision. While most users will not be self-custodying their identity information from the outset due to technical challenges, the system is designed such that users may opt out of Kiva serving as a super custodian. Over time, self-custody and control will become more prevalent, and identities remain globally portable and persistent.

WORLD FOOD PROGRAMME CASE STUDY: SOLVING FOR OPTIMIZATION AND HARMONIZATION OF AID ACROSS U.N. AGENCIES[35]

Background

The World Food Programme (WFP)[36] is the food assistance branch of the United Nations and the world's largest humanitarian organization addressing hunger and promoting food security. WFP provides food assistance to more than 80 million people in more than 80 countries.

In the past several years, the trend has been to enable the people served to make their own purchasing decisions through Cash-Based Interventions (CBI) rather than in-kind food distributions. In 2018, WFP distributed more than USD 1.7 billion in CBI, more than half of the global cash aid distributions[37]. In the right conditions, CBI programs can be more cost-effective and beneficial to the local economies as well as providing an increased element of dignity to the people served.

WFP has pioneered innovation amongst UN agencies, recognizing the potential for blockchain technology in CBI as fourfold: (1) improved efficiencies such as reductions in costs and risks and enhancements in accountability and control, (2) creating a unified view of the people served thereby reducing duplication and fragmentation, creating opportunities for optimization and harmonization, and linking various aid actors through a single connection to the blockchain, (3) multiplying the redemption options (such as ATMs, food stores, health networks, and schools) available to the participating organizations and the people served, and (4) paving the way for blockchain based digital identities by demonstrating the underlying technology in practice and bringing key stakeholders together around a neutral blockchain network.

Building Blocks

In this section, we describe WFP's blockchain-based CBI project called "Building Blocks."[38] Building Blocks was born in January 2017 with a 100-person Proof-of-Concept (PoC) in Pakistan's Umerkot village. At the time, the aim was to demonstrate that blockchain can be used beyond the cryptocurrency application.

For the PoC, beneficiary accounts were created on the blockchain and loaded with tokens representing cash or food and each beneficiary was assigned a random identifier between 1 and 100, which was linked to their public key one-to-one. To redeem their entitlements, beneficiaries would present themselves at cash

[35]MOST OF THE INFORMATION IN THIS SECTION HAS BEEN DRAFTED AS A RESULT OF SEVERAL CALLS AND INTERVIEWS WITH HOUMAN HADDAD FROM THE WORLD FOOD PROGRAMME.

[36]https://www1.wfp.org/overview

[37]https://www.economist.com/free-exchange/2014/03/03/giving-generously

[38]https://innovation.wfp.org/project/building-blocks

or food merchants and provide their random identifier. The merchant would then insert the beneficiary's identifier along with the redemption amount into a web application. The web application would send the request to Building Blocks which would then send a One-Time Password (OTP) to the beneficiary's feature phone via SMS as the authentication mechanism. The beneficiary would then provide the OTP to the merchant who would insert it into the web application and send it to Building Blocks. If the OTP was valid, Building Blocks would check the requested redemption amount against the available blockchain entitlements and, if sufficient, trigger the beneficiary private key held in custody to record a transaction and send a confirmation back to the merchant. Upon seeing the confirmation, the merchant would distribute the requested quantity of cash or food to the beneficiary. WFP would then, based on the Building Blocks record, determine the amount owed to each merchant and settle with them directly.

For the PoC, Building Blocks used the public Ethereum blockchain. This decision was based on the fact that public chains are self-sustaining through crypto-economic incentives and a public network of validators, and therefore not dependent on WFP or the UN. However, the project team observed that major public chains have low transaction throughput and expensive transaction costs due to the prevalence of the Proof-of-Work (PoW) consensus mechanism, which is based on computational power in order to secure transactions to the public ledger.

Jordan Implementation

Having demonstrated the concept of using a blockchain ledger, and incorporating the learnings from the PoC, in May 2017 Building Blocks initiated a large-scale pilot with 10,000 Syrian refugees in Jordan. The concept was similar to the Pakistan PoC. However, for the Jordan pilot, Building Blocks switched to a private, permissioned blockchain using the Parity Ethereum client with a Proof-of-Authority (PoA) consensus algorithm.

The private PoA network provides Building Blocks with a very high transaction throughput at no cost per transaction. The private network also provides higher assurances for data protection privacy. The main downside of the private network is that it is not self-sustaining. However, the smart contract code is identical between private and public networks. Therefore, when the public networks have adequately addressed the throughput, cost, and privacy issues, Building Blocks can switch by merely copy-pasting its code. Another downside is that a private network is less resilient and tamperproof than public networks due to the fewer nodes. However, with each additional independent node on the blockchain, a private chain becomes increasingly closer to the characteristics of public chains in terms of resilience and immutability.

In contrast to the Pakistan PoC whereby authentication was provided through OTP SMS, in Jordan Building Blocks integrated with the existing iris biometric authentication system enabled by the UN Refugee Agency (UNHCR)[39]. Through Building Blocks, refugees only need to scan their irises at the point-of-sale to receive food assistance. All transactions are recorded on a private blockchain-based infrastructure, used as a registry to calculate the balance of every refugee, as well as the amount of funds that must be disbursed by the WFP to the relevant merchants.[40] The advantage of this system is that beneficiaries can access and transfer funds by merely presenting themselves in front of the biometric-based identification system, without the need for a device such as a mobile phone. Indeed, given the precarious situations of Jordan refugees, it is not possible to assume constant internet connectivity or that beneficiaries will always own sufficiently sophisticated phones to handle key management. Facilitating seamless access to critical resources such as food or funds is particularly important for refugees in critical need.

Like Kiva, WFP faces issues with end user smartphone ownership and data connectivity. Hence, Building Blocks also has a guardianship model for custody of keys used to sign transactions. WFP functions as a custodian of the beneficiaries' private keys, which, through the biometric iris authentication, are triggered to sign blockchain transactions related to CBI. Like the Kiva model, the WFP model is also designed to enable self-custody should a user elect to do so when sufficient infrastructure is in place to make this feasible e.g., availability of affordable smartphones with key management capabilities. Eventually, the aim is to provide all beneficiaries with a new set of public-private key pairs (which they will create and have full control over) and transfer their aid credits to these new wallets.

As noted earlier, WFP's Building Blocks uses the UNHCR's Biometric Identity Management System (BIMS)[41] for authentication. Biometric data in BIMS may include original digital scans (such as the iris photographs), feature sets (i.e., biometric template abstracted from the digital scans), and the reduction of feature sets into a data string that functions as a unique identifier. During the registration process, UNHCR collects an individual's biometrics and associates the biometric data (reduced to a data string) with a unique random identifier in the BIMS database. Individuals are then grouped into family units (as a second level abstraction), each with their unique identifier (a 12 characters string).

Authentication in the context of the UNHCR cash aid system requires a beneficiary to provide an iris scan at the point of sale (POS) for every transaction. The process operates as follows: first, the biometric system at the POS is used to collect the biometric data through an iris scan. The scan is then converted to a template and communicated to the UNHCR and matched against the universe of templates in the BIMS database to retrieve the unique identifier associated with the beneficiary's family unit. This identifier is then sent to the WFP's Building Block system to retrieve the public-private key pairs associated with that identifier. The public key will be used to check if the beneficiary's balance is sufficient to make the transaction. If the balance is sufficient to cover the transaction, the private key will be triggered

[39] https://www.unhcr.org/en-us/

[40] https://www.technologyreview.com/s/610806/inside-the-jordan-refugee-camp-that-runs-on-blockchain/

[41] https://www.unhcr.org/en-us/protection/basic/550c304c9/biometric-identity-management-system.html

to sign the transactions on the blockchain, on behalf of the beneficiary. Each communication leg in the entire process is end-to-end encrypted.

For the time being, the system has implemented a series of best practices to mitigate the risk of centralized biometrics, by separating the custody of keys (done by the WFP) from the registry of biometric information linked to the individual's identity (managed by the UNHCR). Hence, from a privacy and security standpoint, WFP's Building Blocks incorporates the necessary safeguards to ensure that the merchant, the bank, the payment processor, the payment network, and other intermediaries are not exposed to information that is not relevant to their function. Indeed, the POS payment processor simply needs to know whether an individual has been enrolled in the system and whether the corresponding account balance is sufficient. It does not need to know the real-world identity, nor even the exact account balance of that individual[42].

Moreover, for reduced security risks, the UNHCR does not store any personal identifying information (such as name, nationality, birthdate, sex, family relations, etc.) together with the biometric data in the BIMS database. All biometrics data is securely stored and completely segregated from any other personal information. Likewise, BIMS does not store the information regarding the beneficiary's private keys—which are only accessible from the WFP's Building Blocks system. The privacy of refugees is therefore protected, since the WFP does not know the actual identity of the individuals whose transactions it processes, and the UNHCR does not have access to the transactions of the individuals it identifies.

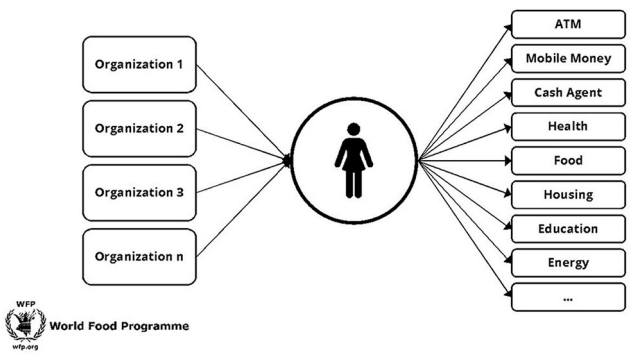
Based on the success of the pilot, in January 2018, Building Blocks was scaled to serve all 106,000 Syrian refugees assisted by WFP in the Jordan camps. It is currently the largest implementation of blockchain technology for humanitarian aid in the world. To date, Building Blocks has processed USD 60 million of CBI through 3 million transactions and saved USD 900,000 in banking fees [43].

Next Steps

Everything described in the previous sections could be achieved with traditional databases. However, as blockchain is a relatively new and often theoretical concept in the humanitarian aid world, Building Blocks was a first step in demystifying some aspects of blockchain technology by demonstrating how the technology works at scale in the humanitarian context. As such, the Building Blocks programme was one of the first of its kind.

Having achieved that preliminary goal, Building Blocks now aims to take the next step by welcoming new members to the network, in order to facilitate seamless interaction with a variety of different agencies. Non-Governmental Organizations (NGOs) have particular security requirements in humanitarian contexts, and international NGOs are often struggling to reconcile the collection of large swathes of personal data for the issuance of digital identities across multiple agencies. In the Jordan refugee camps, for example, more than 45 organizations assist the same beneficiaries. Yet, the various systems are not meaningfully connected and interoperable. This results in duplication of effort and a somewhat fragmented view of the people served, who need to repeatedly disclose their personal information as they move between agencies.

If these organizations channeled their entitlements to each beneficiary's public key, there would be a unified view of the people served, creating opportunities for optimization and harmonization. Program designs and needs targeting could also become more equitable. Furthermore, all actors could be linked through a single connection to the blockchain, and the various outlets (such as food, cash, health, and education) could be combined. The elegance of the solution is that each organization could maintain its proprietary systems for registration, targeting, and entitlements manage, while still avoiding fragmentation.



UN Women[44] is the first organization to join the Building Blocks network, and a joint pilot was launched in June 2019 to demonstrate precisely how two or more organizations can collaborate to assist the same people on a shared blockchain network. The model is intended to serve as the blueprint for broader collaboration.

UN Women (and each subsequent new member) operates an independent Building Blocks node, and each node validates and records every transaction on the network. Given that it cannot currently be assumed that all beneficiaries have smartphones and connectivity, Building Blocks has developed an innovative solution that allows each humanitarian provider on Building Blocks to be the custodian for the private keys related to their entitlements, while still maintaining a unified view of the people served on the blockchain. Building Blocks does not store any personally identifiable information on-chain.

Once the concept of entitlements unification on the blockchain is well-demonstrated and accepted, it is an easy step

[42]Note that in the Building Blocks system, the balance is printed at the bottom of beneficiary transaction receipts; and this is a feature that is much valued by the beneficiaries. However, because the transaction must be biometrically authorized by the beneficiary, the cashier cannot randomly query beneficiary balances, unless the beneficiary has triggered a transaction.

[43]The savings are achieved by performing all the "accounting" on the blockchain and only using the bank for making payments to merchants. The savings may or may not be replicable in other contexts depending on the operational realities on the ground.

[44]http://www.unwomen.org/en

to move to identity attestations. One organization could, for example, attest that the owner of the public key is a nursing mother. Another organization could then search all the public keys for a "nursing mother" attestation and target services to those beneficiaries that fall within their mandate all without needing to know the sensitive personal information of the underlying people.

As the different pieces of a person's identity puzzle are held by different actors, gaining collaboration based on a shared understanding of the technology and its potential for empowering the people served is fundamental in achieving meaningful blockchain-based identity by bringing all the pieces in one place. Building Blocks is taking the approach that the path to a full-fledged blockchain-based identity system is best started with the less sensitive components of identity. For example, insofar as CBI entitlements are determined and distributed in a siloed manner, the related transaction details are also fragmented across various systems and Financial Service Providers (FSP). In such a scenario, if a credit agency wished to analyze the transaction data to assign a credit score for underwriting a loan, they would likely have access to only a portion of all the data. With fewer data points, statistical risk can be determined to a lower degree of accuracy, resulting in beneficiaries being charged a higher interest rate. Instead, if all entitlements were channeled to the unified blockchain wallet for each beneficiary and transactions were authorized from there, the financial transaction histories would also be unified. Based on this, an organization like Kiva, using a zero-knowledge-proof protocol, for example, could establish a credit rating for a beneficiary using all the data, resulting in a more favorable interest rate on the eventual loan. Furthermore, with Building Blocks, the data is portable, so if a Syrian refugee returns home, she could use the data generated in Jordan to get a small business loan in Syria and become self-sustaining again. Otherwise, the data is likely to stay behind with the FSPs in Jordan and would be inaccessible to the refugee back in Syria (or a new destination).

Like the Kiva protocol, Building Blocks also focuses first on the principles of interoperability and minimization, whereby multiple UN agencies can collaborate securely to have a unified view of the same beneficiary, but no personal identifying information is revealed on-chain, thereby protecting the privacy of the identity subject. Also like Kiva, given the conditions of the user population, self-custody is difficult and therefore not a priority at the start. In both cases, a blockchain-based identity infrastructure enables portability of attestations for migrant populations. Over time, additional use cases can be built on top of the identity system, such as using CBI transaction details across multiple UN agencies as data points to predict credit quality.

A question for the future is whether Kiva protocol may be interoperable with Building Blocks. Thus far, interoperability has been focused on actors within the use case e.g., microfinance institutions in Sierra Leone for Kiva and UN agencies for Building Blocks. The users of each identity system may overlap in the future, as these projects scale. For example, a participant (or former participant) in the Building Blocks program may seek microfinance loans in a jurisdiction that uses the Kiva protocol. In bootstrapping her credit worthiness, would her CBI transactions and attestations from Building Blocks be recognized by the microfinance institutions participating in the Kiva protocol? Recognition requires both policy agreements off-chain and technical standards interoperability on-chain. Conversely, a participant in the Kiva protocol may become a participant of Building Blocks. Could her attestations from the Kiva protocol be used in Building Blocks for various UN agencies to better serve her needs? Could both of these identity systems allow other trusted parties outside the initial set of permissioned nodes to become attestors and nodes? Robust interoperability, technical standards and policy alignments enable these identity systems to have *composability* and *stackability*, whereby new applications could be built on top of the base identity layer.

FUTURE PERSPECTIVES

As people become more and more mobile, a working identity system that can operate on a global scale has become a precondition for ensuring equal opportunities in the global economy. As developing economies are rebuilding their identity systems anew, it is important to be mindful of the consequences that an improperly designed system might cause. The current approaches of centralized governmental-based identity systems relying on biometrics have serious limitations with regard to both security and privacy (Prabhakar et al., 2003). A more decentralized and self-sovereign identity system using verifiable credentials and access controls is not only more flexible and efficient, but can contribute to securing fundamental human rights, especially in countries with unstable governments and fragile institutions (Lemieux, 2017). Given their critical situation, migrants, refugees and other vulnerable populations might benefit from a system that enables them to selectively disclose some attributes but not others, depending on the use cases.

Dependence of Self-Sovereignty on Technology Infrastructure

A true self-sovereign identity system would require a certain level of infrastructure, primarily high penetration of affordable smartphones that can securely store private keys and reliable connectivity. Practitioners in the field, such as Kiva and the WFP, recognize the realities of their constituents, who are vulnerable populations in low infrastructure environments, many of whom live below the poverty line. Therefore, it is not possible to assume wide availability of the technical infrastructure and sophistication for self-management of private keys.

Another problem with localized key storage—beyond hardware affordability—is the larger issue of key recovery, since, in a self-managed environment, losing one's phone necessarily entails losing one's private key. Hence, perhaps the most important obstacle to achieving full self-sovereignty is the problem of key recovery, combined with the price of hardware.

In light of these issues, there is a consensus that the best practice at the moment is a custody or guardianship model, whereby program administrators like Kiva or WFP can manage keys on behalf of constituents, but constituents always have

the ability to opt-out of guardianship should they choose to self-manage.

To address these challenges, some companies are moving into building the first generation of blockchain smartphones. HTC Exodus[45] is one of the first blockchain phones on the market, released in October 2018. The Exodus phone has its own trusted execution environment for secure key management and transaction signing. It deploys a social key recovery mechanism to recover private keys when the phone or passphrases are lost, whereby the user splits the private key among three to five trusted contacts.[46] HTC issued a cheaper blockchain phone in Q3 of 2019 called Exodus 1s, which will be priced in the $250 range.[47] While this would still be prohibitively expensive for many of Kiva's or WFP's constituents, it is a step in the right direction.[48]

Digital Money and the Importance of Self-Sovereign Identity

The use of blockchain ledgers for peer-to-peer money transfer has numerous implications in development economics, further highlighting the need for self-sovereign identity solutions. One interesting application of blockchain technology is the digitization of local or complementary currencies as a natively digital cryptocurrency. Community currencies are usually softly pegged to the national currency, and therefore primarily function as a medium of exchange, rather than a store of value or unit of account.

For instance, Grassroots Economics[49] is a non-profit in Kenya that has been implementing a local currency program called Sarafu Credit with rural farmers since 2010. The Sarafu currency is softly pegged to the Kenyan shilling and is accepted by a local community of farmers, traders and schools. In communities where access to cash (Kenyan shillings) is difficult, bank accounts are inaccessible due to lack of identity documents, and mobile money providers like M-Pesa charge exorbitantly high fees, farmers are increasingly relying on local community currencies, as a complementary solution to the national currency (Dissaux and Ruddick, 2017).

Since October 2018, Grassroots Economics has turned Sarafu Credit into a stablecoin transacted on simple feature phones. A stablecoin is a cryptocurrency that is transacted on a blockchain ledger whose value is pegged to a national currency or a reference basket of assets. With the digitization of Sarafu credit as a stablecoin pegged to the Kenyan shilling, the transactions costs are significantly lower than both the paper version of Sarafu, and M-Pesa transactions. For instance, a 101 Kenyan shilling transaction will have a transaction fee of 11 shillings on M-Pesa, but only 2 shillings with Sarafu (the cost of two SMS, a USSD connection and negligible fees to run crypto transactions on an Ethereum side chain).

Most interestingly, transaction information which would otherwise be owned and controlled by M-Pesa, or remain untraceable with paper money, can now be recorded to a blockchain. This data includes statistics on what kinds of goods and services each wallet is spending its funds on, the transaction sizes, and so forth. Such open source transaction data, when tied to a self-sovereign identity system, would provide rich behavioral information for purposes of underwriting microloans, micro-insurance or other humanitarian applications such as needs assessment planning to determine the amount of cash aid to provide to beneficiaries. Traditionally, needs assessment is done through focus groups and surveys. Dynamic data from live transactions would be far more accurate, timely, and insightful in ensuring that beneficiaries receive an adequate amount of cash aid. Furthermore, as described under the Kiva model, if the loans were disbursed and repaid using cryptocurrency, disbursement and repayment claims could be automatically added to the Kiva's identity protocol, thereby strengthening users' credit profile and enhancing the richness of their digital identities.

Grassroots Economics, Sempo (an Australian startup) and the Red Cross are now working together on a new project called Community Inclusion Currencies (CICs), which is a model for channeling cash aid and other sources of philanthropic or private sector cash as reserves that fractionally issue these local currencies. Through a fractional reserve model, cash donations and aid is effectively levered. For example, $100 worth of cash donation may be issued as $120 worth of CICs. If the CICs are circulated within the community at a high velocity, that further amplifies the initial impact of the $100 of cash aid. In order to maintain price stability of the CICs, redemption of CICs for the underlying cash can be gated algorithmically relative to the existing supply of CICs, the issuance and redemption rates of CICs, and the reserve ratio. The CICs would be issued as a stablecoin pegged to the national currency, and ideally the reserve would also be stored as a fiat-pegged stablecoin, with issuance and redemption automated through smart contracts. The CIC model could enable a scalable alternative mechanism to community banks. For example, women's savings and loan groups could deposit their collective savings into a reserve, and whenever members need loans, the smart contract would issue new CICs. Over time, interest and savings rates could be added in order to make various CIC projects economically sustainable. The CIC project was awarded a two year grant from Innovation Norway, an arm of the Norwegian government, to pilot and scale in Kenya and other locations globally[50].

Stablecoins point to a future where money becomes predominantly global and digital, but bankless (Balvers and McDonald, 2017). Until the advent of cryptocurrency, digital money necessarily meant bank-facilitated transactions, with banks or other financial institutions (the gateways to the banking rails) performing KYC and AML checks. Thus, those without identity documents have been left out of the global digital economy (Borio and Disyatat, 2010). As money becomes increasingly global, there may be a concomitant opportunity for the establishment of an equally global and digital identity

[45] https://www.htcexodus.com

[46] https://www.wired.com/review/review-htc-exodus/.

[47] https://mashable.com/article/htc-exodus-1s-blockchain-phone/.

[48] By comparison, the first cell phone from Motorola retailed for $3,995 in 1982. Today, HTC, Samsung and others sell much more powerful smartphones for <$200. See https://www.timetoast.com/timelines/history-of-cellphones-prices.

[49] https://www.grassrootseconomics.org/

[50] http://news.trust.org/item/20191126123058-xtxvz/

management system that preserves the privacy of users (Vigna and Casey, 2016), while adhering to compliance of global regulatory regimes for KYC and AML. In particular, a synergy might emerge between digital money and digital identity, mediated through a blockchain-based infrastructure, whereby transaction data can function as attestations that increase the richness of a digital identity profile. This could contribute to better credit underwriting, humanitarian needs assessment, and more accurate (and ultimately more inclusive) risk assessments for KYC/AML compliance.

Identity Insurance as Backstop and Revenue Stream for Identity Providers?

Innovative ideas and new markets around digital identity have yet to be realized. One interesting proposal explores creating an insurance marketplace for consequential damages related to identity claims[51], which could be built on top of a digital identity management system similar to Kiva's architecture. Such a marketplace could provide the "last mile" assurance against identity errors (e.g., bad data coming into the identity system) and provide a market mechanism for evaluating the accurateness, trustworthiness and usefulness of various claims associated with an identity (Tang et al., 2003). This would enable lenders to feel more comfortable underwriting a loan—particularly to an individual with no formal credit history, if the claims associated with that individual's profile were insured for consequential losses toward the cost of the loan. Over time, traction in lending activity would result in new attestations from the lender, thereby increasing trust and lowering insurance premiums for that particular individual.

Identity insurance could also become a new revenue stream for identity providers such as banks and microfinance lenders, who are, in any case, required by law to conduct diligent KYC checks. In such a semi-decentralized identity management systems, banks, and lenders could underwrite the risk associated with issuing an identity credential on the blockchain, thereby helping subsequent lenders de-risk and creates economic incentives for the lenders of "first resort"—(i.e., the lenders willing to lend or issue identity credentials earlier in a borrower's digital history).

Refugees with little to no attestations might be subject to higher risk premiums (because they have no track history) until the refugees acquire more quality attestations so as to make them more trustworthy. Such a model could encourage refugees to engage as much as possible with specific institutions or organizations, in order to collect a positive track record of verifiable credentials, and therefore reduce the insurance premium associated with their identity. In some cases, risk premiums may even be subsidized by agencies like UNHCR or other relevant organizations. Although such an insurance model might ultimately be beneficial to refugees and displaced individuals, who do not have a strong government to guarantee for their identity, it should only be experimented after extensive research has been done to mitigate any potential downside or systemic risks of such an identity insurance, such as introducing illegal biases, discrimination or arbitrary value judgment into the underlying identity system.

[51] https://identityinsurance.org/

CONCLUSION

Self-sovereign identity is a relatively new area of research, which is only now starting to materialize into real-world applications of new digital identity management systems. This is particularly valuable for applications that have the ability to scale and greatly improve financial and social inclusion of vulnerable populations (Blakstad and Allen, 2018). Yet, it is important to keep in mind that while there are emerging best practice standards and primitives for self-sovereign identity (McMullen et al., 2019), there is no generic identity protocol that solves all use cases. As demonstrated by the Kiva and WFP case studies, identity is inherently use case dependent. Interoperability and standardization will be important for scale, but the success of a particular identity application will depend on how its deployment is tailored to the use cases and local conditions. A successful identity management system will therefore need to be sufficiently flexible to adapt to the inherently malleable nature of human identity.

The development of cryptocurrencies as a new type of open source mobile money, particularly stablecoins, will enable users to benefit from an increased range of economic opportunities brought about by the new financial services built on top of these systems (Thomason et al., 2018). Verifiable credentials issued by trusted actors can function as identity claims. As described above, credentials signed by WFP to specific beneficiaries can serve as alternative credit scores, while organizations like Kiva can provide identity attestations. Likewise, Grassroots Economics, which currently manages the Sarafu program in Kenya, could sign identity claims on behalf of its participants based on Sarafu transactions, which could help its constituents graduate into Kiva's identity protocol and microfinance ecosystem.

Ultimately, Kiva could provide loan capital in a stablecoin to its microfinance partners, via a peer-to-peer transaction that is cheaper and faster compared to international money transfer via correspondent banking (Darlington, 2014). The microfinance lenders could directly disburse loans in a stablecoin denominated in the local currency of the borrower. The microfinance lenders on Kiva's identity protocol would then automatically sign identity claims in regards to disbursements and loan repayments, as such transactions are now verifiable on-chain, thereby reducing potential disputes. Borrowers could subsequently use these loans for their business needs: purchasing inventory for their shop, paying wages to their employees, and so on. As a result, previous and successfully repaid loans would function as identity attestations, further enriching the digital history and credit profile of the borrowers, and creating a virtuous circle for financial inclusion. These new identity business models, such as identity insurance, would likely arise out of this mobile money/identity ecosystem, further enhancing the robustness of the ecosystem as a whole. And while we are still far from having a truly digital, global and self-sovereign identity system, we believe that blockchain

technology could be one of the key building blocks to instantiate this vision.

AUTHOR'S NOTE

FW is a lawyer and entrepreneur, currently serving as Associate General Counsel at the Maker Foundation, which supports the blockchain project, MakerDAO. She co-founded ixo, a blockchain protocol for tokenizing social impact outcomes as digital assets. She started her career on Wall Street before practicing law in New York and London. She received her law degree from Columbia University and her undergraduate degrees from UC Berkeley. She can be reached at fennie@makerdao.com. PD is a legal scholar, whose work focuses on the legal challenges and opportunities of blockchain technology. She is a permanent researcher at the CNRS, Faculty Associate at the Berkman-Klein Center for Internet & Society at Harvard University, and cofounder of the Coalition for Automated Legal Applications (COALA). She can be reached at pdefilippi@cyber.harvard.edu.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

The authors would like to thank Kevin O'Brien and Aaron Goldsmid of Kiva, Houman Haddad of the World Food Programme and Nick Williams of Sempo for invaluable input and feedback, as well as Georgy Ishmaev for his comments on a preliminary version of the paper.

REFERENCES

Alvarez, R. M., Hall, T. E., and Trechsel, A. H. (2009). Internet voting in comparative perspective: the case of Estonia. *Pol. Sci. Polit*. 42, 497–505. doi: 10.1017/S1049096509090787

Androulaki, E., Karame, G. O., Roeschlin, M., Scherer, T., and Capkun, S. (2013). "Evaluating user privacy in bitcoin," in *International Conference on Financial Cryptography and Data Security* (Berlin; Heidelberg: Springer), 34–51. doi: 10.1007/978-3-642-39884-1_4

Araújo, L. C., Sucupira, L. H., Lizarraga, M. G., Ling, L. L., and Yabu-Uti, J. B. T. (2005). User authentication through typing biometrics features. *IEEE Trans. Signal Process*. 53, 851–855. doi: 10.1109/TSP.2004.839903

Aydar, M., and Ayvaz, S. (2019). Towards a Blockchain based digital identity verification, record attestation and record sharing system. *arXiv preprint arXiv:*1906.09791.

Baars, D. S. (2016). *Towards self-sovereign identity using blockchain technology* (Master's thesis). University of Twente, Enschede, Netherlands.

Balvers, R. J., and McDonald, B. (2017). *Designing a Global Digital Currency*. Available online at SSRN: https://ssrn.com/abstract=3049000

Bhargav-Spantzel, A., Squicciarini, A., Bertino, E., Kong, X., and Zhang, W. (2010). "Biometrics-based identifiers for digital identity management," in *Proceedings of the 9th Symposium on Identity and Trust on the Internet* (Gaithersburg, MD: ACM), 84–96. doi: 10.1145/1750389.1750401

Blakstad, S., and Allen, R. (eds.). (2018). "Leapfrogging banks in emerging markets," in *FinTech Revolution* (Cham: Palgrave Macmillan), 121–132. doi: 10.1007/978-3-319-76014-8_7

Blazewicz, J., Kubiak, W., Morzy, T., and Rusinkiewicz, M. (eds.). (2012). *Handbook on Data Management in Information Systems*. Heidelberg: Springer Science and Business Media.

Borio, C., and Disyatat, P. (2010). Global imbalances and the financial crisis: reassessing the role of international finance. *Asian Econ. Policy Rev*. 5, 198–216. doi: 10.1111/j.1748-3131.2010.01163.x

Burge, T. (1988). Individualism and self-knowledge. *J. Philos*. 85, 649–663. doi: 10.5840/jphil1988851112

Bygrave, L. A. (2012). *The Data Difficulty in Database Protection*. Oslo: University of Oslo Faculty of Law Research Paper (2012-18).

Côté, J. E. (1996). Sociological perspectives on identity formation: the culture–identity link and identity capital. *J. Adolescence* 19, 417–428. doi: 10.1006/jado.1996.0040

Campisi, P. (2013). *Security and Privacy in Biometrics, Vol. 24*. London: Springer. doi: 10.1007/978-1-4471-5230-9

Canham, J. (2018). Biometrics: leap of faith or fact of life?. *Biometr. Technol. Today* 2018, 8–10. doi: 10.1016/S0969-4765(18)30024-9

Cap, C. H., and Maibaum, N. (2001). "Digital identity and its implication for electronic government," in *Towards the E-Society*, eds B. Schmid, K. Stanoevska-Slabeva, and V. Tschammer (Boston, MA: Springer), 803–816. doi: 10.1007/0-306-47009-8_59

Christman, J. (2013). Social practical identities and the strength of obligation. *J. Soc. Philos*. 44, 121–123. doi: 10.1111/josp.12024

Darlington, J. K. III. (2014). *The future of Bitcoin: mapping the global adoption of world's largest cryptocurrency through benefit analysis* (Honors thesis project). University of Tennessee, Knoxville, TN, United States.

Davidson, S., De Filippi, P., and Potts, J. (2016). *Economics of Blockchain*. Available online at SSRN: https://ssrn.com/abstract=2744751

De Filippi, P. (2014). Bitcoin: a regulatory nightmare to a libertarian dream. *Internet Policy Rev*. 3. doi: 10.14763/2014.2.286

De Filippi, P. (2016). The interplay between decentralization and privacy: the case of blockchain technologies. *J. Peer Prod*. 7:al-01382006.

De Filippi, P., and Loveluck, B. (2016). The invisible politics of bitcoin: governance crisis of a decentralized infrastructure. *Internet Policy Rev*. 5. doi: 10.14763/2016.3.427

De Filippi, P., and Mauro, R. (2014). Ethereum: the decentralised platform that might displace today's institutions. *Internet Policy Rev*. 25.

De Filippi, P., and Wright, A. (2018). *Blockchain and The Law: The Rule of Code*. Cambridge, MA: Harvard University Press. doi: 10.2307/j.ctv2867sp

Der, U., Jähnichen, S., and Sürmeli, J. (2017). Self-sovereign identity $-$ opportunities and challenges for the digital revolution. *arXiv preprint arXiv:*1712.01767.

Dissaux, T., and Ruddick, W. (2017). "Challenges of collective organization and institution building around community currencies in Kenyan slums," in *4th International Conference on Social and Complementary Currencies* (Barcelona).

Dunphy, P., and Petitcolas, F. A. (2018). A first look at identity management schemes on the blockchain. *IEEE Secur. Priv*. 16, 20–29. doi: 10.1109/MSP.2018.3111247

Duta, N. (2009). A survey of biometric technology based on hand shape. *Pattern Recogn*. 42, 2797–2806. doi: 10.1016/j.patcog.2009.02.007

Eakin, P. J. (1999). *How Our Lives Become Stories: Making Selves*. Ithaca, NY: Cornell University Press.

El Haddouti, S., and El Kettani, M. D. E. C. (2019). "Analysis of identity management systems using blockchain technology," in *2019 International Conference on Advanced Communication Technologies and Networking (CommNet)* (Rabat: IEEE), 1–7. doi: 10.1109/COMMNET.2019.8742375

Friedman, A., Crowley, P., and West, D. (2011). *Online Identity and Consumer Trust: Assessing Online Risk*. Washington, DC: The Brookings Institution.

Ganapathy, V., Thomas, D., Feder, T., Garcia-Molina, H., and Motwani, R. (2011). "Distributing data for secure database services," in *Proceedings of the 4th International Workshop on Privacy and Anonymity in the*

Information Society (New York, NY: ACM), 8. doi: 10.1145/1971690.1971698

Garcia, P. (2018). Biometrics on the blockchain. *Biometr. Technol. Today* 2018, 5–7. doi: 10.1016/S0969-4765(18)30067-5

Garrett, B. (2002). *Personal Identity and Self-Consciousness*. London, UK: Routledge. doi: 10.4324/9780203015667

Geach, P. (1973). "Ontological relativity and relative identity," in *Logic and Ontology*, ed. M. K. Munitz (New York, NY: New York University Press), 287–302.

Gentry, C., and Boneh, D. (2009). *A Fully Homomorphic Encryption Scheme, Vol. 20, No. 09*. Stanford, CA: Stanford University.

Gerard, D. (2017). *Attack of the 50 Foot Blockchain: Bitcoin, Blockchain, Ethereum & Smart Contracts*. CreateSpace Independent Publishing Platform.

Gleick, P. (2014). *Water, Drought, Climate Change, and Conflict in Syria*. Oakland, CA: Pacific Institute. Available online at: https://journals.ametsoc.org/doi/full/10.1175/WCAS-D-13-00059.1 (accessed July 1, 2014).

Goldreich, O. (1998). *Secure Multi-Party Computation*. Preliminary Version, 78. Available online at: http://www.wisdom.weizmann.ac.il/~oded/pp.html

Hammudoglu, J. S., Sparreboom, J., Rauhamaa, J. I., Faber, J. K., Guerchi, L. C., Samiotis, I. P., et al. (2017). Portable Trust: biometric-based authentication and blockchain storage for self-sovereign identity systems. *arXiv preprint arXiv:*1706.03744.

Hardjono, T., and Pentland, A. (2019). Verifiable anonymous identities and access control in permissioned blockchains. *arXiv preprint arXiv:*1903.04584.

Hileman, G., and Rauchs, M. (2017). Global cryptocurrency benchmarking study. *SSRN Electron. J.* 33. doi: 10.2139/ssrn.2965436

ICRC (2017). *Handbook on Data Protection in Humanitarian Action*. International Committee of the Red Cross.

Jacobovitz, O. (2016). *Blockchain for Identity Management*. Beersheba: The Lynne and William Frankel Center for Computer Science Department of Computer Science, Ben-Gurion University, Beer Sheva Google Scholar, 9.

Jain, A. K., Flynn, P., and Ross, A. A. (eds.). (2007). *Handbook of Biometrics*. New York, NY: Springer Science and Business Media. doi: 10.1007/978-0-387-71041-9

Jain, A. K., and Nandakumar, K. (2012). Biometric authentication: system security and user privacy. *IEEE Comp*. 45, 87–92. doi: 10.1109/MC.2012.364

Jain, A. K., Ross, A., and Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Trans. Circuits Syst. Video Technol.* 14, 4–20. doi: 10.1109/TCSVT.2003.818349

Jøsang, A., and Pope, S. (2005). "User centric identity management," in *AusCERT Asia Pacific Information Technology Security Conference* (Brisbane, QLD), 77.

Khare, R., and Rifkin, A. (1997). Weaving a web of trust. *World Wide Web J.* 2, 77–112.

Kosta, E. (2013). *Consent in European Data Protection Law*. Leiden: Martinus Nijhoff Publishers. doi: 10.1163/9789004232365

Kulkarni, M. H., Yadav, A., Shah, D., Bhandari, P., and Mahapatra, S. (2012). Unique id management. *Int. J. Comp. Technol. Appl*. 3, 520–524.

Lemieux, V. L. (2017). "In blockchain we trust? Blockchain technology for identity management and privacy protection," in *Conference for E-Democracy and Open Government* (Krems), 57.

Maesa, D. D. F., Mori, P., and Ricci, L. (2017). "Blockchain based access control," in *IFIP International Conference on Distributed Applications and Interoperable Systems* (Cham: Springer), 206–220. doi: 10.1007/978-3-319-59665-5_15

McMullen, G., De Filippi, P., and Choi, C. (2019). *Blockchain Identity Services: Technical Benchmark of Existing Blockchain-Based Identity Systems*. Toronto, ON: COALA and BRI Big Idea Whitepaper.

Mordini, E., and Massari, S. (2008). Body, biometrics and identity. *Bioethics* 22, 488–498. doi: 10.1111/j.1467-8519.2008.00700.x

Mühle, A., Grüner, A., Gayvoronskaya, T., and Meinel, C. (2018). A survey on essential components of a self-sovereign identity. *Comput. Sci. Rev.* 30, 80–86. doi: 10.1016/j.cosrev.2018.10.002

Muller, B. J. (2010). *Security, Risk and the Biometric State: Governing Borders and Bodies*. London, UK: Routledge. doi: 10.4324/9780203858042

Nagar, A., Nandakumar, K., and Jain, A. K. (2010). "Biometric template transformation: a security analysis," in *Media Forensics and Security II, Vol. 7541*, eds N. D. Memon, J. Dittmann, A. M. Alattar, E. J. Delp III (San Jose, CA: International Society for Optics and Photonics), 75410O. doi: 10.1117/12.839976

O'Gorman, L. (2003). Comparing passwords, tokens, and biometrics for user authentication. *Proc. IEEE* 91, 2021–2040. doi: 10.1109/JPROC.2003.819611

Othman, A., and Callahan, J. (2018). "The Horcrux protocol: a method for decentralized biometric-based self-sovereign identity," in *2018 International Joint Conference on Neural Networks (IJCNN)* (Budapest: IEEE), 1–7. doi: 10.1109/IJCNN.2018.8489316

Prabhakar, S., Pankanti, S., and Jain, A. K. (2003). Biometric recognition: security and privacy concerns. *IEEE Secur. Priv.* 1, 33–42. doi: 10.1109/MSECP.2003.1193209

Proença, H., and Alexandre, L. A. (2010). Iris recognition: analysis of the error rates regarding the accuracy of the segmentation stage. *Image Vision Comput.* 28, 202–206. doi: 10.1016/j.imavis.2009.03.003

Rane, S., Wang, Y., Draper, S. C., and Ishwar, P. (2013). Secure biometrics: concepts, authentication architectures, and challenges. *IEEE Signal Process. Mag.* 30, 51–64. doi: 10.1109/MSP.2013.2261691

Rannenberg, K., Camenisch, J., and Sabouri, A. (2015). *Attribute-Based Credentials for Trust. Identity in the Information Society*. Berlin: Springer. doi: 10.1007/978-3-319-14439-9

Ross, A., and Jain, A. K. (2004). "Multimodal biometrics: an overview," in *2004 12th European Signal Processing Conference* (Vienna: IEEE), 1221–1224.

Sarkar, S. (2014). The unique identity (UID) project, biometrics and re-imagining governance in India. *Oxf. Dev. Stud.* 42, 516–533. doi: 10.1080/13600818.2014.924493

Schartner, P., and Schaffer, M. (2005). "Unique user-generated digital pseudonyms," in *International Workshop on Mathematical Methods, Models, and Architectures for Computer Network Security* (Berlin; Heidelberg: Springer), 194–205. doi: 10.1007/11560326_15

Schneier, B. (1999). The uses and abuses of biometrics. *Commun. ACM*. 42, 136–136. doi: 10.1145/310930.310988

Shrier, D., Wu, W., and Pentland, A. (2016). Blockchain and infrastructure (identity, data security). *Mass. Inst. Technol. Connect. Sci*. 1, 1–19.

Strohminger, N., Knobe, J., and Newman, G. (2017). The true self: a psychological concept distinct from the self. *Perspect. Psychol. Sci*. 12, 551–560. doi: 10.1177/1745691616689495

Suler, J. R. (2002). Identity management in cyberspace. *J. Appl. Psychoanal. Stud*. 4, 455–459. doi: 10.1023/A:1020392231924

Tang, F. F., Thom, M. G., Wang, L. T., Tan, J. C., Chow, W. Y., and Tang, X. (2003). Using insurance to create trust on the internet. *Commun. ACM* 46, 337–344. doi: 10.1145/953460.953519

Thomason, J., Ahmad, M., Bronder, P., Hoyt, E., Pocock, S., Bouteloupe, J., et al. (2018). "Blockchain—powering and empowering the poor in developing countries," in *Transforming Climate Finance and Green Investment with Blockchains* (Cambridge, MA: Academic Press), 137–152. doi: 10.1016/B978-0-12-814447-3.00010-0

Tikkinen-Piri, C., Rohunen, A., and Markkula, J. (2018). EU general data protection regulation: changes and implications for personal data collecting companies. *Comput. Law Sec. Rev*. 34, 134–153. doi: 10.1016/j.clsr.2017.05.015

Tobin, A., and Reed, D. (2016). *The Inevitable Rise of Self-Sovereign Identity*. The Sovrin Foundation, 29.

Toth, K., and Subramanium, M. (2003). "The persona concept: a consumer-centered identity model," in *3rd International Workshop on Emerging Applications for Wireless and Mobile Access (MobEA)* (Budapest).

Uludag, U., Ross, A., and Jain, A. (2004). Biometric template selection and update: a case study in fingerprints. *Pattern Recogn*. 37, 1533–1542. doi: 10.1016/j.patcog.2003.11.012

Unar, J. A., Seng, W. C., and Abbasi, A. (2014). A review of biometric technology along with trends and prospects. *Pattern Recogn*. 47, 2673–2688. doi: 10.1016/j.patcog.2014.01.016

van der Ploeg, I. (2003). "Biometrics and the body as information," in *Surveillance as Social Sorting: Privacy, Risk and Digital Discrimination*, ed D. Lyon (London: Routledge), 57–73.

van Wingerde, M. (2017). *Blockchain-enabled self-sovereign identity* (Doctoral dissertation, Master's thesis). Tilburg University, School of Economics and Management, Tilburg, Netherlands.

Vigna, P., and Casey, M. J. (2016). *The Age of Cryptocurrency: How Bitcoin and the Blockchain Are Challenging the Global Economic Order*. New York, NY: Macmillan.

Werz, M., and Conley, L. (2012). *Climate Change, Migration, and Conflict in Northwest Africa*. Washington, DC: Center for American Progress.

Whitley, E. A., and Hosein, G. (2010). Global identity policies and technology: do we understand the question?. *Global Policy* 1, 209–215. doi: 10.1111/j.1758-5899.2010.00028.x

Wright, A., and De Filippi, P. (2015). *Decentralized Blockchain Technology and the Rise of Lex Cryptographia.* Available online at SSRN: https://ssrn.com/abstract=2580664

Zibran, M. F. (2012). *Biometric Authentication: The Security Issues.* Saskatoon, SK: University of Saskatchewan.

Conflict of Interest: FW was employed by the Maker Foundation at the time of writing this article.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

ORIGINAL RESEARCH
published: 05 March 2020
doi: 10.3389/fbloc.2020.00005

Blockchain Applications and Institutional Trust

Martin Smits and Joris Hulstijn*

Tilburg School of Economics and Management, Tilburg, Netherlands

In the current discussions around Blockchain and distributed ledger technologies, we find a lack of theory to conceptualize and understand application scenarios. In this paper we propose to conceptualize distributed ledger technologies as trust mechanisms. Whereas, previously one had to rely on a trusted third party (e.g., notary), now one must trust a complex software system—the Blockchain and distributed ledger application—as well as the parties that host the software system and ensure its effectiveness. Based on theories of e-commerce, business networks, and trust, we explore relations between trust and Blockchain design. We analyze three case studies of Blockchain applications in the diamond industry. In each case we study two complementary research questions: (1) how does the blockchain application influence trust, and (2) how do trust based requirements affect the design of a blockchain application? We formulate two propositions and find dynamic interactions between trust requirements, blockchain application design, and transaction trust.

Keywords: blockchain, trust, distributed ledger technology, application scenarios, requirements

OPEN ACCESS

Edited by:
Andrej Zwitter,
University of Groningen, Netherlands

Reviewed by:
Michael Shea,
Independent Researcher, Litchfield, United States
Hyojung Sun,
Ulster University, United Kingdom

****Correspondence:***
Martin Smits
m.t.smits@uvt.nl

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *19 August 2019*
Accepted: *30 January 2020*
Published: *05 March 2020*

Citation:
Smits M and Hulstijn J (2020) Blockchain Applications and Institutional Trust. Front. Blockchain 3:5. doi: 10.3389/fbloc.2020.00005

INTRODUCTION

The popularity of Blockchain and distributed ledger technologies for business applications has increased substantially over the past years. Partly, this is due to a hype, fueled by the rising and dropping value of Bitcoin. But apart from the Bitcoin hype, how can we understand the attractiveness of distributed ledger technologies for its use in business applications? A recent claim is that Blockchain applications may enhance trust in inter-organizational relationships and business transactions. For instance, Meijer and Ubacht (2018) reviewed recent publications, and show that Blockchain is often referred to as a "trust mechanism." Regarding Blockchain as trust mechanism suggests that people now trust technology rather than institutions or agencies (e.g., notary; solicitor) and that such institutions may be combined with or even replaced by Blockchain applications. These effects of distributed technology on business networks appear to be similar to dis-intermediation and cyber-mediation effects in e-commerce (Laudon and Traver, 2018). In the case of dis- and cyber-mediation, traditional intermediaries (e.g., notary; solicitor) are augmented by or even fully replaced by technology-based platforms. However, in some cases this may require new intermediaries, e.g., a software certifier. So, by analogy, Blockchain applications may have a variety of effects on business networks and business relations, including effects on trust and effects on the network structure.

In this paper we focus on how Blockchain applications may enhance trust in business relations, and under which conditions trust is or is not established. To analyze these trust aspects, we take two distinct perspectives. First, we analyze recent Blockchain cases in order to identify how trust requirements have been specified and how such specifications affect the design of the Blockchain

application (**Figure 1**, relation A). Second, we analyze how the design of a Blockchain application influences the levels and types of trust in the business network (**Figure 1**, relation B). In this paper we do not focus on how (existing) trust may affect (new) requirements for trust (relation C).

The aim of this paper is explorative: we define key concepts in chapter 2 (including types of trust, Blockchain Technology, and a conceptualization of Blockchain Applications) and we explore three cases to identify relations between "how do trust requirements influence the design of Blockchain Applications" and also between "how does the design of Blockchain Applications influence trust" (which is a design research question, related to A in **Figure 1**) (which is an effectiveness or behavioral research question, related to B).

The remainder of this paper is structured as follows. Section Theory on Blockchain Technology and Trust defines section Blockchain Technology, and provides conceptualizations of section Blockchain Applications, conceptualizations of trust in the e-commerce domain section Trust, and develops hypotheses for testing relations A and B in **Figure 1** section Relations between Blockchain Applications and Trust. Sections Method and Relations Between Trust and Blockchain in the Diamond Industry detail the method and the case studies. The paper ends with a discussion and suggestions for future research (section Discussion and Conclusions).

THEORY ON BLOCKCHAIN TECHNOLOGY AND TRUST

We first define section Blockchain Technology, then section Blockchain Applications, section Trust, and the framework to analyze section Relations Between Trust and blockchain Applications.

Blockchain Technology

Magazzeni et al. (2017) show that Blockchain in its widest sense combines three existing technologies: (1) distributed databases, (2) encryption and (3) consensus protocols. This combination of technologies makes it possible to build applications around a representation of a shared state. In accounting terms, this shared state is a ledger: a repository of data on transactions and the distribution of assets, recorded in accounts. The consensus protocol ensures that parties maintain an identical copy, without the need for a centralized administrator or data storage. So unlike previous automated communication protocols, Blockchain, and general ledger technology make it possible to maintain a so called "*stateful shared state*" of a series of transactions (Magazzeni et al., 2017). "*Shared*" refers to the fact that all participants maintain an identical copy, unlike current systems, in which parties have to rely on their own version of events. In terms of game theory, parties have common knowledge of the state (Fagin et al., 1995). Potentially, this means a huge step forward, as it removes the need for second-guessing misunderstanding and manipulation. "*Stateful*" refers to the fact that each state of the conversation is stored. The system remembers all steps that went on before, unlike stateless communication protocols that only remember the previous step. As the history is shared, the ledger of states becomes immutable and can only be changed in case of consensus.

For a comprehensive introduction to Blockchain technology, we refer to Swan (2015), Magazzeni et al. (2017), and Smits et al. (2020). In short, a Blockchain consists of "*blocks of data*" where each block codifies a set of transactions. A block of transactions is considered valid if the transactions adhere to formal rules that can be verified automatically. For example, a sales transaction is only valid if the seller actually owns the asset to be sold. To avoid the need for a central authority, a Blockchain operates using a consensus protocol. Parties called "*nodes*" verify the validity of the latest block to be added to the chain. To do so, the nodes have to solve a cryptographic puzzle. The solution is represented by a number, called "*nonce*". Essentially the nodes vote by submitting a nonce, and after a majority of nodes have voted a block to be valid, the block is added to the Blockchain, and proof of validity (the nonce) is included in the next block. To make sure that blocks cannot be manipulated without trace, blocks are hashed. Hashing generates for each block a unique number, also called *hash*. Changing a block will result in a different hash. To allow for comparison, the hash of a block is included in the next block. Nodes try to validate the latest block. To keep track of time, also a *timestamp* is added to the next block. In other words, all pieces of evidence needed to verify that blocks of transactions are valid and unchanged, are included on the Blockchain itself.

Blockchain technology can use different consensus protocols to prove validity. The Bitcoin blockchain uses the *proof of work* (POW) protocol. Nodes need to put quite a lot of computing power into solving the cryptographic puzzle. In return, they are rewarded in the currency that is associated with the Blockchain application. Demonstrating validity has value. However, a Blockchain platform based on proof of work consumes enormous amounts of energy. An alternative system is

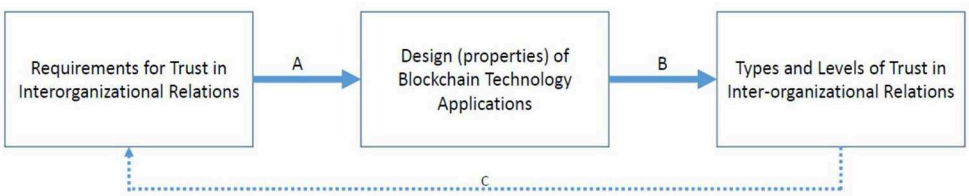


FIGURE 1 | Research model representing two complementary research questions (A,B) (see text).

based on *proof of stake* (POS). In POS, the nodes follow a voting procedure in which nodes that own more of the underlying assets, have a larger voting share. A third alternative is called *validator,* meaning that validity of a block is not determined by voting but by automated verification. A single authority or a selected group of nodes can play the role of validator. Note that such mechanisms re-introduce a form of party trust: the validators need to be trusted.

One can also distinguish *permission less* and *permissioned* Blockchains. The first are open to all actors, the second only to actors with specific permission. For example, if a multi-national firm wants to use a corporate Blockchain for swapping foreign currencies between its country offices, then it makes sense to use a closed (permissioned) Blockchain: only offices of the firm may join. On the other hand, the Bitcoin Blockchain must be open (permission less) to allow all actors worldwide access to the currency. Some authors group these two dimensions into three forms of blockchain: public (permission less, proof of work or proof of stake), consortium (permissioned, selected group of validators), and private (permissioned, single authority), see de Kruijff and Weigand (2017), based on Buterin[1]. See also section Blockchain applications under "Logic Layer."

Blockchain Applications

Blockchain technology can be applied in a business network or in other empirical settings in many different ways. Like all technologies, a Blockchain application must be understood as a sociotechnical system (Clegg, 2000). The sociotechnical system consists of the technological artifact (Blockchain, described in section Blockchain Technology) and the social environment in which the technology is applied, including the interactions between technology and social settings. To analyze the application of Blockchain technology in a specific business network (the socio-technical system), we summarize Smits et al. (2020) who specify three distinct levels at which Blockchain technology may impact a business network (see **Figure 2**).

The business network layers in **Figure 2** are based on e-commerce and business network theory (Van Heck and Vervest, 2007). The bottom layer is the *physical layer*, representing the logistics processes in and between firms (the actors), at specific locations in the network. The *information layer* represents the transactions between firms, and the transaction data stored in information systems (within firms) or in shared ledgers (shared between firms). These shared ledgers may include all data on all transactions, or—depending on the design decisions- only parts of these data. For example, only some crucial financial data or some product properties may be shared in the general ledger. Note that such design decisions may depend on trust requirements specified by actors in the network. The third layer is the *logic layer*. It specifies the business logic, like consensus protocols or validation rules, deployed to control Blockchain operations and automated transactions in other layers. We now specify the three layers in more detail, starting with the information layer (layer 2 in **Figure 2**).

[1]Buterin V. (2015). On Public and Private Blockchains, crypto renaissance salon, August 7, 2015.

The Information Layer

The information layer is where data on transactions are stored in either internal information systems of individual firms or in distributed ledgers shared between firms. Where transactions between organizations used to be stored by each organization internally (represented by separate data silo's in the information layer), transactions can also be stored now only once externally in a Blockchain ledger. Transaction data may include orders, order commitments, as well as payments and deliveries. Internal transactions within the company can be stored in a private (local) blockchain. When transactions are stored (internally or externally) in an irrevocable way in a Blockchain, this not only eliminates duplications (data redundancy), but also related inconsistencies. Another effect of the externalization of data into the shared ledger is mitigation of data heterogeneity. Data representation standards and ontologies will still be needed to enforce a shared definition of crucial concepts, but their reach and effect at the network level will be much stronger, as they are not only used for exchanging data but also for storing the data.

The Physical Layer

The physical layer represents the firms (including intermediaries) and logistics operations involved in the business network. From an organizational perspective, Blockchain-enabled transactions will affect the position of the intermediaries in the physical layer. In particular, *intermediaries supporting information exchange or trust* will be threatened, but this may depend on the type of service offered by the intermediary (e.g., Giaglis et al., 2002). *Search intermediaries* may not be affected. *Trust intermediaries* may be affected if the basis for trust shift to Blockchain security. *Information exchange intermediaries* may be affected because Blockchain aims for single point of storage.

Business transactions are usually related to the movements of goods represented in the physical layer. However, as has been argued in the service science literature, there is an evolution from a goods-dominant logic to a service-dominant logic ("*servicification of goods*"). This not only means that the service sector grows in economic significance, but also a shift from the emphasis on control (ownership) of resources toward use of resources (access right). For example, there is, for instance, less need to own a car if you can have a car or a taxi service, when you need it.

These developments reinforce and are reinforced by Blockchain technology: Blockchain based transactions can be used to transfer money (Bitcoin), but also to transfer access keys for digital products (software and e-books). In the same vein, it can be used to transfer ownership rights on registry goods like houses and ships, and trace consecutive owners along a supply chain.

It is still unclear to what extent transfers of ownership can be turned into valuable services and data. Perhaps Blockchain transactions cannot govern all exchanges at the logistics, physical level. Still, it is expected that Blockchain based transactions will not only record but also govern a large amount of economic exchanges. This may affect operational efficiency (less human effort in the loop) and control efficiency (external control by IT replacing internal control). Together with the savings (and

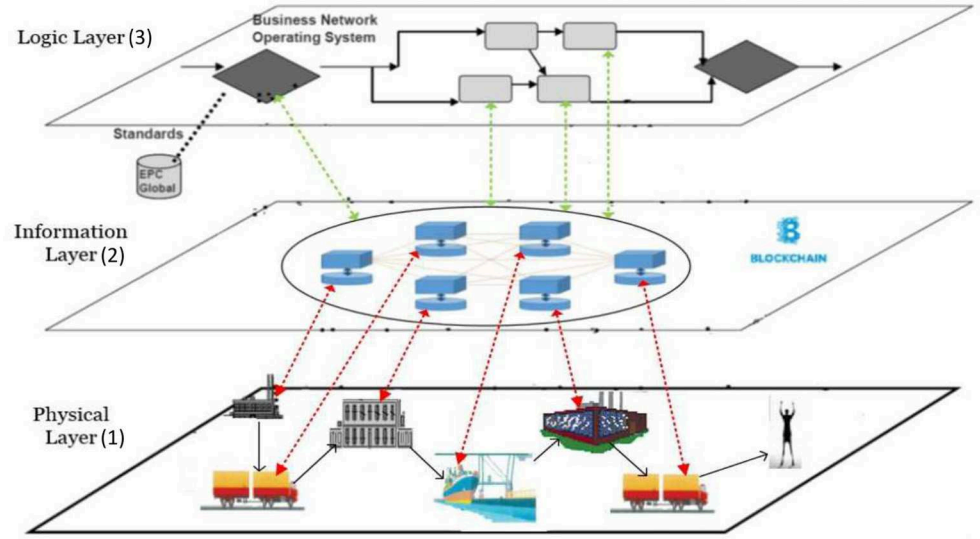


FIGURE 2 | Blockchain applications relate to three business network layers (based on Van Heck and Vervest, 2007; Smits et al., 2020).

costs) at the information layer, this may cause significant savings in transaction costs that in turn may also affect the business network structure.

The Logic Layer

The logic layer is the third business network layer and can become rather complex because it may contain logic and smart contracts that (automatically) do tasks like (i) allowing access to business actors in the network, (ii) executing transactions, (iii) managing risks and rewards, and (iv) assigning roles and responsibilities to business actors (Van Heck and Vervest, 2007). Blockchain transactions can be embedded in smart contracts that are executed automatically. At this moment, smart contracts are still in their infancy, but in principle, there is no computational limit to their scope and smart contracts could take on automated coordination of the other two layers.

We use the 4 × 4 model (Birch et al., 2016) to analyze the logic layer in a blockchain-enabled Smart Business Network. The 4 × 4 model distinguishes four types of logic:

Communication logic: This is the logic for communication between participants in the network. Communication logic includes logic for providing and getting "access to read" and "access to write" for various actors in the Blockchain application (Brennan and Lunn, 2016). Brennan and Lunn (2016) state that in a permission less public Blockchains anyone can read and write on the Blockchain, as long as they meet certain criteria and follow the specified rules. This type of Blockchain is entirely distributed, is a single source of truth and has entirely trustless integrity. A well-known example is the Bitcoin Blockchain. Second, in a permissioned public Blockchain, only permissioned entities may write the ledger, but anyone may view the content. This results in greater accountability and transparency. This form shows great potential in the financial services sector. Third, permissioned private Blockchain, only permissioned entities can read and write on the Blockchain. This form is mostly used in experimental settings where R&D is the main purpose of its existence. A well-known example is the R3CEV consortium (www.R3.com/about).

Content logic: This type of logic is related to the type of goods and services in the business network and the types of assets that are distributed over the network. On a blockchain, many types of assets can be transferred, like cryptocurrencies, letters of credit, or stock bonds. The token value can be simply information, representative of extrinsic value or have intrinsic value. It may also be possible to configure multiple kinds of assets on a single Blockchain.

Consensus logic: To ensure that only legitimate transactions are added to the blockchain, the participating nodes in the network use voting to confirm that new transactions are valid (see above). A new block of data will be added to the Blockchain only if miners in the network reach consensus as to the validity of the transaction. Consensus can be achieved through many different voting mechanisms. The most common is Proof of Work, which depends on probability through the amount of processing power donated to the network (Wright and De Filippi, 2015).

Contract logic: Also defined as the automation logic; the way that transactions are animated to trigger events. Using Blockchain technology, parties have the possibility to confirm that an event or condition has in fact occurred without the need for a third party. A well-known application is a "Smart Contract": a computable contract where the determination of performance and enforcement of contractual conditions occur automatically, without the need for human intervention (Wright and De Filippi, 2015).

Each of the four types of logic can be modified (designed) to optimize the logic layer and to achieve different business objectives (Birch et al., 2016).

Analyzing the Design of a Blockchain Application

To analyze the design of a Blockchain application in a business network setting, we use the three layer model defined above and the nine questions given in **Table 1** (Smits et al., 2020). These questions identify the relevant aspects of the current situation ("As Is") of the Blockchain application in the three layers.

Trust

Trust has been studied in various disciplines. Here we use economic literature (Gambetta, 1988), where trust is related to transactions between buyer and seller. In a (simple) transaction, the buyer needs to trust the seller to deliver the goods or services; and the seller needs to trust the buyer to pay. There are two possible perspectives: trustor (needs reasons to trust the trustee) and trustee (needs to be seen as trustworthy by the trustor). Most literature focuses on the trustor's perspective. Trust is a crucial factor in business relations where there is uncertainty, interdependence, and fear of opportunism, as is the case in online markets (Pavlou and Gefen, 2004). Trust is the foundation of e-commerce (Keen, 1999). Trust between actors has been defined as a "*belief* that the seller will behave in accordance with the consumer's confident *expectations* by showing ability, integrity, and benevolence" (see e.g., Pavlou and Gefen, 2004). Trust is also characterized as "the *willingness* of a party *to be vulnerable* to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" (Mayer et al., 1995, p. 712). So usually, trust refers to a relationship between parties.

TABLE 1 | Questions to assess the three layered design of a Blockchain application.

Physical layer (1)	1. Which firm starts (or started) the Blockchain application, and seeds the first block?
	2. Is the Blockchain application provided by an existing actor in the network or a new entrant (cyber-, dis-intermediation, or re-intermediation)?
	3. Which other firms participate in the Blockchain application?
	4. Is the Blockchain application closed (private blockchain) or open to other firms (public or hybrid blockchain)?
Information layer (2)	5. Which transaction data are stored in the Blockchain (and which data not)?
	6. How is the Blockchain application linked to other (internal and inter-organizational) information systems in the business network?
Logic layer (3)	7. Who (in the network?) decide(s) on the logic applied in the blockchain?
	8. Who may read or write in the blockchain and which control mechanisms are applied?
	9. Which consensus and contract logic is used?

Note that changes in the logic layer may affect the information layer (e.g., which data are shared and stored in the ledger) and the physical layer (e.g., how many organizations will participate; how many transactions will take place).

Parties can trust another, based on reputation or previous contacts (*party or person-based trust*). In modern society, trust relations have often been replaced by formal controls, embedded in institutions (Zucker, 1986). That suggests a category of *institution-based trust*, partly based on reputation and partly based on control mechanisms. We can also trust technology (*technology-based trust*), in the sense that we rely on a mechanism to behave as expected (Vermaas et al., 2010). This depends on a mental model of how the mechanism is supposed to work, and some trust in the party offering the technology, to properly install and maintain it.

Consider the example of a coffee vending machine: we trust the machine to provide coffee when we insert a coin, and not to explode. Is that real trust or merely a metaphor? Upon analysis, it seems that *technology trust* is based on understanding how a system works and on the strength of the prediction of the machine behavior. Note that usually, technology trust also involves party trust and institutional trust. The coffee machine's vendor is trusted to have properly installed and maintained the machine. We may even base our trust on a regulator, to oversee safety of all coffee machines. So even for such a simple case, there is a governance model, involving actors with various roles. By itself, technology cannot be trusted.

In the context of strategic alliances between firms, and in the context of e-commerce, trust has been explored extensively (Das and Teng, 2001; Gefen, 2002; Tan and Thoen, 2002; Perks and Halliday, 2003; Pavlou and Gefen, 2004). Crucial is that e-commerce platforms (like Blockchain applications described above) may enhance trust by adding control mechanisms to the functionality of their platforms, such as an Escrow service, or a reputation rating mechanism. These mechanisms are added to the design of a system to increase trust in other users and reduce possible risks through technological means. Pavlou and Gefen (2004) have shown that in the case of online platforms, trust can be partly based on control mechanisms, such as reputation rating, escrow services, and reviews. Moreover, in the case of e-commerce control mechanisms like reputation rating or reviews, the effectiveness of the (technology based) control mechanism depends on a community of fellow users. The application facilitates and makes use of a social system that provides meaning to it. In a sense, such mechanisms exhibit what has been called socio-materiality: the "social and material aspects of the technology are constitutively entangled" (Orlikowski, 2010). We expect the same to be true for Blockchain applications: effectiveness of a "*stateful shared state*" to generate trust will crucially depend on how the community will accept Blockchain guarantees.

In a series of papers Tan and Thoen explore the notion of *transaction trust*, defined as: "the mental state of the trustor that determines whether he has sufficient trust *to engage in a transaction*" (Tan and Thoen, 2000a,b, 2002). They define transaction trust as the combination of party-based trust and control-based trust. These trust types are defined as follows:

- Party based trust is the belief that the other party (that can be a person or an institution) will behave as expected. This definition fits the definitions above for person based and institution-based trust.
- Control based trust is the belief that the procedures and protocols that monitor and control the successful performance of a transaction, will function properly. Control-based trust also includes the belief that transaction details remain transparent and can be checked. This definition fits the above definition of technology-based trust.

Following decision theory, Tan and Thoen add that the ultimate decision to engage in a transaction for the trustor, depends on a trade-off between "potential gains" of the transaction, and the "transaction risks." The way the trade-off is made, depends on the "transaction trust" as outlined above, but also on the actor's "attitude toward risk" (risk averse, risk seeking). The transaction trust model is depicted in **Figure 3**.

Relations Between Blockchain Applications and Trust

We now use the trust model in **Figure 3** to explore the relations between a Blockchain application (as defined in section Blockchain Applications) and trust (section Trust). Following the definitions of trust, a Blockchain application may affect the decision to engage in a transaction and enter a blockchain based network in four ways:

1. The actor believes the *institution(s)* offering the blockchain based platform to have properly implemented the blockchain, and for each transaction, to faithfully represent the agreement on the blockchain (party-based trust).
2. The actor *believes* the blockchain based network can be *monitored*, and subsequently, that the Blockchain application helps to *reduce* transaction risks (control-based trust).
3. The actor sees *potential gains* because of the Blockchain application in the business network. More potential gains enhance engaging in business network transactions.
4. The actor sees *transaction risks* in the original business network, and believes that a Blockchain application may reduce those risks, through Blockchain based controls.

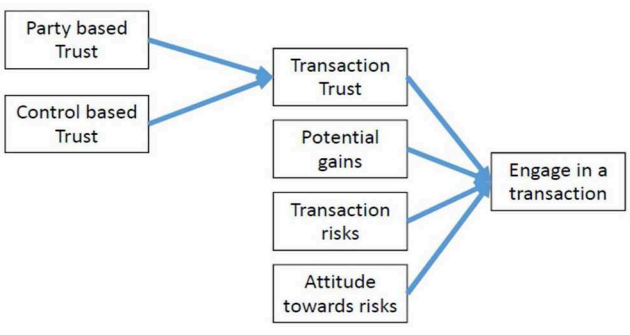


FIGURE 3 | Factors influencing the degree of transaction trust and the decision of an actor to engage in a transaction (adapted from Tan and Thoen, 2000a,b, 2002).

The fifth factor, the actor's *risk attitude* is usually seen as a stable characteristic, and is not likely to be affected by the availability of a Blockchain application.

METHOD

We aim to explore the relations between the design of an artifact, trust in the artifact, and the impact of both design and trust on use of the artifact in a business network. Our research focuses on two related questions, as shown in **Figure 1.** (A) How do the trust requirements in particular application domain influence the design of Blockchain Applications? (B) How does the design of a Blockchain Applications influence the types of trust and trust levels found? Case based research is an appropriate research strategy when it is difficult to separate a phenomenon (blockchain technology effects) from its context (business collaboration, networks, trust, and innovation) (Yin, 2003). In addition, we use observations in case studies to try and develop theory (Eisenhardt, 1989). Specifically, we are interested in design theory for building the artifact (question A), but also in behavioral theory about effectiveness (question B) (Hevner and Chatterjee, 2010).

Cases From Public Sources

As with all emerging technologies, real and mature applications of Blockchain technology are rare. Many organizations have started initiatives to explore the possibilities of Blockchain technologies, but there are few cases in which blockchain is actually deployed in business networks. We have chosen to start by studying publicly known cases, using material from websites, press releases and other public sources, such as news items and technology blogs.

Naturally, this will lead to a bias in the selection of cases. Not many cases of actual implementations of blockchain technologies are known, and even less that have successfully developed in beyond a pilot stage. Moreover, those applications that have been published are likely to be successful ones, or ones that want to be transparent. In addition, there can be bias in the case material itself, because self-published statements are often meant to present a positive image of a project or initiative. Nevertheless, even with this bias toward successful cases and a positive message, the cases provide insight in the aims and choices of a Blockchain application, as we do not use the cases to evaluate success factors, but to search for relations between Blockchain application and trust.

Case Selection

We investigate relations between Blockchain applications and trust in a particular application domain: the diamond industry. We select the diamond industry because trust mechanisms are crucial in this domain. The primary case is Everledger. Everledger offers a Blockchain application focusing on ensuring trust in the *provenance* of diamonds. The term provenance originates from the art and antiques world. It describes means to "relate the value of an object to its origin." The term provenance is also used as a technical term for tracing sources of data in scientific research, and is common in the Semantic Web community where it refers

to the meta-data needed for tracing origin, sources and reliability of data (Simmhan et al., 2004; Janowicz et al., 2015). Observe that the value of objects such as antiques or diamonds depends on the provenance of these objects, the quality and type, the previous owner, and whether the object was lawfully acquired. Such properties can be validated and recorded by Blockchain technology. This characteristic appears to be generic: what is crucial for Everledger, is likely to be crucial for other application scenarios that involve trading objects of value.

When analyzing the Everledger Blockchain in 2018, we found two competing Blockchain initiatives in the same industry: Tracr and Richline. We have included these alternative cases in the analysis, because the three Blockchains provide similar services, but have made different design choices, and induce different types of trust.

We collected the data for all three cases from public sources by doing desk and web research in 2018. We used the official websites[2], as well as additional sources (papers, reports) and blogs. Using the snowball method, we collected 13 documents (65 pages in total) covering the three cases and the diamond industry. We used 18 pages on the diamond industry in general, 17 pages on Tracr, 13 pages on Everledger, and 17 pages on Richline. From these documents, we collected and cross-checked (triangulation) the statements on trust, Blockchain design, and business objectives.

In terms of the research problem (**Figure 1**) we study relationship (A) between requirements for trust in a domain and design choices in Blockchain applications. In the document analysis, these requirements follow from the demands and characteristics of the industry, in this case the diamond industry, and from the type of problem to be solved, in this case trust in provenance of valuable objects. The specific design choices may depend on the context and dependencies of the individual companies and surrounding business networks involved. We also study relationship (B) between the design choices, and the actual types and levels of trust found. This relationship depends on the specific Blockchain application design. As we use public sources, relationships A and B can only be explored to formulate propositions or hypotheses. In depth analysis and hypotheses testing is not possible on the basis of such sources, and needs additional research.

RELATIONS BETWEEN TRUST AND BLOCKCHAIN IN THE DIAMOND INDUSTRY

We present the cases by first describing the industry, including an overview of the key players and business processes. Then we analyze the Blockchain application by using the nine questions (**Table 1**). Subsequently, we explore the relations between the Blockchain application and trust, using **Figure 3** and section Relations Between Blockchain Applications and Trust.

[2]https://www.everledger.io, https://www.tracr.com, https://richlinegroup.com

The Diamond Industry

The diamond industry consists of many small and some large organizations distributed across the supply network illustrated in **Figure 4**. The diamond supply network covers the following five main activities, ranging from mining rough diamonds to selling polished diamonds:

- Mining: Diamond mining takes place in Russia (28% of the total production in 2017), Canada (15%), Botswana (15%), Congo (13%), Australia (11%), and some 20 other countries. Miners sell the rough diamonds to wholesalers. In 2017, 150 million carats of rough diamonds (which equals about 30.000 kilo) were mined for a total value of 15 billion US$.
- Sorting: Wholesalers buy, clean and sort the rough diamonds into "industrial (low) quality" and "gem (high) quality" stones. After that the gem-quality stones are classified in thousands of categories based on size, shape, quality, and color. Wholesalers assign a value to each gem stone. These data are attached to a certificate (under the Kimberly Process Certification Scheme; see below).
- Cutting and polishing: In this phase, the rough diamonds are split and processed into polished diamonds by highly specialized diamond cutting centers (in for instance Amsterdam, Johannesburg, and New York). Polished diamonds are then ready to be sold as gems or to be mounted in jewelry. Diamond cutting centers again classify each diamond, but now on the "four Cs" of the diamond piece: Cut, Color, Clarity, and Carat.
- Diamond Exchanges: Diamonds are sold via registered diamond exchanges. Worldwide, there are about such 25 bourses, all registered by the World Federation of Diamond Bourses (WFDB), which is again supervised by the World Diamond Council (WDC).
- Jewelry Manufacturing and Retail: Jewelers and jewelry manufactures sell the diamonds to end consumers. The total sales value of polished diamonds is about 50 billion US$ per year (www.diamondfacts.org, 2017).

Important actors in the industry are the supervisory authorities KPCS, WFDB, and several large firms. De Beers Group is a large international corporation specialized in diamond exploration, diamond mining, diamond retail, diamond trading as well as in industrial diamond manufacturing (www.debeers.com). Over 70% of the diamond industry is controlled by De Beers via production and purchase agreements with most of the diamond producing countries (Gottlieb, 2006). De Beers provides about one-third of the global supply of diamonds by value (www.tracr.com). Other manufacturers include Alrosa (Bates, 2018) and Diacore, Diarough, KgK group, Rosey Blue, Venus Jewel (Reuters, May 10, 2018).

Two Key Issues in the Diamond Industry

Two key issues in the diamond industry are (i) avoiding trade of so called "conflict diamonds" and (ii) providing trust in provenance ("assuring the origin") of valuable and polished diamonds.

The first issue relates to the trade in rough diamonds. This trade is strictly regulated under the supervision of the

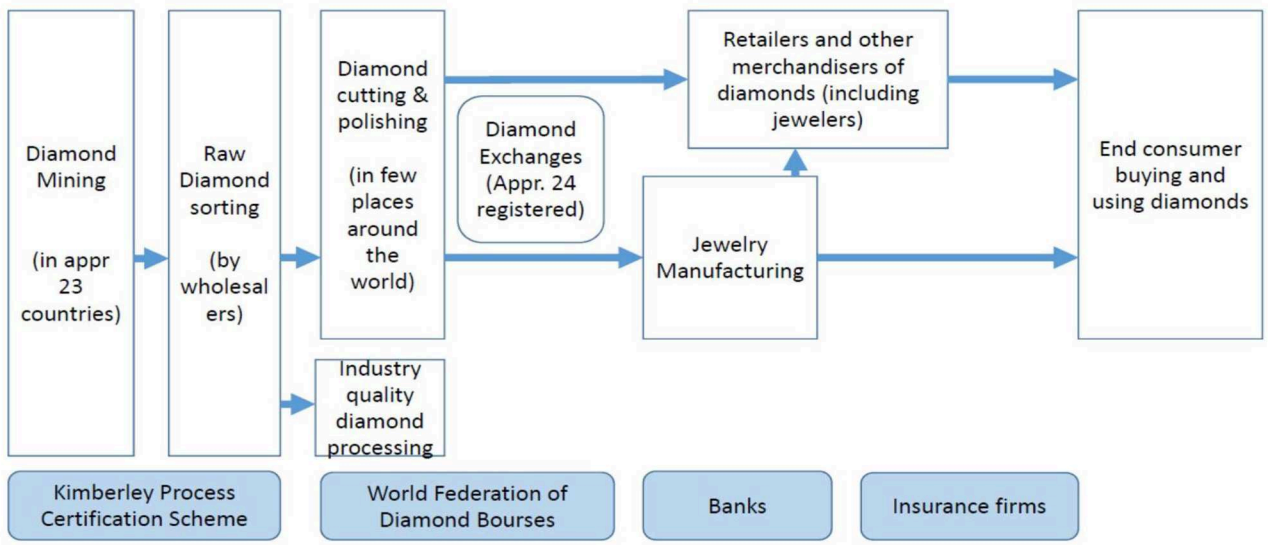


FIGURE 4 | The diamond supply network, including the key actors.

Kimberley Process Certification Scheme (KPCS), which aims to fully eliminate "conflict diamonds." Conflict diamonds are "rough diamonds used by rebel movements … to finance conflict aimed at undermining government." KPCS relies on the financial contributions of participants, supported by industry and civil society observers. The Kimberley Process is, strictly speaking, not an international organization: it has no permanent offices or permanent staff. Neither can the Kimberley Process be considered as an international agreement from a legal perspective, as it is implemented through the national legislations of its participants (www.kimberleyprocess/com; November 2018). KP participants are the states and regional economic integration organizations that are eligible to trade in rough diamonds. As of November 2013, there are 54 KP participants representing 81 countries (including the EU, counting as one participant). KP participants include all major rough diamond producing, exporting, and importing countries. The diamond industry, through the World Diamond Council, and civil society groups are also part of the Kimberley process. These organizations have been involved since the start of KPCS and continue to contribute to its growth and monitoring. As much as 81 governments have enshrined the KPCS into national law. For example, the US adopted the Clean Diamond Trade Act in 2003 (Executive Order 13312). The act requires that all diamonds imported to and exported from the United States have a certificate of origin, according to the Kimberley process, adopted by the UN. In 2018, 99.8% of the world's diamonds are said to come from conflict-free sources. Governments, NGOs and the UN continue to strengthen the Kimberley Process and its system of warranties (www.kimberleyprocess/com; November 2018).

To execute and enforce the KPCS, *rough diamonds receive a unique serial number* that makes it possible to store essential data about a diamond and link the data to the KP certificate. Strong physical control measures exist in the mining and testing process to ensure that data stored for each diamond (type, cut, color, weight, origin, quality) corresponds with the real diamond. In all subsequent processing, the system ensures that the data and the actual diamond remain aligned. For instance, diamonds may be packaged in tamper-proof containers, sealed with an identification code. After that, the transaction history is traced, making it possible to establish legal ownership. In 2017, 70.000 KPCS certificates were issued for a total production of 150 million carat, implying that one KPCS certificate includes -on average- 2.000 carats (about 0.4 kilo) of rough diamonds (if all rough diamonds are certified).

The second issue relates to provenance of polished diamonds. Diamond supply chains are complex and fragmented, resulting in a lack of transparency and trust amongst stakeholders, despite KPCS certificates. The lack of trust, the high value of the assets, and the need to prove that diamonds are legitimately obtained, mean that actors continuously need to *prove provenance of diamonds*. Provenance refers to "the place of origin or earliest known history of something." As stated above, the term originates in the art world, where it means "a record of ownership of a work of art or an antique, used as a guide to authenticity or quality" (online dictionary). The analogy is clear. Buyers of antiques or diamonds usually do not have the expertise to recognize authenticity and quality of the object; they have to rely on evidence from experts. For example, if a retailer wants to sell a valuable diamond, proof is needed on who has cut and polished the diamond and where the original raw diamond came from. Currently, this proof (as far as it exists) is based on linking the KPCS certificate to data provided by the Diamond Exchanges and other actors in the network. However, as the supply network is fragmented, and there are no standards for packaging, identifying or tracing diamonds, it remains hard to establish a full trace

of origin. This explains the huge difference in value between certified and non-certified diamonds.

Three Blockchain Applications to Enhance Trust

In 2018, at least three Blockchain applications aim to solve the issues above: the blockchain applications of Everledger, Tracr, and Richline, a USA based jewelry producer and retailer. We compare the Everledger, Tracr, and Richline blockchain applications and the relation with trust.

Everledger is a rapidly growing business and IT service provider, based in London, and founded in April 2015 (www.everledger.io). In 2018, Everledger had 70 employees across 6 countries. In March 2018, Everledger raised 10 million US$ to expand its global business. Everledger presents itself as an "independent, emerging technology-based enterprise focused on addressing real-world challenges through breakthrough solutions [. . .] to industries where transparency, trust and provenance matter most." This phrase confirms a focus on generation of trust as the main purpose, and a focus on provenance and real-world problems and solutions.

The core issue that Everledger claims to address is "provenance of valuable objects." In 2018, Everledger offers services in six business domains: diamonds, gemstones, minerals, wines, luxury goods, and art (www.everledger.io). Everledger provides services via six different platforms: each domain has its own designed blockchain-based services. To design and provide the services, Everledger collaborates with local experts in mining countries or wine growing areas, or with art experts and artists in the art world. The value proposition of Everledger is based on combining traditional domain knowledge and modern technologies to record, trace and certify transactions and to store evidence on immutable general ledgers. Everledger claims that the Everledger Blockchain application has created an ecosystem of trust within the diamond industry by means of digital provenance tracking and certification.

Everledger makes use of the *IBM Blockchain Platform*, also called *Hyperledger*, for building its blockchain application[3] Everledger is a permissioned system: it is open only for a community of users, who are known in advance and are therefore identifiable and traceable. Nevertheless, the Blockchain application is distributed, avoiding a single point of failure and increasing transparency. The consensus protocol is specifically used to ensure immutability and non-repudiation of transaction records. How validation is done in practice, is not disclosed. Everledger uses expertise of its local partners (as a single authority), in establishing authenticity of a diamond or other valuables. After that, tracking and tracing of the transactions can be done by a regular distributed ledger, i.e., without centralized authority.

Summarizing, the Everledger Blockchain application aims to ensure the following properties:

- *Identification and authentication of diamonds*: diamonds are identified and authenticated, based on a unique number, description of type and origin, and evidence like photographs.
- *Identification and authentication of KPCS certificates*: certificates are uniquely identifiable and traceable, and linked to the diamonds they are about. These properties, if recognized in the market, make it hard to sell two separate diamonds under the same certificate. Note the similarity to the double-spending problem, for which blockchain was originally designed.
- *Data integrity*: no manipulation or deletion of records, after initial recording.
- *Non-repudiation*: once recorded, it is impossible to deny a transaction in which a specific diamond occurs. These two properties makes it possible to trace ownership in a reliable way. If enough traders demand verification of ownership before a transaction, this will make it harder to sell stolen diamonds.

Figure 5 illustrates how the Everledger blockchain application provides services (A–G) to the various actors involved. Note that the Everledger application (like all multi-sided platforms and electronic markets) needs to design and develop customized interfaces and services for each actor type. For instance, service A enables mining experts to add and check information on a certain set of raw diamonds. Service G allows regulators like KPCS to add KP certificates to raw diamonds. Note that Everledger aims to convince each actor (group) to use the platform by offering a specific value proposition for that group, based on customized interfaces.

Tracr is another Blockchain application in the diamond industry. Tracr was conceived by De Beers in 2017 as a mine-to-customer traceability solution. In a pilot project in 2018, Tracr reports that it has identified and tracked 200 diamonds from rough diamond until sales. Tracr claims to have solved the key problem "to determine the characteristics that uniquely identify a rough diamond," "to determine the characteristics that uniquely identify a polished diamond," and "to match the polished with the rough piece" (Bates, 2018).

Tracr is dominated by the mining side of the supply network (actors to the left). Given the large market share of De Beers, it is likely that involvement of De Beers will help to reach a critical mass for diamond certificates. On the other hand, De Beers also represents the "vested interests" in the industry. Blockchain initiatives like Everledger compete with reputation-based trust in provenance. It could be possible that the initiative was started in order not to lose market share.

Richline is a US based company specialized in manufacturing, distribution, marketing and retail of jewelry and luxury goods. The company was founded in 1982 and is based in Florida (USA). Richline has a strong position in lab-grown diamonds. These artificial diamonds are claimed to be chemically, physically, and optically identical to mined diamonds. Diamonds from a lab are guaranteed to be conflict-free and naturally comply with the Kimberley process. In 2018, Richline started a blockchain application, called TrustChain, to ensure provenance of diamonds used in rings and other jewelry (https://www.

[3]https://www.ibm.com/blogs/think/2018/05/everledger/

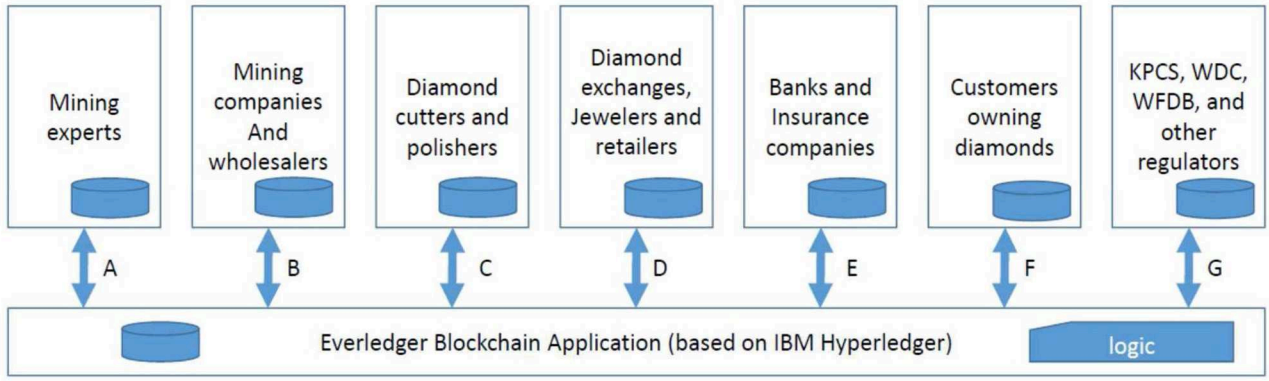


FIGURE 5 | Overview of the Everledger Blockchain application in the diamond industry.

trustchainjewelry.com). The aim is to track and trace diamonds and precious metals from mining (or growing) through to refining, polishing, manufacturing, and delivery. Trustchain is a collaboration between IBM (technology provider), Richline (jewelry manufacturing and distribution), Rio Tinto (diamond supplier for the proof-of-concept), Leach Garner (precious metals supplier), Asahi Refinery (precious metal refinery), and Helzberg (US jewelry retailer). Together these actors cover the entire supply chain. In 2018, they are in the development phase, with the purpose of establishing a proof-of-concept. After that, a full trial will be held with a wider set of industry parties (expected 2019; no further details).

Comparing the Three Blockchain Applications

Before analyzing relations between trust and Blockchain design, we first compare and analyze the designs of the three Blockchain initiatives in the diamond industry using the nine questions on the three application layers in **Table 1**.

The Physical Layer

1. *Which firm started the blockchain application and seeded the first block?* The three cases have different initiators: Everledger is initiated by a London based newcomer in the industry. By contrast, Tracr is initiated by a large, well-known worldwide producer and retailer in the industry (De Beers), and Richline is initiated by a large retailer in the USA, and also a trader in artificial diamonds, who will benefit from increased demand for diamonds with known origin. All three initiatives are in the start-up phase. Developments in party-based trust (does the industry accept the actors?) and control-based trust (does the application provide the right services?) will determine further growth of the initiatives.
2. *Is the blockchain application provided by an existing actor in the network or by a new entrant (cyber-, dis-intermediation, or re-intermediation)?* The Everledger application is an example of a new technology-based intermediary entering the industry. All three initiatives are in principle examples of cyber-mediation: an IT-based intermediary is taking a position in the diamond supply chain. Ultimately, the cybermediary may take over the position of (some) diamond bourses or exchanges, or may lead to bankruptcies of testing agencies (disintermediation). It is also possible that Tracr and Richline involve window-dressing of incumbents in order to retain or regain market share (re-intermediation).
3. *Which other firms participate in the blockchain application?* All three initiatives are in the start-up phase and only a limited number of actors participate in 2018. Note that Everledger provides an infrastructure that allows other actors to enter into the industry, in particular insurance companies ("providing insurance services to diamond owners") and banks ("providing financial services to diamond owners"). This move may be a potential disruptor since banks and insurance providers may require strict certification of diamonds, thereby potentially reducing the power of incumbent firms like De Beers.
4. *Is the blockchain application closed (private blockchain) or open to other firms (public or hybrid blockchain)?* All three blockchain applications are permissioned, but are open to known actors in the diamond industry and also to "all customers that own diamonds" and some to "providers of banking or insurance services." All Blockchain applications require participants to be identified and authenticated. No anonymous users are allowed.

The Information Layer

5. *Which transaction data are stored on the Blockchain, and which data are not?* The initiatives cover data on rough as well as polished diamonds and aim to provide provenance proof by tracking origin, type, quality, and ownership. The material we studied does not provide details on the exact data elements stored in the distributed ledgers. The data architecture of the applications and the uptake of standards for identification and authentication, and representation formats for crucial properties, will affect the further development of services, thereby influencing the potential gains and transaction risks for actors to engage in the blockchain application.

6. *How is the Blockchain application linked to other (internal and inter-organizational) information systems in the business network?* All three applications provide interfaces to information systems maintained by the various actors in the industry (e.g., for recording KP certificates), suggesting functionalities for linking the blockchain application to (some) internal systems. No information is provided on, for instance, the automatic or manual linking process to KP certificates.

The Logic Layer

7. *Who (in the network?) decide(s) on the logic applied in the blockchain?* The initiators of the three applications decide on the logic. The logic is embedded in the application services for different stakeholders (A–G in **Figure 5**), but is based on common properties (physical provenance, identification and authentication, traceability, integrity, non-repudiation). Transparency of the logic, and impact of this transparency on control- based trust remain unclear. Some actors must remain secret (as indicated by De Beers) and therefore the logic must allow for partial disclosures. Successful development of each initiative will depend on how the initiator (focal actor of the network), handles this sensitive issue and how this development will affect party based and control-based trust.
8. *Who may read or write on the Blockchain and which control mechanisms are applied?* The Blockchain access control logic in each initiative determines who exactly will be permitted to enter, and which data may be read or written by which actors. Details on identification and authentication of actors are not provided. Also, no details are provided on how the Blockchain validators verify the certification of the physical diamond mining processes.
9. *Which consensus and which contract logic is used?* It is likely, that all three initiatives work with a validator consensus logic, although the records themselves are distributed. The differences between Tracr and Everledger, illustrate how different perspectives of the focal actors are influencing the decision rights embedded in the blockchain logic. Everledger appears to be a cooperative, whereas Tracr has a clear dominant player. The development of decision rights will further influence the actors' perception of gains and risks, and the subsequent decision to engage in the network.

We now use the observations in the three cases to analyze relations between trust and the design of the Blockchain applications.

Analyzing Relations Between Trust Requirements and Blockchain Applications

Using the trust definitions of **Figure 3**, we analyze (A) the influence of trust requirements on Blockchain application design, and (B) the influence of Blockchain application design on trust. We present our findings in **Table 2** where columns one to three illustrate six observations (A1–A6) on relation A and four observations (B1–B4) on relation B. Our observations in the three cases provide support for the impact of four trust requirements (T1–T4) on six Blockchain design aspects (BC1–BC6). We identify 15 examples (1–15 in the right column) of the impact of (four) Blockchain design choices on trust, gains and risks.

Our observations in the three cases and **Table 1** lead to the following propositions. Specifically, observations A1–A6 appear to support P1, and observations B1–B4 appear to support P2. This provides reason for these propositions to be further developed and tested in additional research.

P1: Trust requirements influence the design choices for the physical, information and logic layers of the Blockchain application.

P2: Blockchain design properties influence party based trust, control based trust, expected gains, and expected risks of using the Blockchain application.

Our observations in the three cases were only made in 2018, which is the year that the three initiatives started offering their services in the diamond industry. More follow-up research is needed to evaluate the combined impact of design choices and trust on Blockchain application and business network success.

DISCUSSION AND CONCLUSIONS

The aim of this paper was to explore the relations between trust and the design of Blockchain applications. We first defined a *Blockchain application* as an application of Blockchain technology in a sociotechnical setting, also known as a business network. To analyze the design of a Blockchain application, we use the three layer model (**Figure 2**) consisting of the (i) physical layer specifying the firms and logistics in the business network, (ii) the information layer specifying the data architecture of transactions and shared ledgers, and (iii) the logic layer specifying four types of logic (communication, content, consensus, and contract logic). Second, we define *trust* using an (adapted) model of Tan and Thoen, which analyzes transaction trust in terms of party-based trust, control-based trust, potential gains, transaction risks, and risk attitude.

We analyzed three Blockchain applications in the diamond industry. The diamond industry is characterized by assets, whose value depends on ensured provenance. This need for provenance is strengthened by regulatory compliance (Kimberly Certification Process). Hence, trust mechanisms are crucial in this domain. The three Blockchain applications differ in their design choices on each of the three layers. In the physical layer, we observe different numbers and types of actors who participate in the Blockchain applications; in the information layer, we observe different types of data shared; in the logic layer, we find different types of business logic.

One key question is about how trust requirements in a business setting affect the design of a Blockchain application. In the three cases we find six examples of the impact of trust requirements on design. The other key question is about the effect of Blockchain application design on types and levels of trust. In the three cases we find 15 examples of the impact of design on party trust, control trust, expected gains, and risk. We formulate two propositions to be developed in future research.

We conclude from our observations that trust requirements do indeed influence the design of a Blockchain application and also, vice versa, that the design of a Blockchain application

TABLE 2 | Six observations (A1–A6) on how trust requirements affect Blockchain application design and four observations (B1–B4) on how application design affects trust.

Trust requirements	A	Blockchain application design	B	Party trust/control trust/gains/risks
T1. Needs to track provenance of valuable goods	A1	BC1. Link IDs to rough and polished diamonds (identification and authentication in the logic layer)	B1	1. Belief in mechanisms for identification and authentication of rough and polished diamonds (control trust) 2. Belief in mining and cutting experts from Everledger partners, De Beers or Richline participants to execute these mechanisms (party trust) 3. Expected gains: reduced costs of testing downstream 4. Expected risks: increased dependency on limited number of certifiers for trading
	A2	BC2. Identification of pieces (in physical, information, and logic layer)		
	A3	BC3. Single source of truth on origin, quality, and ownership of objects (shared data in the information layer)	B2	5. Belief in mechanisms for tracking and tracing objects (control trust) 6. Belief in blockchain immutable records of ownership (control trust) 7. Belief in Blockchain platform providers to execute these mechanisms properly (party trust) 8. Expected gains: increased certainty of origin, quality and ownership, reduced costs of insurance 9. Expected risks: increased dependency on limited number of certifiers for trading
T2. Needs for valid data entry; and compliance with audit criteria (KPCS)	A4	BC4. Permissioned blockchain with validation protocol in the logic layer	B3	10. Belief in mechanisms for data entry and KPCS compliance (control-based trust) 11. Belief in Blockchain platform providers to execute these mechanisms properly (party-based trust) 12. Expected gains: reduced transaction risks, reduced compliance risks 13. Expected risks: increased bureaucracy and administrative burden
T3. User needs to control their data	A5	BC5. Contract logic and communication logic		
T4. User needs to check with other users	A6	BC6 Ability to tell and share stories using the Blockchain application (information and logic layer)	B4	14. Platform based belief in Blockchain data and provenance of the diamonds (party trust) 15. Expected risks: reduced risks because of shared risks and protection by the community

influences the trust induced. These vice versa relations suggest dynamic interactions between application design choices and trust over time (**Figure 1**). For example, if a new-comer offers a Blockchain application in a network, the design may enhance trust for those organizations that decide to start using the application. After some time, those trust levels may have become "de facto" mandatory for all actors in the network. This may trigger other actors, such as incumbents, to formulate different or stronger (trust) requirements that will force the original new-comer to adjust the information, physical, or logic-layers of the design. If the subsequent design is taken on, and effective, this will again lead to changes in trust and perceptions of trustworthiness.

The possibility of such a trust dynamic shows that the current discourse of Blockchain replacing trust by means of technology, is too simplistic. At best it will replace some forms of trust by other forms of trust. In particular, party trust in traditional institutions is replaced in technology-based control trust combined with some residual party trust, namely in those parties who execute the control mechanisms.

These dynamic relations between trust, Blockchain application design, and further business developments make it hard to predict which Blockchain application design will be most commonly adopted. Prediction is even more difficult when multiple Blockchain applications are competing for dominance. We advise to follow the developments of the three Blockchain applications in the diamond industry to evaluate interactions between trust, design, and adoption.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: www.jckonline.com/editorial-article/tracr-de-beers-blockchain-platform/, footnote 2 in other web sources see article.

AUTHOR CONTRIBUTIONS

MS and JH made substantial contributions to the conception and design of the work, revising the work critically, approved for publication of the content and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved.

REFERENCES

Bates, R. (2018): *Inside TRacr, the De Beers developed Blockchain platform. In Luxury*. Available online at: www.jckonline.com/editorial-article/tracr-de-beers-blockchain-platform/ (accessed November 2, 2019).

Birch, D., Brown, R., and Parulava, S. (2016). Towards ambient accountability in financial services: shared ledgers, translucent transactions and the technological legacy of the great financial crisis. *J. Payments Strategy Syst.* 10, 118–131.

Brennan, C., and Lunn, W. (2016). *Blockchain: The Trust Disruptor*. Report Credit Suisse.

Clegg, C. W. (2000). Sociotechnical principles for system design. *Appl. Ergono.* 31, 463–477. doi: 10.1016/S0003-6870(00)00009-0

Das, T. K., and Teng, B.-S. (2001). Trust, control and risk in strategic alliances: an integrated framework *Organ. Stud.* 22, 251–283. doi: 10.1177/0170840601222004

de Kruijff, J., and Weigand, H. (2017). "Towards a blockchain ontology," in *Advanced Information Systems Engineering (CAiSE 2017),* eds E. Dubois and K. Pohl (Berlin: Springer), 29–43.

Eisenhardt, K. M. (1989). Building theories from case study research. *Acad. Manage. Rev.* 14, 532–550. doi: 10.5465/amr.1989.4308385

Fagin, R., Halpern, J. Y., Moses, Y., and Vardi, M. (1995). *Reasoning About Knowledge*. Cambridge, MA: MIT Press.

Gambetta, D. G. (1988). "Can we trust?" in *Trust,* ed D. G. Gambetta (New York, NY: Basil Blackwell), 213–237.

Gefen, D. (2002). Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM Sigmis Database* 33, 38–53. doi: 10.1145/569905.569910

Giaglis, G., Klein, S., and O'Keefe, R. M. (2002). The role of intermediaries in electronic marketplaces. *Inform. Syst. J.* 12, 231–246. doi: 10.1046/j.1365-2575.2002.00123.x

Gottlieb, M. S. (2006). *Jewelry Retail: An Industry Study, MSG Accountants, Consultants and Business Valuators.*

Hevner, A., Chatterjee, S. (2010): *Design Research in Information Systems: Theory and Practice*. Berlin: Springer. doi: 10.1007/978-1-4419-5653-8

Janowicz, K., van Harmelen, F., Hendler, J. A., and Hitzler, P. (2015). Why the data train needs semantic rails. *AI Mag.* 36, 5–14. doi: 10.1609/aimag.v36i1.2560

Keen, P. G. W. (ed.). (1999). *Electronic Commerce Relationships: Trust by Design.* Englewood Cliffs, NJ: Prentice-Hall.

Laudon, K. C., and Traver, C. G. (2018). *E-commerce, 13 Edn*. London: Pearson education.

Magazzeni, D., McBurney, P., and Nash, W. (2017). Validation and verification of smart contracts: a research agenda. *Computer* 50, 50–57. doi: 10.1109/MC.2017.3571045

Mayer, R. C., Davis, J. H., and Schoorman, F. D. (1995). An integrative model of organizational trust. *Acad. Manage. Rev.* 20, 709–734. doi: 10.5465/amr.1995.9508080335

Meijer, D., and Ubacht, J. (2018). "The governance of blockchain systems from an institutional perspective, a matter of trust or control?" in *Proceedings of the 19th Annual International Conference on Digital Government Research (DG.O 2018),* eds M. Janssen, S. A. Chun, and V. Weerakkody (Delft: ACM), 90:91–90:99.

Orlikowski, W. J. (2010). The sociomateriality of organisational life: considering technology in management research. *Camb. J. Econo.* 34, 125–141. doi: 10.1093/cje/bep058

Pavlou, P. A., and Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Inform. Syst. Res.* 15, 37–59. doi: 10.1287/isre.1040.0015

Perks, H., and Halliday, S. V. (2003). Sources, signs and signalling for fast trust creation in organisational relationships. *Eur. Manage. J.* 21, 338–350. doi: 10.1016/S0263-2373(03)00049-5

Simmhan, Y. L., Plale, B., and Gannon, D. (2004). A survey of data provenance in e-science. *ACM Sigmod* 34, 31–36. doi: 10.1145/1084805.1084812

Smits, M. T., Weigand, H., and Kruijff, J. D. (2020). "How blockchain technology affects performance of financial services," in *Digital Transformation*, eds B. van Gils and E. Proper (Berlin: Springer Verlag).

Swan, M. (2015). *Blockchain: Blue Print for a New Economy*. Boston, MA: O' Reilly

Tan, Y.-H., and Thoen, W. (2000a). An outline of a trust model for electronic commerce. *Appl. Artifi. Intell.* 14, 849–862. doi: 10.1080/08839510050127588

Tan, Y.-H., and Thoen, W. (2000b). Towards a generic model of trust for electronic commerce. *Int. J. Electro. Commer.* 5, 61–74. doi: 10.1080/10864415.2000.11044201

Tan, Y.-H., and Thoen, W. (2002). Formal aspects of a generic model of trust for electronic commerce. *Decis. Sup. Syst.* 33, 233–246. doi: 10.1016/S0167-9236(02)00014-3

Van Heck, E., and Vervest, P. H. M. (2007). Smart business networks: how the network wins. *Commun. ACM* 50, 28–37. doi: 10.1145/1247001.1247002

Vermaas, P. E., Tan, Y.-H., van den Hoven, J., Burgemeestre, B., and Hulstijn, J. (2010). Designing for trust: a case of value-sensitive design. *Knowl. Technol. Policy* 23, 491–505. doi: 10.1007/s12130-010-9130-8

Wright, A., and De Filippi, P. (2015). Decentralized blockchain technology and the rise of Lex Cryptographia. *SSRN*. doi: 10.2139/ssrn.2580664

Yin, R. K. (2003). *Case Study Research: Design and Methods*. Newbury, CA: Sage Publications

Zucker, L. G. (1986). "Production of trust: institutional sources of economic structure, 1840-1920," in *Research in Organizational Behavior,* eds B. M. Staw and L. Cummings (Greenwich: JAI Press), 53–111.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

HYPOTHESIS AND THEORY
published: 12 March 2020
doi: 10.3389/fbloc.2020.00010

Digital Identity and Distributed Ledger Technology: Paving the Way to a Neo-Feudal Brave New World?

Oskar J. Gstrein[1]* *and Dimitry Kochenov*[2]

[1] Campus Fryslân Data Research Centre, University of Groningen, Leeuwarden, Netherlands, [2] Faculty of Law, University of Groningen, Groningen, Netherlands

While the digital layer of social interaction continues to evolve, the recently proclaimed hopes in the development of digital identity could be both naïve and dangerous. Rather than just asking ourselves how we could digitize existing features of identity management, and corresponding financial transactions on a community or state level, we submit that truly useful and innovative digital identities need to be accompanied by some significant rethinking of the essential basics behind the organization of the world. Once digital technologies leave the realm of purely online or deeply local projects, the confrontation with the world of citizenship's biases and the random distribution of rights and duties precisely on the presumption of the lack of any choice and absolute pre-emption of any disagreement comes into a direct conflict with all the benefits Distributed Ledger Technology purports to enable. Some proponents of Distributed Ledger Technology-based identity systems envisage "cloud communities" with truly "self-sovereign" individuals picking and choosing which communities they belong to. We rather see a clear risk that when implemented at the global scale, such decentralized systems could be deeply harmful, reinforcing and amplifying the most repugnant aspects of contemporary citizenship. In this contribution, we present a categorization of existing digital identity systems from a governance perspective and discuss it on the basis of three corresponding case studies that allow us to infer opportunities and limitations of Distributed Ledger Technology-based identity. Subsequently, we put our findings in the context of existing preconditions of citizenship law and conclude with a suggestion of a combination of several tests that we propose to avoid the plunge into a neo-feudal "brave new world." We would like to draw attention to the perspective that applying digital identity without rethinking the totalitarian assumptions behind the citizenship status will result in perfecting the current inequitable system, which is a move away from striving toward justice and a more dignified future of humanity. We see the danger that those might be provided with plenty of opportunities who already do not lack such under current governance structures, while less privileged individuals will witness their already weak position becoming increasingly worse.

Keywords: digital identity, self-sovereign identity, citizenship, human dignity, discrimination

OPEN ACCESS

Edited by:
Jane Thomason,
University of Queensland, Australia

Reviewed by:
Michael Shea,
Independent Researcher, Litchfield, CT, United States
Jon Crowcroft,
University of Cambridge, United Kingdom
Anwaar Ali,
University of Cambridge, United Kingdom, in collaboration with reviewer JC

***Correspondence:**
Oskar J. Gstrein
o.j.gstrein@rug.nl

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *16 August 2019*
Accepted: *11 February 2020*
Published: *12 March 2020*

Citation:
Gstrein OJ and Kochenov D (2020) Digital Identity and Distributed Ledger Technology: Paving the Way to a Neo-Feudal Brave New World? Front. Blockchain 3:10. doi: 10.3389/fbloc.2020.00010

INTRODUCTION

The World Bank set up an Identification for Development program (ID4D) in 2014 (World Bank, 2018, p. 1). In the 2018 report of this program, it is claimed "that an estimated 1 billion people globally face challenges in proving who they are because they lack official proof of their identity. As a result, those people struggle to access basic services—including healthcare, education, financial, and mobile services—and may miss out on important economic opportunities, such as participating in the digital economy or formal employment" (World Bank, 2018, p. 3). Accordingly, the World Economic Forum (WEF) established a "Platform for Good Digital Identity" at the beginning of 2018 (WEF, 2018a). While this initiative remarkably focuses on "good" identities with the objective to "ensuring that everyone can participate in the digital society through identity and access mechanisms" (WEF, 2018b, p. 8), the question of technological feasibility remains largely open. It seems promising to explore, however, the role Blockchain and other Distributed Ledger Technologies (DLT; including Ethereum, IOTA, Hyperledger and others) could play in underpinning such systems of fully or at least largely decentralized identification (Verhulst and Young, 2018, pp. 30–31; Wagner et al., 2018)[1]. In a report from December 2018, the WEF presented research estimating that, by 2022, 150 million people will have "blockchain-based" digital identities (WEF, 2018b, p. 17). Additionally, the market for identity verification is projected to be between 16 and 22 billion dollars (Pike and Dickson, 2018). Much of this discussion and the associated hopes focus on enhancing the capabilities of inefficient identification systems in the Global South, as well as aiding structurally suppressed groups within developing countries. As is typically the promise when it comes to digitization in public administration (Fleer, 2018, pp. 1350–1354), DLT-based systems should be able to make public services more efficient. For developed countries, this can mean that it is quicker, more convenient for the citizen, and more cost-effective to provide them. For developing countries, however, the promise is that it is possible to provide "proper" public administration in many areas for the first time. In the context of how the use of innovative technology can aid in bridging the gap between the global north and south, the term "leapfrogging" is used (Parry, 2011), and it is not difficult to imagine such disruptive strides could be made in the area of digital identity once DLT systems are applied in large scale.

However, "digital identity" and "self-sovereign identity" are also "buzz" terms. The hopes vested in them could be both naïve and dangerous, unless accompanied by some significant rethinking of the crucial basics behind the organization of the world. In this submission, we will particularly focus on this tension in connection to the allocation of citizenship and "innate" individual rights. We argue that when implemented at the global scale, DLT-based digital identity systems could be deeply harmful, reinforcing and amplifying some of the most repugnant aspects of contemporary citizenship. While particularly people from developed countries take their privileged status for granted, citizenship remains one of the most crucial global instruments for upholding and reinforcing inequalities through installing (often impenetrable) barriers in a world where inequalities are rooted more in space than in class (Milanovic, 2012). Arguably, manifested in traditional identity management systems such as passports, glass ceilings are distributed among the human population, in many ways emerging as the core element of the contemporary world order. Therefore, one might propose that such behavior is opposed to the enlightenment ideal of equal human worth, the idea of deserving and rationality (Carens, 2015; Kochenov, 2019), as well as the concept of human dignity, which is at the core of modern human rights law (Petersen, 2012). In other words, the current citizenship system can be considered as a rigid cast system. This claim is supported by empirical evidence collected and analyzed by Kochenov and Lindeboom (2019) and Harpaz (2019). If technology is uncritically taking the side of the current status quo, instead of offering new rationales to question it, it will most probably emerge as yet another, immensely effective tool of oppression and injustice. Given the current trends and ongoing discussions, we perceive the likelihood of realization of such a grim perspective as high. Since we are currently living in a world where the majority of features associated with citizenship amount to liabilities—rather than bundles of rights as Kochenov (2019) outlines in detail throughout his monograph—it can be assumed that more and better identification is not necessarily a desirable way forward. The improved policing of the random distribution of privilege with the help of new technologies could result in less justice in the world.

Some proponents of DLT-based identity systems envisage "cloud communities" with truly "self-sovereign" individuals picking and choosing which communities they belong to (Orgad, 2018, pp. 251–260). While there is no universally acknowledged definition of self-sovereign identity, Allen (2016) has described it as "[…] the next step beyond user-centric identity and that means it begins at the same place: the user must be central to the administration of identity." Allen goes on to propose 10 principles that should be associated with self-sovereign identities. Wagner et al. (2018, p. 27) have proposed to define it as "a model of digital identity where individuals and entities alike are in full control over central aspects of their digital identity, including their underlying encryption keys; creation, registration, and use of their decentralized identifiers […] The architecture gives individuals and entities the power to directly control and manage their digital identity without the need to rely on external authorities." In that sense, it is even imaginable that DLT-based systems would allow individuals to freely choose the communities they associate themselves with for different purposes and for a limited time (e.g., You prefer the education system of country A, but health care in country B suits your needs better? Why not have both if you meet the basic requirements?).

In this contribution, we propose a governance-focused categorization of existing approaches to digital identity systems, use three case studies of existing digital identity systems to infer opportunities and limitations of DLT-based identity

[1]For a detailed technical definition of standards for a decentralized identifier that is not necessarily based on DLT, see https://www.w3.org/TR/did-core/ (accessed November 23, 2019).

specifically, put these in the context of the existing preconditions of citizenship law, share our broader concerns on recent developments, and conclude with a suggestion of relevant tests for DLT-based identity systems that we put forward to avoid the plunge into a neo-feudal "brave new world"[2]. While only one of our digital identity case studies uses DLT, we describe the other already implemented large-scale digital identity systems to highlight salient aspects that are also relevant to the development and use of DLT-based systems. We see tensions embedding DLT in existing and undeniable power structures, using digital identities cross-border in a societally meaningful way, and backing digital identities up using biometrical data as anchor.

DIGITAL IDENTITY AND DISTRIBUTED LEDGER TECHNOLOGY AS APPLIED AT THE STATE OR LOCAL LEVEL

As we investigate innovative digital identity programs from a governance perspective and with a focus on assigning rights and duties in the public sphere, we can distinguish three categories:

1. Centralized Top-Down; e.g., Aadhaar, India[3]
2. Individual Incentive Based; e.g., E-Residency, Estonia[4]
3. Community Based Bottom-Up; e.g., Forus.io/"*Kindpakket*", Netherlands[5]

Not all of the examples mentioned use DLT or Blockchain as underpinning technologies. However, since they were built with digital technologies at the core, even those not using a Blockchain-like system share common characteristics, opportunities, and risks relevant for DLT-based and decentralized identity management systems. It is therefore useful to consider them in this submission, especially since some of them have already been implemented in very large scale. In this section, we will describe the context and main features of these three categories before outlining the main opportunities and risks we see.

Example No. 1: Aadhaar

The "Aadhaar" program in India is arguably one of the most prominent examples of a "Centralized Top-Down" approach to digital identity management. Since India, while not so highly developed, is a country with one of the largest populations worldwide, it understandably presents a considerable challenge to implement a smoothly working identification mechanism. It was estimated that, by 2008, the four most frequently used traditional identity programs in India were passports that were used by 40 million, Permanent Account Numbers (PAN) for use by the Indian Income Tax Department with 70 million registrations, the "Ration Card" (issued by states governments to allow for the purchase of essential commodities such as wheat) with 220 million registrations, and finally 500 million voter IDs issued by the Electoral Commission (Zelazny, 2012, p. 6). Given these digits and the knowledge of the current population of India, it is clear that identification management in India in 2008 was not working comprehensively covering the entire population. It was and continues to be difficult for the country to register citizens at birth (Masiero, 2018, p. 7). In such a situation, digitization is attractive to build a safer, quicker, more efficient, and transparent system.

The Indian government started to draw up a plan for a new digital identity in 2006 and founded the Unique Identification Authority (UIDAI) in 2008 (Zelazny, 2012)[6]. Subsequently a Unique Identity (UID) was developed, which consists of 12 numbers. In order to link this identifier to a person, large amounts of personal data about it and its family are being collected (e.g., date of birth, parents' names, etc.). In particular, the number of biometric measurements is extensive and includes fingerprints as well as iris scans (Masiero, 2018, p. 4). This is also relevant for DLT-based digital identity systems, since the use of biometrics is considered as one possible solution to link "anonymous" digital wallets containing digital identities to their rightful owners (De Filippi and Wright, 2018, pp. 14–16). Although heavily disputed by some developers (Burt, 2019), biometrics continue to be an option for backup mechanisms in cases where users lose access to their digital identities, or if devices storing them have been destroyed or lost. This fits into a larger trend of increasingly using biometrical information to identify users on smartphones and mobile devices (Rattani et al., 2019, pp. 12–18).

At the end of 2018, it was estimated that 90.1% of the Indian population or more than 1.2 billion individuals were registered with the system[7]. Their UIDs are stored and managed in a centralized database system managed by the UIDAI. Although the Aadhaar system was lacking a specified list of purposes at the time of its inception, it was primarily intended to facilitate the delivery of social welfare (particularly nutrition) and to address concerns about ineffective distribution of the subsidies or fraudulent behavior (Masiero, 2018, pp. 6–7). Before turning to opportunities and risks of Aadhaar and similar Top-Down digital identity systems, we will continue to introduce case studies for the second and third category as indicated above.

Example No. 2: Estonian E-Residence

When it comes to digital identity based on individual incentives, the Estonian E-Residency program has gained a lot of attention[8]. Estonia has become one of the most innovative countries in the area of digital governance over the last decades. To promote the country as an economic hub within the European Union being open to business from everywhere, the government launched an E-Residency program on December 1, 2014[9]. This "new digital nation" supposedly consists of individuals from across the world

[2]With reference to Aldous Huxley's dystopian novel from 1932, where the allocation of roles in society was strictly predefined and controlled through advanced technological systems.

[3]https://uidai.gov.in (accessed August 8, 2019).

[4]https://e-resident.gov.ee (accessed August 8, 2019).

[5]https://forus.io (accessed August 8, 2019).

[6]The UIDAI has a website at: https://uidai.gov.in (accessed August 7, 2019).

[7]https://uidai.gov.in/images/state-wise-aadhaar-saturation.pdf (accessed August 8, 2019).

[8]https://e-resident.gov.ee/ (accessed August 8, 2019).

[9]Identity Documents Act and State Fee Act Amendment Act (Isikut tõendavate dokumentide seaduse ja riigilõivuseaduse muutmise seadus), RT I, 29 October 2014, 1.

who decide to establish their business in Estonia (Poleshchuk, 2016). Individuals interested in registering for E-Residence in Estonia can administrate their business remotely and use Estonia with its state-of-the-art digital technologies as their hub. They might consume Estonian services and products along the way, do business via the Estonian companies they found, and might eventually wish to move to Estonia, raising the country's profile. The program was launched with the target to have 10 million E-residents by the year 2025. At a later stage, the goal was added to attract 20,000 companies by the year 2021. On December 1, 2018, Estonia's population of E-residents is composed of a∼50,000 people from 157 countries, while a large portion of this growth reportedly occurred in 2018. Additionally, around 6,000 new companies have established themselves through the program (Korjus, 2018a). E-Residence essentially first intended to attempt to offset the deficiency of the original jurisdictions the E-Residence reside in and of which they hold the citizenship. It would be premature to report any real success, however: banking due diligence rules and frequent actual residence requirement meant that—for example—an Iraqi with an E-Residence is still first and foremost tied to her country and the possibility of doing business in or via Estonia is *de facto* very limited. To potentially mitigate some of these issues, the Estonian administration is looking into developing version 2.0 of this program (Korjus, 2018b).

Example No. 3: Kindpakket

The third category of digital identity programs we propose to consider consists of "Community Based Bottom-Up" approaches. Such programs might first seem limited in scope and impact. Indeed, they focus on significantly smaller populations. Communities such as the city of Zug in Switzerland (Kohlhaas, 2017) or the community of Zuidhorn/Westerkwartier in the province of Groningen in the Netherlands have successfully experimented with digital identity based on DLT in community settings (Velthuijs, 2018). In the case of the latter, a community child welfare program was realized (*Kindpakket*). The identity of a potential applicant gets stored in a digital wallet that is controlled through a smartphone application (called "Me"). Once the identity and the essential credentials are confirmed by the community/state or a trusted third party (e.g., certifying notary), the individual can "shop" for benefits that are tendered in different funds administered by the community or offered by other benevolent actors (e.g., humanitarian organizations). If the individual is interested in a specific program or fund (e.g., childcare benefits) that can be found on a platform, it is possible to apply directly. One of the additional benefits of the design of the system is that no raw personal data are being exchanged in the application process. The technology is able to assess the application just based on whether the criteria of the fund meet the credentials of the person. This is made possible by implementing a method called "Zero Knowledge Proof" (ZKP). According to Kulkarni (2018, p. 60), "ZKP allows a user to construct a mathematical proof so that, when a program is executed on some hidden input known only to that user, it has a particularly publicly known output, but without revealing any other information beyond this." Kulkarni goes on to explain that ZKP has been further developed with the emergence of "Zero-knowledge Succinct Non-Interactive Argument of Knowledge" (Zk-SNARKS), which allows one to prove something is true without revealing the reason for why it is true. With the implementation of such technologies in digital identity systems, the individual gains more control over the management of her own digital identity: she chooses under which circumstances she shares personal data, and the system is designed in a way that limits the exchange of raw personal data considerably. Therefore, it is often being claimed that such systems operationalize the concept of self-sovereign identity, solving the data protection-related "identity crisis" of the digital age (Toth and Anderson-Priddy, 2019, pp. 17–18). The individual and citizens become less dependent on intermediaries such as governments or other institutions. Once the individual meets the requirements, it receives either currency, or purpose-bound vouchers ("tokens") that can be used at merchants or for specific services (e.g., entrance to public swimming pool, sport lessons, etc.) that the backer of the fund wants to promote[10]. Since pilots in this area seem promising, it is not unlikely that such programs will become more common in many communities across the world in the years to come.

Opportunities and Limitations

It has been claimed that Blockchain is a solution searching for a problem (Frederik, 2018), and in the years 2018 and 2019, the "disruptive" potential of Blockchain and other DLT is questioned considerably. In particular, the financial sector seems to be disappointed after having made significant investments in the development of "proof of concepts." In that light, analysts claim that DLT solutions either work mainly as niche applications (e.g., supply chain sector), have modernization value replacing long-outdated systems, or have reputational value guaranteeing prestige (Higginson et al., 2019). To investigate the usefulness of DLT-based applications, Zwitter and Boisse-Despiaux (2018, p. 6) have proposed four guiding questions to find out whether DLT solutions are appropriate for the envisaged use case. Paraphrased, these are as follows:

1. Do the benefits of a DLT solution justify the costs of development and the scaling process?
2. Is the application demanding decentralization through distribution and built-in trust through transparency?
3. Does a ledger created by the application need to be immutable?
4. Does the final application comply with legal norms, relevant codes of conducts, ethical principles, and human rights?

If these questions cannot be answered affirmatively, the use of DLT might be unnecessary and other technologies might be better to achieve sustainable progress. However, despite these critical aspects, it is certainly too early to state that DLT as such have failed. There are still plenty of promising pilot projects[11],

[10]The project has been discussed with the developers in September and October 2018.

[11]https://blockchan.ge/curatedexamples.html (accessed August 8, 2019).

and the technology keeps developing beyond the "original" Blockchain system underpinning the cryptocurrency Bitcoin, which, at the time of writing, was already more than 10 years old (Nakamoto, 2008). In other words, DLT are "not monolithic concepts," with changing attributes that offer great potential in the area of "disintermediation, transparency, and accessibility." Therefore, it is still being widely believed that DLT can have a significant impact in the area of identity management, even if this might take longer than many have proposed in the past years (Verhulst and Young, 2018, p. 16).

If we consider the presented case studies and aim at identifying opportunities, we see the following: First, digitization of identity clearly offers a venue to be more precise, effective, and comprehensive in identity management. The sheer number of issued UIDs under the Aadhaar regime is impressive, and India might actually have found a tool to comprehensively issue identities for its entire population for the first time in its history. Secondly, it seems likely that the cost of administration of identities can be reduced due to the advantages of automatization. Thirdly, particularly the use of DLT and the implementation of the self-sovereign identity concept have the potential to put the individual in control of its own credentials. This could result in a profound culture change in an area in which individuals typically depend on the state, public institutions, or corporations to administer their identity. This can be enabling for individuals, offering them more choice and possibilities as the Estonian model and the study in Zuidhorn show. Additionally, many have associated hopes that this will increase the level of data protection and privacy, reducing the likelihood of large data breaches containing millions of personal credentials (Toth and Anderson-Priddy, 2019, pp. 17–18). Fourthly, DLT seems to facilitate cross-border cooperation. It is imaginable that not only communities or the state provide funds in the case of Community Based Bottom-Up approaches, but also humanitarian organizations or private parties do so to provide aid in areas that were struck by natural or man-made disaster. All of these opportunities are significant and explain the interest in the subject.

Nevertheless, for these opportunities to be realized, the following limitations need to be overcome: First, the development of DLT needs to be based on fundamental values and human rights. This can be tied to "human centered design" approaches (Giacomin, 2014). Concretely, DLT-based digital identities need to support governance structures respecting, protecting, and promoting privacy and more generally the autonomy of the individual. Poorly designed digital identity systems can seriously threaten the enjoyment of privacy, as the Aadhaar example demonstrates. In particular, biometrical data are very sensitive and difficult to protect with legal frameworks (Jasserand, 2018, p. 155), and the identifiers used relate to integral parts of the body of each individual. While it is possible to start or stop using a key, password, or any other credential used to create trust, the management of data relating to the physical shape of a human being requires much more refined frameworks. In this respect, it is also necessary to consider how attractive a centralized database containing credentials of more than 1 billion people is for private parties, as well as all types of cybercrime, cyberattacks, cyberespionage, or cyberwarfare. Not only from this perspective, the underlying regulatory and governance framework of Aadhaar seemed inappropriate since specified purposes, as well as safeguards and individual remedies for the use of the system, were not identified in a specific law at the inception of the system. In a judgment from 26 September 2018, the Indian Supreme Court tried to respond to these challenges by limiting the purposes the UID has to be used for (Indian Supreme Court, 2018). Following this judgment, it is no longer mandatory to use a UID when opening a bank account, buy mobile phone cards, in an admission process to a school, or for appearance in boards or common entrance examinations (Mahapatra, 2018; Privacy International, 2019).

This highlights a second limitation that needs to be addressed. As powerful as digital identities might be, it is important to make sure existing governance structures are precise, adequate, and have the capability to link them to "the real world." The ambitious Estonian E-Residence program ran into this limitation at the point at which individuals started to apply for bank accounts in the country. Tax authorities and other actors along the value creation chain find it currently difficult to work with digital identity, which in turn makes these identities practically useless. Estonia aims at addressing this issue in version 2.0 of the E-Residence program (Korjus, 2018b), but the underlying issue here might be the different requirements in different areas of the regulatory space (e.g., Anti-money-laundering frameworks), which all have to be proportionate and aligned in order for digital identity to work (Kaiser, 2018, pp. 578–587). Community-Based Bottom-Up approaches seem to be less sensitive to this problem since the scope of their operations is smaller and the technology is applied "closer" to the individual, which allows one to tailor it more carefully, taking the concrete problems into account.

A third limitation we see is the impact on the development of groups and social equality in general. While the Indian government has claimed that Aadhaar particularly helps the poor, careful observers such as Usha Ramanathan claim that inaccurate use of biometric data, the spreading mandatory nature of the UID, and other shortfalls create challenges for weak and sick people living in rural areas and can result in life-threatening situations for members of the transgender community or others whose identity may now be clear, but still not accepted widely in society (Bhardwaj, 2018). The datafication of social interaction reshapes the relationship not only between the individual and the government but also between groups and the rest of society (Taylor et al., 2017, pp. 226–235). Possibly, this aspect raises one of the most important aspects when discussing digital identities and the use of DLT in this area. If digital identities will fully replace existing concepts such as citizenship, they will not be able to do so in an environment that is free of customs, traditions, and power structures. This has also significant implications for deciding how centralized or decentralized the architecture of a DLT-based digital identity system can be. A fully decentralized system might be potentially empowering for the individual on the one hand, but the necessity to keep the link to society (and the resources it controls) remains on the other.

DIGITAL IDENTITY AT THE GLOBAL LEVEL: TOTALITARIAN NEO-FEUDALISM

What would happen if DLT-based digital identity was applied globally and became the standard tool of choice, eventually digitizing citizenship? To develop an answer to this question, it is useful to clarify that citizenship does not depend on the need of documenting identity, which many of the predominantly technology-focused proponents arguably aspire to solve. Virtual nations or "cloud communities," as long as they aim at replicating existing national structures, will probably make the world worse off, particularly those individuals who are less privileged already. The framing of this complex topic that academics such as Orgad (2018, pp. 251–260) propose seems unhelpful, especially in the context of his aspirational concept of a "global" citizenship, and leaving beside what such a global citizenship would ultimately mean in detail. If digital identity management driven predominantly by concerns relating to technological feasibility was to replace traditional identity management and citizenship, this arguably random segregation of the global population into relatively closed groups of varying value will continue (Kochenov, 2019). Some of these DLT-based digital identities—just as currently citizenship—will come with far-reaching rights, whereas others will predominantly represent liabilities. Hence, digitized identity management will first and foremost make this segregation process more granular, and effective.

To illustrate this with a concrete example, if someone is assigned a humiliating set of liabilities in real life—e.g., a Central African Republic citizenship—instead of a noble and democratic status—e.g., citizenship of France—virtual nations will not change anything from the perspective of individual rights and human dignity. The lack of any rights worldwide coming with some citizenships as opposed to a bundle of rights coming with others can be measured. By comparing the gross domestic product (GDP), Human Development Index (HDI), travel freedom, settlement, and work rights abroad, it is easy to see why being French—a status welcoming you to the job market of 41 countries—is infinitely better than being a citizen of the Central African Republic (Kochenov and Lindeboom, 2017, 2019). Hence, the actual problem derives from already existing real-world inequalities between identities and citizenships. It is not only that citizenships by definition exclude, the difference between citizenships matters (Kochenov, 2018, pp. 321–324). The question is how the digitization of identity management will implicitly or explicitly affect and interact with this reality.

To elaborate on this point, citizenship's core function throughout history has been to establish and police global race- and wealth-based hierarchies. This unfolds in many different perspectives, such as gender: it took US women almost a 100 years to get the right to vote, and women in the Swiss canton of Appenzell-Innerrhoden had to wait until 1991 and a decision of the Federal Supreme Court was necessary (Swiss Federal Supreme Court, 1990). Compared with women in "developed countries," individuals living in colonial territories fared even worse. While African Americans have not been enjoying the same rights as "Caucasian" US citizens historically, the same is true for those with different ethnic backgrounds living in European and Asian empires. Emmanuelle Saada has researched how arbitrary—based entirely on skin color—the ascription of French citizenship in the colonies of the Republic was (Saada, 2012). After decolonization was finished following the Second World War, the former colonial subjects are now confined to places around the world reserved uniquely for the losers of Ayelet Shachar's infamous "birthright lottery" (Shachar, 2009). Hence, the world has both changed and remained the same. It changed, because in the second half of the 20th century, the Western world has started to accept women's rights. Furthermore, racial and indigenous minorities within "first world" states are also respected in many cases. Nevertheless, in other aspects, the world has remained the same. Milanovic (2012) has outlined that inequalities can now been found between states, rather than within national borders. Hannah Arendt's concept of a "right to have rights" citizenship for individuals who would otherwise be stateless is a status associated with rights in a handful of countries only (Oman, 2010, pp. 280–289). In many others, it is a severe and undeserved liability with sometimes fatal consequences. Those locked into the poorest former colonies do not inhabit the same narrative as privileged individuals of the global north. Citizenships are thus about preserving inequality worldwide. If cloud communities, digital identity projects, and virtual nations do not address this issue in their design, these fundamental realities will remain the same.

In other words, before considering potential benefits of a set of quasi-citizenships and the deployment of digital identity to create virtual nations, it is crucial to be fully aware of the drastic differences between citizenships in "real life." This has to be considered in the light that many see digital identities as an opportunity to fully "identify" populations in countries that fail to register individuals at birth, or comprehensively throughout their lives. To be identifiable is not necessarily "a good thing." If the development and deployment of DLT-based digital identities do not recognize and take into account the circumstances, the promised benefits will remain a dream, and digitized identity management might ultimately see a considerable societal push-back.

CONCLUSION: A HOLISTIC ASSESSMENT OF DIGITAL IDENTITY

Coming back to the start of this submission, and the question how a "good" digital identity can be achieved, we hope to have convincingly made the point that such an achievement can only be realized if the current culture and understanding of identity and citizenship can be significantly improved. In other words, what is needed for true and meaningful progress is technology-enabled change of the status quo (Grinbaum and Groves, 2013, pp. 139–140), rather than the mere and value-neutral digitization of existing paradigms and power structures. Potentially, such an attempt to provide an ethically sound version of digital identity can also be inspired by the discussion around the ethical valuable use of artificial intelligence (Gath, 2018). Unless this cannot be guaranteed by the proponents and implementing actors of DLT-based identities, it might be overall better for society to stick with the current systems despite their "gray" areas and incomplete features. In the end, this might result in

more freedom and opportunities for the individuals concerned and enable more societal development than artificially restricting frameworks based on immature technological systems.

Nevertheless, we believe that digital identities based on DLT have potential if designed in the right way. For example, the GENESIS Design principles for Blockchange seem capable as guidelines (Verhulst and Young, 2018, pp. 74–77). The acronym is composed of (G)overnance legitimacy based on (E)thically sound intentions. The aim should be to produce solutions for real problems, (N)ot to promote technology as such. In this submission, we have shown that the Community-Based Bottom-Up approach seems particularly promising in this respect. This may also be one of the main reasons why the studies in this area deliver immediately tangible and solid results. Still, the (E)cological footprint of DLT-based systems remains an open question (De Vries, 2018, p. 804). However, as mentioned earlier in this submission, we think that it is important to keep in mind that DLT are still developing and, as other technologies in the past, might also get more energy effective over time. In particular, as we follow the discussion about the transition of "Proof of Work" to "Proof of Stake" consensus mechanisms, this seems not unlikely (Xu, 2018). The next principle is aimed at making sure that DLT use is (S)ynchronized with existing initiatives. We have alluded to this aspect in this submission at various stages, but believe that particularly the four-step test proposed by Zwitter and Boisse-Despiaux (2018) adds a useful perspective to this consideration. Additionally, when designing identity systems, (I)nteroperability and open standards are crucial to avoid vendor lock-in or dependence on large players. It is hard to imagine a truly self-sovereign identity based on proprietary technological standards. Finally, the last principle is to (S)ecure first block accuracy, which can also be interpreted as making sure that once personal data (especially biometrical data) are put on an immutable ledger, these data are accurate and do not cause unnecessary harm for the respective individual or citizen.

To conclude, we suggest that implementing DLT-based systems for identity management needs a holistic approach taking all of the aforementioned aspects into account and putting them at the center of the design process of applications. As we aimed at demonstrating throughout, it seems particularly useful to take into account existing knowledge, inequalities, and the limitations of citizenship law. Since the Centralized Top-Down approach and Individual Incentive programs particularly tend to make overgeneralized assumptions on individual (and collective) identity, we see less space for their success in the short- to mid-term. Still, we believe that it is useful to consider the development of such systems since they provide the background for a discussion on what identity should and could mean in the Digital Age. Nevertheless, the immediate future seems to belong to Bottom-Up digital identity approaches with their ability to improve gradually and incrementally, taking the complex social environment into account on a much more granular and practical level.

DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article. For further information and data please contact the corresponding author.

AUTHOR CONTRIBUTIONS

Initially, OG took the lead on section Digital Identity and Distributed Ledger Technology as Applied at the State or Local Level, while DK took the lead on section Digital Identity at the Global Level: Totalitarian Neo-Feudalism. At a later stage, the authors worked on all of the sections of this piece together, reviewing and reformulating each other's arguments throughout.

ACKNOWLEDGMENTS

The authors are grateful to Prof. Liav Orgad and Prof. Rainer Bauböck for their critical engagement with some of the ideas, which DK contributed to the EUI online debate on Cloud Communities: The Dawn of Global Citizenship, which are developed further in this paper. Additionally, we would like to thank Maarten Velthuijs and Jamal Velij for taking the time to discuss and explain the underpinning technological paradigms and features. Finally, Carolin Kaiser deserves a special mention for the numerous inspiring exchanges on the subject.

REFERENCES

Allen, C. (2016, April 26). The Path to Self-Sovereign Identity. *Life With Alacrity*. Available online at: http://www.lifewithalacrity.com/previous/ (accessed November 22, 2019).

Bhardwaj, A. (2018, September 26). Here's what Prashant Bhushan and Usha Ramanathan have to say on #AadhaarVerdict. *Newslaundry*. Available online at: https://www.newslaundry.com/2018/09/26/heres-what-prashant-bhushan-and-usha-ramanathan-have-to-say-on-aadhaarverdict (accessed August 8, 2019).

Burt, C. (2019, October 4). Self-sovereign identity community discusses the future of digital ID at IIW XXIX. *Biometricupdate*. Available online at: https://www.biometricupdate.com/201910/self-sovereign-identity-community-discusses-the-future-of-digital-id-at-iiw-xxix (accessed November 22, 2019).

Carens, J. (2015). *The Ethics of Immigration*. New York, NY: Oxford University Press.

De Filippi, P., and Wright, A. (2018). *Blockchain and the Law*. Cambridge, MA: Harvard University Press.

De Vries, A. (2018). Bitcoin's growing energy problem. *Joule* 2, 801–809. doi: 10.1016/j.joule.2018.04.016

Fleer, P. (2018). Digitization and the continuities of change in administrative information processing. *Adm. Soc.* 50, 1335–1359. doi: 10.1177/0095399718791540

Frederik, J. (2018, August 25). De blockchain: een oplossing voor bijna niets. *de Correspondent*. Available online at: https://decorrespondent.nl/8628/de-blockchain-een-oplossing-voor-bijna-niets/519071687772-2a5ee060 (accessed October 30, 2018).

Gath, C. (2018). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Phil. Trans. R. Soc. A* 376:20180080. doi: 10.1098/rsta.2018.0080

Giacomin, J. (2014). What is human centred design? *Design J.* 17, 606–623. doi: 10.2752/175630614X14056185480186

Grinbaum, A., and Groves, C. (2013). "What is "responsible" about responsible innovation? Understanding the ethical issues," in *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society,* eds R. Owen, J. Bessant, and M. Heintz (New York, NY: Wiley & Sons), 119–142.

Harpaz, Y. (2019). *Global Citizenship 2.0*. Princeton NJ: Princeton University Press.

Higginson, M., Nadeau, M. C., and Rajgobal, K. (2019, January). Blockchain's Occam problem. *McKinsey & Company*. Available online at: https://www.mckinsey.com/industries/financial-services/our-insights/blockchains-occam-problem (accessed August 8, 2019).

Indian Supreme Court (2018). *Justice K.S. Puttaswamy (Retd) vs Union of India (2017), Writ Petition (Civil)*. W.P. (C) No.-000494-000494/2012.

Jasserand, C. (2018). Law enforcement access to personal data originally collected by private parties: Missing data subjects' safeguards in directive 2016/680? *Comput. Secur. Rev.* 34, 154–165. doi: 10.1016/j.clsr.2017.08.002

Kaiser, C. (2018). *Privacy and Identity Issues in Financial Transactions* (dissertation Thesis). University of Groningen, Gronigen, Netherlands.

Kochenov, D. (2018). "Escapist technology in the service of neo-feudalism," in *Debating Transformations of National Citizenship. IMISCOE Research Series,* ed R. Bauböck (Cham: Springer), 321–326.

Kochenov, D. (2019). *Citizenship*. Cambridge, MA: MIT Press.

Kochenov, D., and Lindeboom, J. (2017). Empirical assessment of the quality of nationalities: The quality of nationality index (QNI). *Eur. J. Compar. Law Governance* 4, 314–336. doi: 10.1163/22134514-00404007

Kochenov, D., and Lindeboom, J. (2019). *Kälin and Kochenov's Quality of Nationality Index*. Oxford: Hart Publishing.

Kohlhaas, P. (2017, December 7). Zug ID: Exploring the First Publicly Verified Blockchain Identity. *Medium*. Available online at: https://medium.com/uport/zug-id-exploring-the-first-publicly-verified-blockchain-identity-38bd0ee3702 (accessed August 8, 2019).

Korjus, K. (2018a, November 30). E-Residency is 4 Years Old so Here's 4 Surprising Facts About the Programme. *Medium*. Available online at: https://medium.com/e-residency-blog/e-residency-is-4-years-old-so-heres-4-surprising-facts-about-the-programme-c3a9d64c988d (accessed August 8, 2019).

Korjus, K. (2018b, September 11). E-Residency 2.0: What do Estonians think of the programme? *Medium*. Available online at: https://medium.com/e-residency-blog/e-residency-2-0-what-do-estonians-think-of-the-programme-99853274a55b (accessed July 22, 2019).

Kulkarni, K. (2018). *Learn Bitcoin and Blockchain : Understanding Blockchain and Bitcoin Architecture to Build Decentralized Applications*. Birmingham, UK: Packt Publishing Ltd.

Mahapatra, D. (2018, September 27). *Times of India*. Available online at: https://timesofindia.indiatimes.com/india/aadhaar-stays-minus-fangs-and-pangs/articleshow/65972588.cms (accessed August 8, 2019).

Masiero, S. (2018). Explaining trust in large biometric infrastructures: a critical realist case study of India's Aadhaar project. *E J Info Sys Dev Countries* 84:e12053. doi: 10.1002/isd2.12053

Milanovic, B. (2012). *Global Income Inequality by the Numbers: In History and Now*. Policy Research Working Paper No. 6259 of The World Bank. Available online at: http://documents.worldbank.org/curated/en/959251468176687085/Global-income-inequality-by-the-numbers-in-history-and-now-an-overview (accessed July 26, 2019).

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. Bitcoin.org. Available online at: https://bitcoin.org/en/bitcoin-paper (accessed August 8, 2019).

Oman, N. (2010). Hannah Arendt's "Right to Have Rights": a philosophical context for human security. *J. Hum. Rights* 9, 279–302. doi: 10.1080/14754835.2010.501262

Orgad, L. (2018). "Cloud communities: the dawn of global citizenship?" In: *Debating Transformations of National Citizenship. IMISCOE Research Series*, ed R. Bauböck (Cham: Springer), 251–260.

Parry, J. (2011). Leapfrogging into the future. *BMJ* 342:d2990. doi: 10.1504/IJISD.2005.008087

Petersen, N. (2012). *Human Dignity, International Protection*. Max Planck Encyclopedia of Public International Law. Available online at: https://opil.ouplaw.com/abstract/10.1093/law:epil/9780199231690/law-9780199231690-e809?rskey=DtxXVW&result=1&prd=EPIL (accessed July 22, 2019).

Pike, S., and Dickson, F. (2018). *Identity on the Blockchain Special Report: Forecast and Analysis for Blockchain-Based Identity Solutions, Jul 2018 - Special Study - Doc # US44070617*. International Data Corporation.

Poleshchuk, V. (2016). *"Making Estonia Bigger": What E-Residency in E-Estonia Can Do for You, What It Can Do for Estonia*. Investment Migration Working Papers. Available online at: https://investmentmigration.org/download/making-estonia-bigger-e-residency-e-estonia-can-can-estonia/ (accessed August 8, 2019).

Privacy International (2019, September 26). Privacy International, Initial analysis of Indian Supreme Court decision on Aadhaar. *Privacy International Website*. Available online at: https://privacyinternational.org/long-read/2299/initial-analysis-indian-supreme-court-decision-aadhaar (accessed August 8, 2019).

Rattani, A., Derakhshani, R., and Ross, A. (eds.). (2019). "Introduction to selfie biometrics," in *Selfie Biometrics*, *Advances in Computer Vision and Pattern Recognition* (Cham: Springer), 1–18.

Saada, E. (2012). *Empire's Children – Race, Filiation, And Citizenship in the French Colonies*. Chicago, IL: University of Chicago Press.

Shachar, A. (2009). *The Birthright Lottery - Citizenship and Global Inequality*. Cambridge, MA: Harvard University Press.

Swiss Federal Supreme Court (1990). *116 Ia 359*.

Taylor, L., van der Sloot, B., and Floridi, L. (eds.). (2017). "Conclusion: what do we know about group privacy?" in *Group Privacy, Philosophical Studies Series* (New York, NY: Springer), 225–237. Available online at: https://www.springer.com/gp/book/9783319466064

Toth, K. C., and Anderson-Priddy, A. (2019). Self-Sovereign digital identity – a paradigm shift for identity. *IEEE Secur. Priv.* 17, 17–27. doi: 10.1109/MSEC.2018.2888782

Velthuijs, M. (2018, March 30). *Forus, Gemeente Zuidhorn, Berenschot, Platform Forus garandeert de betrouwbaarheid, SB1GL17005*. Veiligheid en toegankelijkheid.

Verhulst, S. G., and Young, A. (2018). Field Report - On the Emergent Use of Distributed Ledger Technologies for Identity Management. *GovTech report*. Available online at: https://blockchan.ge/blockchange-fieldreport.pdf (accessed August 7, 2019).

Wagner, K., Némethi, B., Renieris, E., Lang, P., Brunet, E., and Holst, E. (2018). *'Self-sovereign identity' Position Paper*. Blockchain Bundesverband. Available online at: https://www.bundesblock.de/wp-content/uploads/2018/10/ssi-paper.pdf (accessed August 7, 2019).

WEF (2018a). *Platform for Good Digital Identity*. World Economic Forum Web Portal. Available online at: https://www.weforum.org/projects/digital-identity (accessed August 7, 2019).

WEF (2018b). *Our Digital Future – Building an Inclusive, Trustworthy and Sustainable Digital Society*. World Economic Forum Web Portal. Available online at: https://www.weforum.org/reports/our-shared-digital-future-building-an-inclusive-trustworthy-and-sustainable-digital-society (accessed August 7, 2019).

World Bank (2018). *Identification for Development 2018 Annual Report*. World Bank ID4D program webpage. Available online at: https://id4d.worldbank.org/sites/id4d.worldbank.org/files/2018_ID4D_Annual_Report.pdf (accessed November 23, 2019).

Xu, B. (2018, April 5). Blockchain vs. Distributed Ledger Technologies. *Medium*. Available online at: https://media.consensys.net/blockchain-vs-distributed-ledger-technologies-1e0289a87b16 (accessed August 8, 2019).

Zelazny, F. (2012). *The Evolution of India's UID Program: Lessons Learned and Implications for Other Developing Countries, Policy Paper 008*. Washington, DC: Center for Global Development.

Zwitter, A. J., and Boisse-Despiaux, M. (2018). Blockchain for humanitarian action and development aid. *J. Int. Humanit. Action* 3, 1–16. doi: 10.1186/s41018-018-0044-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

PERSPECTIVE
published: 07 April 2020
doi: 10.3389/fbloc.2020.00015

Identity Management Systems: Singular Identities and Multiple Moral Issues

Georgy Ishmaev and Quinten Stokkink*

Department of Software Technology, Delft University of Technology, Delft, Netherlands

The paper examines some of the competing normative claims surrounding the development of Identity Management (IM) systems in general and Self-Sovereign Identity (SSI) systems in particular. It is argued that SSI developments should be assessed against the backdrop of IMs attempting to implement a global identity layer based on aggregated singular identities and reputation scores. It is also argued that this trend defines key ethical issues pertaining to the development of SSI systems. In order to explicate and evaluate these issues, the paper looks at the desirability of singular aggregated identities through the lens of moral-philosophical theories. It is argued that such an analysis strongly suggests moral desirability of a plural identities approach in SSIs that have built-in advantage for the implementation of the practical separation of identities.

Keywords: identity, privacy, autonomy, ethics, blockchain, decentralization

OPEN ACCESS

Edited by:
Oskar Josef Gstrein,
University of Groningen, Netherlands

Reviewed by:
Michael Cooper,
Independent Researcher, Denver, CO, United States
Jia Xu,
Independent Researcher, San Francisco, CA, United States

****Correspondence:***
Georgy Ishmaev
g.ishmaev@tudelft.nl

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *17 October 2019*
Accepted: *05 March 2020*
Published: *07 April 2020*

Citation:
Ishmaev G and Stokkink Q (2020) Identity Management Systems: Singular Identities and Multiple Moral Issues. Front. Blockchain 3:15. doi: 10.3389/fbloc.2020.00015

INTRODUCTION

Even within the scope of a single discipline the concept of identity often falls apart into numerous meanings and interpretations (Martin and Barresi, 2006). Any attempt to tackle and unify this concept into a single label within the scope of an interdisciplinary study is even less tangible task. It is unsurprising then that in the field of Self-Sovereign Identity (SSI) systems development quite often we encounter suggestions to abandon it altogether in favor of a more palpable definition like an identifier or an attribute (Grigg, 2019). The other proposed strategy to tackle this conceptual ambiguity is to claim the particular technical interpretation of identity as the most fitting one and simply go along with it (Ma et al., 2018). From the ethical perspective, both of these strategies are problematic in the context of systems managing human identities. Any such identity management (IM) system—no matter how narrow and technically focused the ambitions of its creators are—inevitably cuts into a gordian knot of ethical concerns regarding autonomy, self-determination, and self-identification of its users (Manders-Huits and Hoven, 2008).

In order to try and address these issues we might consider the relation between descriptive and normative concerns regarding the concept of personal identity. While being analytically distinct, these sets of problems are related in the form of a feed-back loop. Shoemaker and Tobia (2019) contemplate this strategy as a sort of "reflective equilibrium" where both the conceptualizations of personal identity and relevant ethical concerns are built in the light of one another. Our ethical concerns inform the strategies of conceptualization, and ontological insights on the nature of identity highlight how these ethical concerns should be addressed. Historically, the development of earlier modern identification solutions, such as passports, have been predominantly driven by the consideration of societal goods, sometimes expressed as government needs or wider

communitarian values (Lloyd, 2005). The later developments of digital identity management systems have also highlighted moral concerns pertaining to the individual values and human rights (Chaum, 1985; Shoemaker, 2010).

As the more recent developments demonstrate, the interplay between individual and communitarian moral values is still very much a defining characteristic of this field. The Aadhaar—universal and de-facto obligatory identity system based on biometric identification rolled out by the Indian government—is one such example. Dixon (2017) points out that this system, problematic from the privacy perspective, was justified to the general public largely on moral grounds, such as the necessity to prevent fraud in the distribution of state subsides. An even more vivid example of this trend is presented by the Chinese government project—"Social Credit System" (SCS). Unlike other state identity management systems, SCS goes beyond mere forensic purposes and implements an explicit system of scores for profiled citizens designed to reflect their social "trustworthiness" and eliminate "black sheep from the society" (Ohlberg et al., 2017; Engelmann et al., 2019).

SSI systems seemingly occupy a middle ground in this contest between communitarian and individual values. Proponents of these solutions argue that SSI systems can bring enhanced privacy, data security and full controls over their digital identities to individuals, combined with the reliable mechanisms of identification (Allen, 2017; Tobin and Reed, 2017; Ma et al., 2018). With the help of minimized private data disclosures and enhanced individual control over identity data these solutions, argue SSI designers, will reconcile social needs for the working identity management systems with individual rights to privacy and autonomy[1]. Interestingly enough, this aspiration to reconcile conflicting values mirrors the central point of arguments in the debates between the different moral-philosophical approaches to personal identity.

Of particular interest here is the narrative theory of identity most notably championed by MacIntyre (2007) and Speight (2015). The moral focus of his theory lies with the concerns of a distinctively communitarian character—responsibility for one's actions, accountability, and obligations toward others. In the imagined opposite corner, philosophical approaches to personal identity that highlight self-focused moral concerns: questions of self-determination and moral autonomy (Sen, 2007; Strawson, 2015). It would be of course a crude simplification and a great disservice to these intricate and elaborate theories to represent them as simply aligned along the axis of individual—communitarian values. Rather, it would be more appropriate to say that as the very phenomenon of personal identity itself reflects both individual and social aspects of a human life, these theories illuminate different aspects of the same phenomenon[2].

However, it is possible to highlight one particular point where these theories seem to be at odds. That is the question of whether singular identity—as opposed to the plural, multiple identities—could provide reconciliation between self-focused and others-focused moral concerns (Strawson, 2015). We will argue that the examination of this aspect of conflicting value claims can provide some helpful insights in the context of SSI systems. Through the lens of these moral-philosophical arguments we identify competing normative claims behind the development of IM systems and highlight ethical issues in this field that can and should be addressed by SSI solutions.

IDENTITY MANAGEMENT SOLUTIONS

To provide insights on the tension between the competing moral claims pertaining to identity management systems it is helpful first to consider key technological trends in this area. In fact it is possible to identify a single trend largely definitional both for the technical developments in the area of IM systems and social and ethical concerns associated with this field—the identity resolution problem. This problem has emerged as a rather innocuous and purely technical issue in the data base management and statistics as a problem of classification task whereby two or more entities (collections of attributes)—often from different databases—are matched together based on the similarity of their features (Edwards et al., 2016). This problem has also motivated the development of novel identity resolution techniques and tools assisted by the advancements in artificial intelligence.

An increasing volume of big data available from social networks and online services has enabled advertising companies such as Google and Facebook among others to track individuals both online and offline with ever-increasing precision (Zuiderveen Borgesius, 2016; Venkatadri et al., 2018). Furthermore, such tracking is combined with profiling—the aggregation of individuals' profiles enriched with demographic, financial, social, and behavioral data—performed without consent. Advanced identity resolution tools, the wealth of private data, and near monopolistic market positions have enabled the move by advertising companies and data brokers toward the development of global identity solutions based on singular aggregated identities (Wolfie and Spiekermann, 2016)[3]. Despite some public backlash, this global private data industry, which spans different industries and private-public partnerships with government agencies, only continues to grow (Cleland, 2018).

This background largely defines many of the normative claims surrounding justification for the SSI development such as bringing the ownership of online identities back to individuals, or taking control of identities away from corporations (Tobin and Reed, 2017). It can be argued though, that while such claims carry certain emotional and intuitive appeal, basic scrutiny reveals certain inconsistencies, given that the idea of identity

[1] At the same time, these systems are not fundamentally different from legacy identity management (IM) systems, considering that the identification of individuals is an explicit purpose of the SSIs, as compared to anonymity systems.

[2] For broader overview of a narrative theories of identity see Speight (2015).

[3] Both Facebook and Google should also be noted for their efforts to introduce end-user identity solutions, built on top of their massive private data silos—"Facebook connect" and "Google sign-up" respectively. These are sets of Application Programming Interfaces (API), that can be implemented by third party web-services (websites, apps, etc.) to let their visitors authenticate themselves using Facebook or Google identities.

ownership seems both conceptually and ethically problematic (Floridi, 2006). SSI systems, however, carry technical potential to address some of the more specific ethical issues pertaining to the field of IM systems. Minimization of private data disclosures, decentralization of private data storage, and practical separation of context specific identities—all those measures that can help to tackle non-consensual profiling of individuals by third parties.

Self-Sovereign Identity

To unfold promises of SSI solutions we need first to look into the basics of these systems. Unlike in the field of blockchain-based cryptocurrencies it is difficult to highlight one single project that could be representative of SSI technology in the same way as Bitcoin[4]. At the moment there are over a 100 different projects that employ blockchain technology to provide the functionality of digital identity in one form or another[5]. And considering that any SSI at this point is a bleeding edge technology, there are also no clearly established standards. Some noteworthy work in this area, however, is accomplished by the W3C Credentials Community Group[6]. Several concepts comprising the general idea of SSI technology present specific interest in the W3C model.

The starting point here is to consider that public/private key encryption underlying most of the online interactions (such as messaging) can also be used to establish identities of the interacting parties[7]. This can be done with the help of a Public Key Infrastructure (PKI) which enables the exchange of keys between the parties and links names to the specific keys. Traditional PKIs are managed by the centralized trusted parties, such as certificate authorities or messaging service providers. The first crucial concept in the SSI schema is the Decentralized Public Key Infrastructure (DPKI)—essentially a data base containing public keys. The main novelty of DPKI is that, using blockchain as a decentralized database, it can radically reduce reliance on trusted parties while at the same time ensuring security from manipulation, censorship, or compromise (Allen et al., 2015).

With the help of DPKI, identity owners can register their decentralized identities associated with public keys on the blockchain without dependance on any centralized registrars (thus "self-sovereignty"). Schematically it can be said that DPKI forms the base layer allowing for another key component of SSI system—decentralized identifier (DID). Defined as a technical standard, in its idea DID is similar to a Uniformed Resource Identifier[8]. DID, however, points to entities (endpoints associated with natural persons or organizations for instance) rather than Web resources. In itself, generic DID contains an identifying string of symbols as an ID index and metadata, together called the DID document—a machine readable structured piece of data—and metadata called the DID document. In its most basic form, this identification scheme can include ID strings as a designation of the owner, information about the context of identification, cryptographic methods of authentication (specific public keys), and pointers to the method of authentication (specific blockchain).

Such identities in themselves provide limited functionality of course. The third crucial concept of SSI, however, makes a significant difference: the capacity to issue verifiable credentials. From the user's point of view, a verifiable credential is a digital, cryptographically signed document containing certain claim(s) about its holder—such as being a of certain age or being licensed to operate a vehicle—essentially similar to physical credentials. Practically, verified credential implementation proposed by W3C uses DIDs as subjects of claims and DID documents as root records for digital identities. This scheme allows individuals in a privacy-preserving manner. An individual can potentially generate multiple DIDs for interactions with different parties, choose different parties to sign his/her verifiable credentials, and to present only specific verified claims (such as age) to minimize private data disclosures.

Singular Identities

This scheme highlights a crucial difference between SSI systems and centralized identity management systems where a single authority (whether a government office or company) serves as a root of trust for all identities and credentials within the system. More importantly, such an identification scheme provides an alternative to the model where an individual has to use a single identifier such as legal name, mobile number, or government-issued number, through a range of relations and interactions. Thus, with minimized private data disclosures and the generation of disposable identifiers, the SSI model can make identity resolution and consequent profiling by third parties more costly (but not impossible).

It is crucial to point out, however, that the problem of identity resolution has no purely technical solution as it ultimately rests on a number of social factors. As an interoperable and open-ended standard (and like any other software solution—malleable) SSI can be also implemented in a way that makes aggregation of profiles easier[9]. Economic and social adoption of particular SSI schemes, practicalities of users' behavior, design of user interfaces, and finally a resistance of entities interested in the preservation of their profiling capabilities—all these factors can have profound effects on the adoption of standards. This is a problem closely related to the much larger ongoing problem of "crypto-wars"—the continuing struggle between entities with different interests over the establishment of encryption standards (and regulation) on a global scale (Gasser et al., 2016).

[4]While conceptually SSI is conceived as technologically agnostic, all practical implementations currently are based on blockchain technology and in this paper term SSI refers to these solution.

[5]See list by Markus Sabadello: https://github.com/peacekeeper/blockchain-identity

[6]See https://www.w3.org/TR/vc-data-model/

[7]The method of two-key encryption (or asymmetric cryptography) can be used both to encrypt messages and sign them. For instance owner of key pair (public and private key) Alice publishes her public key, so that Bob or anybody else can use it to encrypt message in such a way that only Alice can decrypt it using private key. Or alternatively, Alice can sign a message with her private key, so that Bob using public key can verify that the message was indeed signed by her (given that Alice is a unique holder of private key).

[8]Common example of a Uniformed Resource Identifier is a simple URL, e.g. "www.example.com"

[9]See "Blockchain in Ad Tech," available online at: https://www.acxiom.com/wp-content/uploads/2017/12/AC-1752-17-3-Point-of-View-Blockchain-in-Ad-Tech.pdf

However, compared to the debate on the moral desirability of strong encryption, the debate on the moral desirability of multiple identities has not gained similar scale yet. Up to date it remains predominantly one-sided, presented mostly by the position of the proponents of a singular identity approach. One example of such justificatory reasoning is a widely cited statement by Facebook's founder Mark Zuckerberg: "Having two identities for yourself is an example of a lack of integrity" (Kirkpatrick, 2011). This thesis on the moral desirability of a singular identity—a "real name" policy—is also a recurring topic in the criticism of anonymity online. Government policy proposals on the mandatory identification for internet services can be found across a range of countries with very different legal and cultural traditions such as Austria and China[10]. There are good reasons, thus, to examine the moral-theoretical foundations of these claims.

VALUE OF IDENTITY FOR WHOM?

To gain some clarity on the question of singular identities we can consider basic concepts and normative premises. First point of consideration here is an epistemic asymmetry between what can be called a first-person and a third-person view of one's identity. Indeed, no matter how accurate a description of a person can be including one's appearance, behavior, habits, and beliefs such a description is inevitably incomplete compared to the sum of experiences, memories and beliefs about oneself experienced by an individual (Manders-Huits and Hoven, 2008). In the context of IM systems this principle highlights the risk of an imposition of purely administrative notion of identity and a reductionist treatment of individual users as mere objects of computation (Manders-Huits, 2010). This observation on the epistemically privileged position translates into the claim that an individual should have a say in the construction or interpretation of one's identity in IM.

This principle in itself, however, does not provide arguments on the moral value of a singular identity. It can be argued that as long an individual has a say in the information associated with one's identity the principle is satisfied, whether it is a singular aggregated identity or not. What is at stake here is the practical question of person's re-identification across a range of contexts and scenarios. In that sense it is not only the question of a tension between first-person and third-person views, but also the question of tension between self-centered moral concerns and others-focused moral concerns. And it would be wrong to consider this distinction in an adversarial framework of a naive Hobbesian world dominated by the clash between competing egoistic concerns. It is more of a question whether the singular identity approach can strike a balance between self-focused and others-focused ethical concerns.

[10]See https://www.derstandard.at/story/2000101677286/government-seeks-to-eliminate-internet-anonymity-with-severe-penalties; https://techcrunch.com/2017/08/27/china-doubles-down-on-real-name-registration-laws-forbidding-anonymous-online-posts/

Moral Value of a Singular Identity

Probably some of the most influential moral-philosophical arguments in the support of such a view are proposed by MacIntyre (2007). He takes a radical stance on the necessity of a singular narrative identity as a focal point of moral concerns, grounded in the ideals of the virtues of antiquity. According to MacIntyre, there is simply no moral identity for the abstract individual, since the self finds its moral identity in and through membership in communities. A unified narrative—the story of one's life—is something that both defines and addresses the tension between self-regarding concerns of moral autonomy and concerns regarding one's accountability for past actions. Through the prism of shared norms and associated beliefs narrative self provides the intelligibility of an individual's action for others and for the owner of these actions.

Building on this reasoning MacIntyre makes his arguments in favor of a unified, singular identity as morally good and desirable, in juxtaposition to the idea that one can entertain multiple roles and multiple identities. Fragmentation of self-identity into a set of demarcated areas of role-playing, argues Macintyre, allows no scope for the exercise of dispositions which could genuinely be accounted as virtues in any sense. Only those traits of character that can be manifested consistently throughout the range of contexts and relations amount to something that contributes to the moral self. It is not difficult to see a parallel in this line of thinking with the proposals on the development of social reputation systems (Ohlberg et al., 2017; Engelmann et al., 2019). Indeed, any society-wide IM system based on singular identities and reputations provides a unified, cross-context prism for the normative assessment of an individual's actions and behaviors. More so, many of the ongoing developments in the area of such identity management systems seem to mirror the same moral arguments implicitly or explicitly.

Accordingly, the critique of the narrativist arguments on the moral desirability of a unified identity, can help to highlight key moral issues of IM systems built around persistent singular identities. The first problematic issue here is the question of a choice of a unifying normative framework for the evaluation of one's identity. An immediate concern here is that such a prism can be unfair and unacceptable, if it is designed or distorted in such a way that it serves the interests of particular parties only. Indeed, as Grigg (2019) notes, too often the interests of entities controlling IM systems seem to replace genuine community-defined values. The deeper issue here, however, is that even in the absence of a self-interested entity defining a normative framework for the assessment of identities, such a singular framework in itself is morally problematic.

Moral Autonomy of Identity

Sen (2007) illustrates this problem with the observation that any singular framework for the evaluation of identity can be reductionist, biased, or meaningless once it is translated into a different context. As he argues, each of the collectivities (professional, religious, cultural etc.) to all of which an individual may simultaneously belong, give him or her a particular identity. Accordingly, each of these particular identities may presuppose varying or even

competing evaluative frameworks. This problem becomes apparent when we consider the cases when individuals' activities on social media cause them to lose their jobs or make them victims of misguided legal repercussions (Mantouvalou, 2019). Similarly, morally problematic conflicts between evaluative frameworks occur when economic evaluations come into contradiction with human rights (Rotenberg and Seon Kang, 2018).

A unified, singular set of evaluative norms, formalized in a reputation system, simply cannot grasp the complexity and multiplicity of contexts in which individuals make choices and exercise their moral autonomy. This is a fundamental issue going back to the need of respect for the uniqueness of a first-party perspective of oneself. It is too easy to classify others, but valid moral judgement respectful of the principles of moral autonomy is hard. As Strawson (2015) aptly notices, very often the reasoning on the value of identity goes together with a: "fabulously misplaced confidence that the elements of experience that people consider fundamental for their own experiences must be also fundamental for everyone else." And as some empirical studies suggest this bias might be widespread and inseparable from the deeply embedded evaluative character of social identities (Strohminger et al., 2017).

More importantly, this is not merely an issue of biases, or unfair judgments but an issue that goes to the core of the principle of moral autonomy (Manders-Huits, 2010). The more diachronically persistent an identity is across the range of social contexts, the more likely it is to accumulate conflicting normative judgments. An individual burdened with an ever-increasing weight of conflicting moral judgments on the value of one's identity either falls into conformity or becomes paralyzed by the inability to make genuine moral choices. The only feasible way to address this issue is to provide viable alternatives to singular persistent identities that follow individuals across all contexts of their lives. Multiple identities separated by the contexts of social interactions can provide an escape from this impasse, and contrary to the arguments on the lack of moral integrity attributed to such multiplicity, it is in fact a necessary prerequisite for the construction of a moral self in the globalized world of conflicting normative frameworks.

CONCLUSION

This paper has highlighted the connection between the question of a singular identity in practical IMs development and some of the established traditions in the moral theories of identity. The engagement with the moral-philosophical approaches to personal identity helps to map and disentangle some of the ethical concerns related to SSI solutions. The prominent position here takes the problem of conflicting claims on the moral desirability of a singular persistent identity. On one hand this is the focal point of ethical concerns associated with the development of IMs in general, highlighting the issues of profiling, privacy, autonomy, and freedom. On the other, in the context of SSI, this issue pinpoints both the main promise and the most realistic pitfall that can undermine moral desirability of such systems.

SSIs are often presented under the general promises of bringing controls and ownership of identities to the individuals. We have argued that such overly generalized claims tend to mislead the debate and attention from some of the more specific considerations. The most desirable feature of SSI systems is an ability to provide individuals with the freedom to exercise multiples identities in different contexts and relations. This capacity can help to address issues of non-consensual profiling and detrimental effects of reputation systems. And conversely, in the absence of such functionality, all other features such as minimized data disclosures, local storage of private data, and decentralized key management will lose ethical significance in being reduced to marketing slogans. The provided theoretical analysis gives a justification to this argument, and also aims to steer the debate in the direction of the more explicit considerations of the ethical aspects of SSI development.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

Allen, C. (2017). *The Path to Self-Sovereign Identity*. Available online at: https://github.com/WebOfTrustInfo/self-sovereign-identity/blob/master/ThePathToSelf-SovereignIdentity.md

Allen, C., Brock, A., Buterin, V., Callas, J., Dorje, D., Lundkvist, C., et al. (2015). *"Decentralized Public Key Infrastructure." A White Paper from Rebooting the Web of Trust*. Available online at: https://www.weboftrust.info/downloads/dpki.pdf

Chaum, D. (1985). Security without identification: transaction systems to make big brother obsolete. *Commun. ACM* 28, 1030–1044. doi: 10.1145/4372.4373

Cleland, S. (2018). *The Stunted State of U.S. Antitrust Enforcement of Internet Platforms. Submission for: U.S. FTC Fall 2018 Hearings on "Competition & Consumer Protection in the 21st Century"*. Available online at: https://www.ftc.gov/system/files/documents/public_comments/2018/08/ftc-2018-0048-d-0023-151008.pdf

Dixon, P. (2017). A failure to "do no harm"—India's Aadhaar biometric ID program and its inability to protect privacy in relation to measures in Europe and the U.S. *Health Technol.* 7, 539–567. doi: 10.1007/s12553-017-0202-6

Edwards, M., Wattam, S., Rayson, P., and Rashid, A. (2016). "Sampling labelled profile data for identity resolution," in: *2016 IEEE International Conference on Big Data (Big Data)*, 540–547.

Engelmann, S., Chen, M., Fischer, F., Kao, C., and Grossklags, J. (2019). Clear sanctions, vague rewards: how China's social credit system currently defines 'good' and 'bad' behavior. *Proc. Confer. Fairness Account. Trans. FAT** 19, 69–78. doi: 10.1145/3287560.3287585

Floridi, L. (2006). Four challenges for a theory of informational privacy. *Ethics Information Technol.* 8, 109–119. doi: 10.1007/s10676-006-9121-3

Gasser, U., Gertner, N., Goldsmith, J. L., Landau, S., Nye, J. S., O'Brien, D., et al. (2016). *Don't Panic: Making Progress on the "Going Dark" Debate.*

Available online at: https://cyber.harvard.edu/pubrelease/dont-panic/Dont_Panic_Making_Progress_on_Going_Dark_Debate.pdf

Grigg, I. (2019). *Why We Must Ask the Why of Identity*. Available online at: https://github.com/WebOfTrustInfo/rwot9-prague/blob/master/topics-and-advance-readings/ask-why.md

Kirkpatrick, D. (2011). *The Facebook Effect: The Inside Story of the Company That Is Connecting the World*. New York, NY: Simon & Schuster Paperbacks.

Lloyd, M. (2005). *The Passport: The History of Man's Most Travelled Document*. Sutton.

Ma, M., Remore, C., Gisolfi, D., Kussmaul, W., and Greening, D. (2018). *SSI: A Roadmap for Adoption*. Avialble online at: https://github.com/WebOfTrustInfo/rwot6-santabarbara/blob/master/final-documents/a-roadmap-for-ssi.pdf

MacIntyre, A. C. (2007). *After Virtue: A Study in Moral Theory, 3rd Edn*. Notre Dame, IN: University of Notre Dame Press.

Manders-Huits, N. (2010). Practical versus moral identities in identity management. *Ethics Information Technol.* 12, 43–55. doi: 10.1007/s10676-010-9216-8

Manders-Huits, N., and Hoven, J. (2008). "Moral identification in identity management systems," in *The Future of Identity in the Information Society*, eds S. Fischer-Hübner, P. Duquenoy, A. Zuccato, and L. Martucci (Springer US), 77–91.

Mantouvalou, V. (2019). 'I Lost My Job over a Facebook Post: Was that Fair?' Discipline and dismissal for social media activity. *Int. J. Comp. Lab. Law Ind. Relat.* 35, 101–125.

Martin, R., and Barresi, J. (2006). *The Rise and Fall of Soul and Self: An Intellectual History of Personal Identity*. New York, NY: Columbia University Press.

Ohlberg, M., Ahmed, S., and Lang, B. (2017). *Central Planning, Local Experiments. The complex implementation of China's Social Credit System*. MERICS Mercator Institute for China Studies. Available online at: https://www.merics.org/sites/default/files/2017-12/171212_China_Monitor_43_Social_Credit_System_Implementation.pdf

Rotenberg, M., and Seon Kang, S. (2018). *Comments of the Electronic Privacy Information Center Federal Trade Commission Hearings on Competition and Consumer Protection in the 21st Century Question 9: Consumer Welfare Implications Associated with the Use of Algorithmic Decision Tools, Artificial Intelligence, and Predictive Analytics*. Available online at: https://epic.org/apa/comments/EPIC-FTC-Algorithmic-Transparency-Aug-\penalty-\@M20-2018.pdf

Sen, A. (2007). *Identity and Violence: The Illusion of Destiny, 1 st Edn*. New York, NY: Norton.

Shoemaker, D., and Tobia, K. P. (2019). *Personal Identity. Oxford Handbook of Moral Psychology, Forthcoming*. Available online at: https://ssrn.com/abstract=3198090

Shoemaker, D. W. (2010). Self-exposure and exposure of the self: Informational privacy and the presentation of identity. *Ethics Information Technol.* 12, 3–15. doi: 10.1007/s10676-009-9186-x

Speight, A. (2015). "The Narrative Shape of Agency: three contemporary philosophical perspectives," in *Narrative, Philosophy and Life*, *Vol. 2*, ed A. Speight (Dordrecht: Springer Netherlands), 49–60.

Strawson, G. (2015). "Against narrativity," in *Narrative, Philosophy and Life, Vol. 2*, ed A. Speight (Dordrecht: Springer Netherlands), 11–31.

Strohminger, N., Knobe, J., and Newman, G. (2017). The true self: a psychological concept distinct from the self. *Perspect. Psychol. Sci.* 12, 551–560. doi: 10.1177/1745691616689495

Tobin, A., and Reed, D. (2017). *The Inevitable Rise of Self Sovereign Identity*. A White Paper From the Sovrin Foundation. Available online at: https://sovrin.org/wp-content/uploads/2017/06/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf

Venkatadri, G., Andreou, A., Liu, Y., Mislove, A., Gummadi, K. P., Loiseau, P., et al. (2018). "Privacy risks with Facebook's PII-based targeting: Auditing a data broker's advertising interface," in *2018 IEEE Symposium on Security and Privacy (SP)*, 89–107.

Wolfie, C., and Spiekermann, S. (2016). *Networks of Control: A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy*. Facultas. Available online at: https://crackedlabs.org/dl/Christl_Spiekermann_Networks_Of_Control.pdf

Zuiderveen Borgesius, F. J. (2016). Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. *Comput. Law Security Rev.* 32, 256–271. doi: 10.1016/j.clsr.2015.12.013

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

ORIGINAL RESEARCH
published: 07 April 2020
doi: 10.3389/fbloc.2020.00004

The Private Governance of Identity on the Silk Road

*Catalina Goanta**

Law & Tech Lab/Studio Europa, Maastricht University, Maastricht, Netherlands

Long before the creation of blockchain platforms, the rise of personal computing, and Internet connectivity brought with it a digital, online dimension of the material world, leading to the socio-technical construct known as "digital identity." After the online discussion boards and emailing lists of the early 1990s, individuals started socializing via the Internet more predominantly using social networks. One specific type of platform links this online socializing and transacting to blockchain-based spaces: dark web marketplaces. Identified as second-generation cryptocommunities, dark web marketplaces deployed cryptography for the use of pseudonymous identity, for communication, but also currency. This paper explores two questions in this fascinating space: what was the role of identity on the Silk Road, and what governance lessons can be drawn from this illustration for the purpose of applying them to more recent cybercommunities such as Ethereum? The paper is structured as follows. The first part describes the Silk Road and sketches its essential characteristics. The second part looks at how individuals could become platform users on the Silk Road, by analyzing the contractual relationship between the Silk Road and an individual user based on the rights and obligations enshrined in the Silk Road terms of service (ToS). The third part critically reflects on arbitrariness as the main pitfall arising out of the private regulatory framework created by the Silk Road, and contributes to existing narratives surrounding the regulatory nature of code by proposing a code-as-procedure perspective for analyzing this regulatory framework. Part four concludes.

Keywords: dark markets, private governance, code is law, Silk Road, procedural law

OPEN ACCESS

Edited by:
Oskar Josef Gstrein,
University of Groningen, Netherlands

Reviewed by:
Richard Tighe,
Oxfam, San Francisco,
United Kingdom
Jia (Carol) Xu,
Independent Researcher, San
Francisco, CA, United States

***Correspondence:**
Catalina Goanta
catalina.goanta@
maastrichtuniversity.nl

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *25 September 2019*
Accepted: *28 January 2020*
Published: *07 April 2020*

Citation:
Goanta C (2020) The Private Governance of Identity on the Silk Road. Front. Blockchain 3:4. doi: 10.3389/fbloc.2020.00004

INTRODUCTION

The inescapable interest in blockchain technology seen in the past years has ignited a lot of debate surrounding the decentralization of established legal concepts and institutions. One such example reflects discussions around the concept of identity (e.g., nationality, citizenship, or broadly speaking membership to a legally defined group), as well as the institutions administrating various aspects of identity (e.g., state agencies conferring or depriving individuals of nationality). As decentralization is seen as empowering individuals to give up the use of or dependence on intermediation (be it private or public), it gave rise to the notion of self-sovereign identity systems which ought to preserve an individual's self-determination in providing, or even expanding, the benefits of record-keeping.

This discussion has prompted a lot of interdisciplinary literature, looking at the broader themes of e-government (Reijers et al., 2016; Augot et al., 2017; Hou, 2017; Sullivan and Burger, 2017) and

smart cities (McMillan, 2014; Biswas and Muthukkumarasamy, 2016; Ibba et al., 2017; Jaffe et al., 2017; Rivera et al., 2017; Sharma et al., 2017; Marsal-Llacuna, 2018) or particularly the use of self-sovereign identity systems in development aid [e.g., the United Nations High Commissioner for Refugees (UNHCR) using blockchain to manage the identity of refugees] (Biometric Technology Today, 2017; Mears, 2018), the privacy issues posed by the use of public blockchains in the management of identity (Zyskind et al., 2015; Zhang et al., 2017; Yang et al., 2018), or the use of decentralization as a means of breaking socio-legal constructs that lead to, for example, global inequality (Freund, 2017; Michaels and Homer, 2017). Most of this literature, whether reflecting legal, sociological, or economic analyses, focuses on recent platforms such as Pavilion.io, Mattereum, or Stampery (Casino et al., 2019). However, long before the creation of these blockchain platforms, the rise of personal computing and Internet connectivity brought with it a digital, online dimension of the material world, leading to the socio-technical construct known as "digital identity" (Lemieux, 2016; Dunphy, 2018). After the online discussion boards and emailing lists of the early 1990s, individuals started socializing via the Internet more predominantly using social networks (Can and Alatas, 2019). One specific type of platform links this online socializing and transacting to blockchain-based spaces: dark web marketplaces. Identified as second-generation cryptocommunities[1] (Goanta and Hopman) dark web marketplaces deployed cryptography for the use of pseudonymous identity, for communication, but also for currency[2].

The most prominent example of such a marketplace is the Silk Road, a space only reachable through the use of The Onion Router browser (TOR) (AlQahtani and El-Alfy, 2015), where an administrator with cyberlibertarian views by the name of Ross Ulbricht managed the first iteration of a multimillion dollar illegal marketplace. While dark web user identity is not as such self-sovereign, understanding how Ross Ulbricht's platform managed the identity of registered users can provide useful insights into blockchain governance problems. Just like many well-known "idols" of the contemporary blockchain space, Ross Ulbricht's libertarian views made him especially allergic to the notion of state law limiting individual freedoms and argued not only for the reduction of government interventionism but also for the potential replacement of state law with the private rules of a community that took individual freedom as the most important value in determining its own functioning[3] (Greenberg, 2012; Bartlett, 2015).

This paper explores two questions in this fascinating space: what was the role of identity on the Silk Road[4] and what governance lessons can be drawn from this illustration for the purpose of applying them to more recent cybercommunities such as Ethereum? The paper is structured as follows. The first part describes the Silk Road and sketches its essential characteristics. The second part looks at how individuals could become platform users on the Silk Road, by analyzing the contractual relationship between the Silk Road and an individual user based on the rights and obligations enshrined in the Silk Road terms of service (ToS). The third part critically reflects on arbitrariness as the main pitfall arising out of the private regulatory framework created by the Silk Road and contributes to existing narratives surrounding the regulatory nature of code by proposing a code-as-procedure perspective for analyzing this regulatory framework. Part 4 concludes. From a methodological perspective, this paper is based on the qualitative analysis of all the documents identified by the Government of the United States as the Silk Road ToS in Ulbricht's initial indictment from 2014[5].

FEATURES OF THE SILK ROAD AS A CRYPTOCOMMUNITY

According to the United States government, the Silk Road was a dark web marketplace created by a United States citizen (Ross Ulbricht) in order to facilitate the transacting of illegal (and legal) items such as drugs[6]. The first iteration of the Silk Road was online between 2011 and 2013 (Christin, 2012), when Ulbricht started popularizing the nascent platform on various web forums to facilitate the sale of self-produced hallucinogenic mushrooms[7]. At the height of its popularity, the Silk Road managed to bring together up to 150,000 active users, mostly from the United States (see **Figure 1**). The most commonly sold product on the Silk Road was, by far, weed (see **Figure 2**).

[1] A cryptocommunity is a virtual community where cryptography is used to ensure the "security of identity, communication, currency, or more recently, value" and for the creation and/or support of political ideologies. See C. Goanta and M. Hopman, "Cryptocommunities as legal orders" 3.

[2] *Ibid.*

[3] For a general overview of the Silk Road, see A. Greenberg, *This Machine Kills Secrets: How Wikileakers, Cypherpunks, and Hacktivists Aim to Free the World's Information* (Dutton 2012); J. Bartlett, *The Dark Net* (Melville House 2015).

[4] It is important to mention that all references to Silk Road in this paper are used to designate the first iteration of this dark web marketplace. There were at least three more iterations.

[5] *United States of America v Ross William Ulbricht*, Indictment, District Court, Southern District of New York, 21 August 2014, 14 Cr. 68.

[6] *Ibid.*, Government exhibit 226D.

[7] Bartlett, fn 12, at 137.

Origin		Acceptable destinations	
Country	**Pct.**	**Country/Region**	**Pct.**
U.S.A.	43.83%	Worldwide	49.67%
Undeclared	16.29%	U.S.A.	35.15%
U.K.	10.15%	European Union	6.19%
Netherlands	6.52%	Canada	6.05%
Canada	5.89%	U.K.	3.66%
Germany	4.51%	Australia	2.87%
Australia	3.19%	World. excpt. U.S.A.	1.39%
India	1.23%	Germany	1.03%
Italy	1.03%	Norway	0.70%
China	0.98%	Switzerland	0.62%
Spain	0.94%	New Zealand	0.56%
France	0.82%	Undeclared	0.26%

FIGURE 1 | Shipping origin and destination (Christin, fn 14, at 9).

Category	#. items	Pct.
Weed	3338	13.7%
Drugs	2194	9.0%
Prescription	1784	7.3%
Benzos	1193	4.9%
Books	955	3.9%
Cannabis	877	3.6%
Hash	820	3.4%
Cocaine	630	2.6%
Pills	473	1.9%
Blotter (LSD)	440	1.8%
Money	405	1.7%
MDMA (ecstasy)	393	1.6%
Erotica	385	1.6%
Steroids, PEDs	376	1.5%
Seeds	374	1.5%
Heroin	370	1.5%
DMT	343	1.4%
Opioids	342	1.4%
Stimulants	291	1.2%
Digital goods	260	1.1%

FIGURE 2 | Top 20 product categories of items available (Ibid).

The Silk Road was partially fueled by a revolutionary vision. For the cypherpunks of the late 1980s who were the first to set libertarian ideals in cyberspace (May, 1992), the libertarian vision of freedom entailed removing the state from the affairs of its citizens. This very idea was taken over by the Silk Road, where it further developed in the wake of new tools (e.g., cryptocurrencies and hidden network services). Using those tools, Ulbricht and his helpers managed to usher in a new expression of libertarianism, where the community was mostly free to enter into transactions that states would not otherwise recognize as lawful. Still, not all members of the community shared the revolutionary vision. Given the behavioral diversity of the Silk Road's members, it comes as no surprise that not all of them believed in the platform's core philosophy. Some members show abundant support for the movement behind the platform; yet others see it as a one-stop-shop for drug commerce, and nothing more[8].

The Silk Road's effectiveness was primarily based on trust. The essential "technology" that made transactions possible between strangers who did not know or trust each other was not necessarily the cryptocurrency they were paying with—although this did make their interactions possible—but rather the reputational mechanisms that created behavioral incentives for users, both sellers and buyers, to conduct business within the parameters set by the creator of the system.

In addition, the Silk Road operated in a very intense adversarial environment. Cryptocommunities are innately built on the premise that there is a malicious entity trying to prevent the system from achieving its functions, and this is expressed in a cat-and-mouse setup between the actors of the system. For every solution an actor comes up with, there will be others trying to undermine it. For cypherpunks, the adversary was the arm of the state, which at times was real and frightening and threatened the livelihood of the group's members[9]. This tension also extended to dark markets, with one difference: because dark markets started gathering and trading in wealth, in the form of cryptocurrencies, this drew the interest of a new type of adversary—individuals or groups, with no allegiance to the state, who either were direct competitors or simply followed personal purposes (whether for entertainment, financial gain, or both) in hacking market participants, including platforms[10]. These attacks often took place in the form of phishing, where hackers would for instance make mirrors of the Silk Road website and ask users to log in, gaining access to their accounts, as well as any information seen by that user's account. This feature was further consolidated into decentralized platforms as well. At the moment, Ethereum's main adversary is not the state, but the overabundance of similar platforms and developers who might have a stake in bringing the platform down or simply drying it of its funds, as was the case with the decentralized autonomous organization (DAO) attack in 2016 (Metjahic, 2018).

On the Silk Road, only a small community had high technology literacy. Whether it entailed knowing how to operate the different cryptographic tools available on the hidden network (e.g., not falling prey to phishing attacks on TOR) or understanding the algorithms calculating the seller reputation rate or the Silk Road's fees, it becomes very clear from the forum posts of the Silk Road's first iteration that the overwhelming majority of users are in the dark[11]. This effect was most likely worsened by the operation of constant changes by the platform, as well as by the high volatility of the Bitcoin market. All these features together divided the community into two categories: the core users who understood the infrastructure of the system and its components and the users who gave up trying to understand these matters and simply relied on the user-friendly interface to get their business done. The more sophisticated these tools get in different iterations of cryptocommunities, the bigger the gap between those who know how to work with and around them and those who remain illiterate, because the cost of becoming educated on this matter might be too high, thus leading to an indirect knowledge centralization creep. In addition, as a direct result of their high behavioral heterogeneity, cryptocommunities based on decentralized platforms have been developing a fuzzy jargon that frustrates the process of gaining technological literacy (Walch, 2017).

Although a lot of the activity on the Silk Road was based on human decision making, the platform developed strong technocratic institutions. The Silk Road administrators made an effort to develop an elaborate set of legal rules to keep order in the community. By October 2011, Ulbricht had between a team of two and five administrators to run the platform, deal with complaints, resolve disputes, moderate the forum, and

[8]Goanta and Hopman, fn 9.

[9]May, fn 19.

[10]Bartlett, fn 12, at 138.

[11]Goanta and Hopman, fn 9.

track down law enforcement infiltration[12]. However, given that this team was running a website, its policy implementations were fundamentally technocratic. This led to the development of technocratic institutions, such as the algorithmic reputational mechanisms. The community was based on a set of rules, but the procedural implementation of these rules in the form of enforcement mechanisms such as algorithmic reputation systems and account management was inconsistent, as clear procedures were absent.

THE PRIVATE GOVERNANCE OF USER TRANSACTIONS ON THE SILK ROAD: RULES

That the Silk Road has libertarian roots is undeniable. Given that no recognized state in the world currently in existence can be described as a libertarian legal system, the principles behind it retain a highly philosophical dimension. In the light of this characteristic, one would expect whatever system of rules was created by the platform's operators and adhered to by the community to be a set of social and legal norms consistent with this core libertarian vision. Yet a large part of the rules applicable to the Silk Road are rules that are currently in force in national and supranational legal orders and are not just philosophical in nature. For this reason, some rules might conflict with the libertarian order, which raises the question of whether this confusion was a result of misunderstanding the role and infrastructure of legal systems in the first place. This part looks into the nature of the rules found at the core of the Silk Road operations, which can be entirely found in the Supplementary Annex 1–5 to this paper.

Charter

The Silk Road Charter is the equivalent of the platform's constitution, as it lays down its purpose and fundamental values. The purpose of the Silk Road is to provide "systems and platforms" to "customers," in order to empower them to "live as free individuals[13]." The Charter also mentions that in doing so, the Silk Road engages in the protection of "basic human rights," although these rights are not defined any further. What is, however, defined, is a set of five fundamental values.

Self-Ownership

This value revolves around the property rights cast upon individuals, which include their bodies, thoughts, and will but also "anything they create with their property or obtain without coercion." The concept of owning one's person echoes Locke's writing: "though the earth, and all inferior creatures, be common to all men, yet every man has a property in his own person: this nobody has any right to but himself. The labor of his body, and the work of his hands, we may say, are properly his" (Locke, 1821). The same ideas are also reflected in the work of Nozick and Rothbard (Nozick, 1977; Rothbard, 1978; Cohen, 1995; Van Parijs, 1995). There are however, no legal systems that embraced the idea of a right of self-ownership as described by libertarian philosophical literature, partially because the libertarian ideal of self-ownership has been labeled as a "self-defeating theory when we consider the operability and usefulness of the rights it bestows upon those who have no original resources to trade" (Cleaver, 2011). In the context of the Charter, self-ownership is an expression of an absolute personal freedom.

Responsibility

The Charter clarified the notion of responsibility by placing it in the context of accountability: "If one infringes on another's rights, they should be held accountable." The Silk Road system thus acknowledges that an infringement of rights created in this order must lead to punishment. What is unclear, however, is how the punishment is determined and who establishes and enforces it. In formal legal orders, the state holds the monopoly over the exercise or threat of violence, which it implements through, for instance, access to justice. As formal legal orders are not acknowledged in the Charter, there are two ways in which this fundamental value can be interpreted. It can first be interpreted to say that users do have access to justice, albeit private justice. In this reading, accountability takes place through a right of retribution of the wronged party. Another option is to interpret this value as authorizing the operators of the platform to act as the guardians of these values and thus penalize behavior going outside of its mandated limitations.

Equality

Equality is used in this ecosystem to express its decentralized nature, or namely, the fact that the platform's sovereign is not considered to have the same authority as the state in a peer-to-peer platform. The value of equality is an example of internal inconsistency within the Charter, as it contradicts the notion of self-ownership. The latter is based on the consideration that you can own your own body, as well as what you make and what you conquer without using force. Applied to a reality of limited resources, this idea entails that at some point in time, all resources will have been conquered and that newcomers will not have anything to conquer anymore. However, this conflicts with the principle of equality, as it expressly envisages property rights.

Integrity

The Charter defines integrity as honoring one's word. Put differently, this fundamental value embodies the centuries-old contractual principle of *pacta sunt servanda* (Wehberg, 1959; Jeremy, 2000; Mazzacano, 2011), namely, that all promises made need to be kept. In the UNIDROIT Principles of International Commercial Contracts, for instance, this principle has led to the development of specific performance remedies (Gebhardt, 1947; Vlavianos, 1993; Dizgovin, 2016). In the context of the Silk Road, this entailed that buyers could have a right of replacement in case ordered products did not meet the buyer's expectations.

Virtue

The last fundamental value is phrased in a rather confusing way: "to improve one's self and the lives of others in all actions." On the

[12] Bartlett, fn 12, at 138.

[13] Supplementary Annex 1.

one hand, it seems to refer to personal development, but on the other hand, it also refers to development as communal progress.

Terms of Service

The Silk Road ToS comprise three different documents: the Seller's Contract (Supplementary Annex 2), the Seller's Guide (Supplementary Annex 3), and the Buyer's Guide (Supplementary Annex 4). These resources bring to light quite a considerable volume of contractual rules, primarily between the platform and the sellers, as well as between the sellers and the buyers.

Seller's Contract

The Seller's Contract was a nine-line piece of text displayed to a platform user when they registered as a seller and focused on essential obligations the seller was going to be held accountable for. Most prominently, data protection and information duties (e.g., packaging and product information) references can be found therein, as follows:

Data Protection

In a way, sellers can be considered the processors of buyers' personal data, such as the buyer's shipping address. This is referred to as client anonymity. While some buyers would give fake addresses, some would use their real ones, so this posed great risks for the whole operation. The seller's role was thus to safeguard this information, and they were under a strict obligation to destroy the client's shipping address "as soon as it is used to label the package[14]."

Information Duties

In their contract with the sellers, the platform mandated the latter to bear the burden of obtaining information on matters such as staying "up-to-date on the latest stealth shipping methods" or informing themselves of the further obligations outlined in the Seller's Guide (Eisenberg, 2003; Bar-Gill and Porat, 2017).In addition, not only did sellers bear the burden of obtaining information, but they were also subject to mandatory disclosures, since they were expected to "describe the items accurately and truthfully." In a way, this clause is reminiscent of an obligation of performing their contract with the Silk Road in good faith. This is seconded by a moral norm imposed on the sellers, that of treating customers with respect, so to "go above and beyond for them."

Failure to comply with these rules resulted in the vendor account being banned. The Silk Road seemed to have formalized this agreement with its sellers by retrieving consent in a very explicit manner: "By clicking 'I agree' at the bottom, you agree to abide by the guidelines and terms below when selling on Silk Road."

Seller's Guide

As mentioned in the Seller's Contract, sellers had to inform themselves of the obligations outlined in an additional document which was available on the Silk Road's wiki. The need for the centralization of rules and procedures used on the Silk Road was felt by the community as a whole, and the emergence of a wiki page was announced by Dread Pirate Roberts (DPR) at an early stage of the platform's existence: "Hey gang, I want to set up a Silk Road wiki so we can have all of the FAQ's answered in one spot and hopefully remove some of the clutter of repeated questions from the forum. Anyone who's set up a wiki before want to administer this? I could do it, but I want to get more community members involved and a wiki is a great community project anyway[15]."

The final wiki was ready by November[16], and it served as an information management resource users could inspect to understand how different aspects of the platform worked. At the same time, the administrators who had been appointed by DPR to compile the wiki also added—most likely with his permission—the Seller's and the Buyer's Guides.

As far as the seemingly legal obligations in the Seller's Guide are concerned, they were not few.

Data Protection

The obligation of destroying any shipping address information once the package was shipped is complemented with the prohibition of obtaining any personal information from the buyers. The action of saving customer addresses could lead to the revocation of seller privileges.

Obligations Relating to Payment

The processing of payment was the Silk Road's monopoly, expressed in the escrow system, which entailed that any payment had to be held by the platform until the buyer notified the receipt of the goods, at which point the payment would be released to the seller. The Seller's Guide makes it clear that if payments do not go through the escrow system, this can cost sellers their accounts. There were two exceptions to when party agreement overruled this principle: (i) the website being down and (ii) "closing early," namely, releasing funds from escrow before the arrival of the goods.

Fiscal Policies

The Silk Road's success was based on the possibility of using Bitcoin. However, Bitcoin was a highly volatile currency, and the platform came up with two options for its sellers: (i) the possibility to pegging listings to either the dollar or Bitcoin and (ii) the possibility to use "escrow hedging," namely, to reduce the losses of fiscal depreciation for payments that were placed in escrow.

Restricted Items

The Seller's Guide specifies that any items that serve to "harm or defraud, such as stolen items or info, stolen credit cards, counterfeit currency, personal info, assassinations, and weapons of any kind" were indirectly considered to be immoral and therefore not permitted.

Customer Service

Sellers are encouraged to behave in good faith; otherwise, they are warned that the platform's reputational mechanisms will not

[14]Supplementary Annex 2.

[15]https://antilop.cc/sr/users/dpr/threads/20110826-0208-Silk_Road_wiki.html

[16]https://antilop.cc/sr/users/dpr/threads/20111110-1711-Announcing_the_official_Silk_Road_Wiki.html

help their business goals. Tampering with these systems (e.g., leaving feedback for yourself as a seller from a dummy account) would be sanctioned with the revocation of privileges. The same goes for threatening customers "even if it is a veiled threat" and lying about shipping out goods. In other terms, defects of consent equivalent to misrepresentation or fraud would result in the harshest punishment, namely, killing the account.

Buyer Statistics

This is a reference to the algorithmic reputational mechanism used to determine the reliability of the buyer (Resnick and Zeckhauser, 2002; Mudambi and Schuff, 2010; Motoyama et al., 2011). The role of these systems is to remedy the trust issues arising out of the context of concluding pseudonymous transactions.

Seller Pages

An example of a mandated disclosure identified in the documents was the reputation system created by the platform. Available in different versions throughout the life span of the Silk Road, an algorithm would calculate consumer feedback, leading to a score (e.g., 100% positive feedback), which also allowed the seller to be ranked vis-à-vis the other sellers on the platform. Sellers could not opt out of this system, and as such, it would not lead to any consumer rights, because its role was to flag bad actors. Voluntary disclosures related to transaction details such as guarantees for seized orders, pricing, and shipping; office hours for incoming orders; shipping options; or the payment and escrow policy. Voluntary disclosures were made at the seller's discretion. However, interestingly, these disclosures did not only include transactional limitations (e.g., "I only accept payment through the Silk Road") but also generated rights. This is an example of a consumer right of replacement arising out of the "Seized Orders Guarantee" practices by the seller going by the name Variety Jones: "Any orders stolen by customs will be replaced and re-shipped at absolutely no charge to you. It's not your fault if the thieving bastards intercepted your order, and I believe it is my responsibility to package your order stealthily enough to make it to your door. If they steal it again, I will replace it again, once again at absolutely no charge to you. I will not stop until you get your order. I may request that you use a different mailing address for replacement orders. I fucking hate those goddam customs wankers, and want you to be confident you will receive what you pay for[17]."

Buyer's Guide

As we have seen above, the Seller's Guide includes a number of obligations the seller is bound to by agreeing to the Seller's Contract. The Buyer's Guide reproduces a lot of information from the Seller's Guide (e.g., escrow hedging and buyer statistics). However, this information does not seem to give rise to specific rights, nor are there similar obligations as for the sellers. For instance, there is no specific reference to the revocation of buyer privileges if the escrow system is not respected. From this perspective, the Buyer's Guide seems to be a collection of voluntary disclosures made by the Silk Road; yet it is not clear from the wiki whether the platform also considered the relationship with the buyer to have a contractual nature.

One aspect that is much more detailed in the Buyer's Guide rather than the Seller's Guide is the reference to the escrow. In addition to the information on the release of the payment, this part elaborates on what happens if the seller and the buyer do not agree on whether the shipment has been made. The guide specifies that should the package never arrive or arrive not in the expected condition, there might be a right—most likely given by the sellers themselves—to ask for a full or partial refund. To be eligible, buyers would have to click a "resolve" button, and that would give them access to the "resolution center," where the two parties would initially try to solve the dispute bilaterally, and "'in the rare event that an agreement can't be reached, a Silk Road admin would be right there to mediate and investigate if necessary." This demonstrates the existence of a platform-led dispute resolution body adjudicating potential issues arising out of problematic deals (Ortolani, 2016).

This section analyzed and classified the various types of legal rules that can be extracted from the contractual relationship between the Silk Road as a platform and its users, to exemplify the rights and obligations undertaken within this transaction framework. In what follows, these rules are further put into context from the perspective of identity management.

IDENTITY MANAGEMENT ON THE SILK ROAD AS A START-UP STATE

A lot has been written about the pseudonymous identity of Silk Road users (Huang, 2015; Kozinski, 2015; DiPiero, 2017; Holm, 2017). As a place where users themselves considered they broke the laws of their own states either because they would not recognize their legitimacy or simply want to shop for other legal standards, the Silk Road would make use of digital aliases as a way to hedge users from the risks incurred by engaging in commerce on the platform. However, what is less explored is the fact that in becoming users on the Silk Road, individuals would have to create accounts which were under the direct control of the Silk Road administrator. The management of these accounts would thus become an administrative task of implementing the platform's main rules and principles, albeit in a seemingly arbitrary way. As the administrator, Ulbricht had the possibility to demote sellers, to ban users from the forum, or most importantly to "kill users" (see Supplementary Annex 5). It remains unclear how exactly Ulbricht has made use of this discretion. However, what can be determined is the fact that the Silk Road operated in an organic way, namely, by dealing with issues as they came along. For instance, when Ulbricht writes the forum post regarding the creation of a Silk Road wiki, he does not mandate it to specific users but rather asks for their opinion and collaboration. When forum participants indicate they believe a wiki would not necessarily solve the questions most users would repetitively turn to the forum to (e.g., how the reputation system worked; how escrow worked; and what intermediation fee was charged by the Silk Road), Ulbricht shared

[17] https://antilop.cc/sr/vendors/

with the community his doubts in moving further with the idea: "hmmm . . . figured someone would want to do this. I guess I'll figure out how to set it up. reply to this thread if you want to be a contributor to the wiki[18]." This interaction is illustrative of the many *ad hoc* ideas and decisions that needed to be made in the course of the platform's life. Unlike a traditional legal system, systematically coupled with procedures which allow for the implementation of rights and obligations, the Silk Road was a victim of the consideration that libertarianism does not require rules. In addition, none of the admins, including Ulbricht himself, had any experience or training in governance. In such a context, the Silk Road legal concepts come across as the creations of a start-up state, with little to no systematization in rule-making, as well as with questionable consistency.

This raises two main points that are pertinent for the digital identity discussion, stemming from the same consideration, namely, that not acknowledging the importance of procedural rules affects the delivery of justice in cyberspace: first, as a private, hidden platform acting as an administrative institution keeping records of its users, the Silk Road can be an (extreme) illustration of the pitfalls of the private governance of identity; second, the Silk Road generated its own legal standards and also created the infrastructure necessary for the enforcement of these standards, which can be a relevant illustration for the code-is-law narrative coined decades ago by Lessig (2000).

As far as the first point goes, the Silk Road example seems to have an almost prescient nature as it unfolded years before the conspicuous content moderation debates which currently weigh heavily on the shoulders of social media platforms (Klonick, 2018; Langvardt, 2018; Witt et al., 2019). The Silk Road used a contractual relationship to impose various rights and obligations to its users. Yet while a cyberlibertarian orientation allegedly formed the basis of the platform's activities, the platform administrator had a literal kill switch to deal with accounts promoting undesirable activities. Such actions would arguably be based on the violations of the platform's ToS. However, as it is currently clear also in the context of social media content moderation, the content of community guidelines and its enforcement are two separate issues that do not always overlap. The discretion Ulbricht seems to have enjoyed in taking measures against platform users is a random enforcement mechanism undermining the vision and principles expressed, for instance, in the Silk Road charter (see Supplementary Annex 1). The measures the administrator would be able to take against platform users can be considered a restriction of the user identity in the given socioeconomic context. This is comparable to the ancient practices around Roman citizenship (Koops, 2012). Roman citizenship can perhaps reflect one of the clearest examples of how access to rights and privileges can influence a person's identity within a defined social group (Kunkel, 1975): during the Republic and the Principate, only Roman citizens would live according to the so-called *ius civile* (Van den Bergh, 2011). All the rights that defined *ius civile* would only apply to citizens, and their application was overseen by a *praetor urbanus*, in front of whom citizens would have recourse to actions or defenses protecting their economic interests. Non-citizens were thus invisible from the perspective of *ius civile*[19] (Hitch, 1932). The issue of access to justice brings with it the question of exclusion on the basis of identity and the inherent ways of gaming this system through identity theft. In an online environment where a pseudonymous account is the only way of interacting with the community, banning users, or killing their accounts—and with that any reputation or community standing the user may have earned—leads to the same arbitrary exclusion from the identity management system that may have warranted the creation of multiple identities or other forms of retaliation. On the one hand, the administrator believed he was administering justice when he exercised his powers. On the other hand, whatever justice was served to the community, it did not systematically apply to all its members in the same way, because while he mimicked the creation of market-based institution to protect trade, it can be argued that Ulbricht failed—or was unwilling to—to mimic the rule of law. This is the very same problem content moderation platforms currently deal with, even when they try to design specific access to justice institutions like Facebook's new appellate court for content oversight (Constine, 2019). The lack of justice fora and principles for the delivery of justice also remains one of the main problems in current blockchain governance debates. In one of his governance statements, Ethereum core developer Vlad Zamfir states that "the blockchain should be governed on a basis of global cooperation between self-selecting members and entities from the global public" (Zamfir, 2018). Still, the institutionalization of this cooperation and the functioning of the Ethereum blockchain as a public good in terms of administering justice when harms occur remain ideals that have so far not materialized.

The second point to be made in relation to identity management systems has to do with the nature of the rules on the Silk Road. The Silk Road reputation system, together with its pseudonymous user registration, was the expression of identity certification on the dark market place. In Lessig's seminal piece proposing computer codes as a regulator of cyberspace[20], he warned that for instance privacy can be coded in the identification architecture (which is the actual practice of the Silk Road) and that cyberspace would end up being regulated by cyberspace (De Filippi and Wright, 2018). This narrative is increasingly used in contemporary cryptocommunities, like those formed around various blockchains, especially in the light of self-enforcing tools such as smart contracts, where the code is law, literally[21]. Calling the code the regulator of cyberspace, however, takes away from how law really works in society: substantive rules are made to define the body of rights and obligations benefitting or imposed on legal subjects, and procedural rules determine how substantive rules are applied in practice. While the formalism attached to procedural rules is considered to be

[18] Fn 36.

[19] They would carry with them their own laws; see R. M. Hitch, 'Our Debt to Roman Law' (1932) 13 Loy LJ 66, 71.

[20] Lessig, fn 32.

[21] *Ibid.*

a trait of continental civil law[22] (La Porta et al., 2008). the same formalism can serve to enrich the code-is-law narrative by adding the perspective of the code as procedure. Rights and obligations themselves cannot be regulated through the code, while their implementation may very well be. When implementation rules are lacking, it makes the expression of rights or obligations difficult. For instance, the data protection and obligations related to payment which were embedded in the Seller's Guide have no equivalent expression in procedures that could have made the application of these obligations more transparent or systematic. It is true that "[b]y translating laws into technical rules, legal provisions are automatically enforced by the underlying technological framework[23]." However, the legal provisions automatically enforced are not the substantive standards, but inconsistent procedural rules. In comparison, consumer protection laws in Europe establish that the consumer must be protected from unfair commercial practices, and unfairness is a substantive rule that looks at the potential manipulation of the consumer through misleading or omissive practices[24]. To implement this rule in practice, additional national rules outlining specific judicial or extrajudicial procedures (e.g., injunctions or other judicial measures and access to alternative dispute resolution) needed to be drafted to guarantee the consistent application of the fairness principle. It is in vain that platforms such as the Silk Road, but similarly also Facebook and even Ethereum, draft community guidelines or governance principles, if these guidelines and principles lack a clear procedural framework which can make their application transparent and conducive to legal certainty. Absent such procedural rules (called "administrative rules" in **Figure 3** below), a control panel like the one used by Ross Ulbricht (see Supplementary Annex 5) is nothing more than the expression of randomized management, where the administrator would—when available—be tagged in forum posts bringing issues to his attention (e.g., users being disrespectful on the forum), and if he decided to take any action, this action would be mostly left at his discretion, should it fall under a category of rules which were not preset in the ToS.

[22] See for instance the debate on legal origins, Rafael La Porta, Florencio Lopez-de-Silanes, and Andrei Shleifer, 'The Economic Consequences of Legal Origins' (2008) (46)2 Journal of Economic Literature 285.

[23] De Filippi and Wright, fn 52 at 194.

[24] Directive 2005/29/EC concerning unfair business-to-consumer commercial practices [2005] OJ L149/22.

			Rule category		
	Legal rule	Consequence	Constitutional	Contractual	Administrative
Charter	Self-ownership Accountability Equality Pacta sunt servanda	-	✓		
Seller Contract	Data protection (client anonymity)	If saving or asking for personal data, then vendor account removed	✓	✓	
	Information duties & mandatory disclosures (shipping, Seller's Guide, describing items accurately)	If not abiding by these rules, then vendor account removed		✓	
	Performance in good faith	If description not 'truthful', then vendor account removed		✓	
Seller's Guide	Payment in Escrow	If payment outside Escrow or new vendors finalizing early, then vendor account removed	✓	✓	
	Restricted items (e.g. weapons, CP)	-	✓	✓	
	Prohibition of defects of consent (fraud, threat)	-		✓	
Buyer's Guide	Tampering with reputation systems Buyer statistics (reputation)	If posting fake reviews, then vendor account removed	✓	✓	✓
	Resolution center for Escrow	-	✓	✓	✓

FIGURE 3 | Overview of the most prominent legal rules in terms of service.

Blockchain governance does not raise the discretion issue to the same extent, given that accountability is supposedly left up to the consensus protocol deployed by specific blockchain networks. Yet perhaps the most important point to be made in this respect is that even in a decentralized, self-enforcing system, procedural rules are vital in determining the path of decision making. In a way, a consensus protocol is a procedure in itself. Still, the scaling of blockchain ecosystems from performing a function of currency exchange to delivering a broader category of transactions (e.g., self-sovereign identity systems) ultimately depends on how such ecosystems will deal with harms that may arise within their scope and how such harms ought to be remedied. That entails setting clear expectations regarding the balance of rights and obligations between the participants to these transactions, but also their standing in relation to the network itself.

CONCLUSION

This article looked at the Silk Road dark market as a cryptocommunity that deployed a unique identity management system. This identity was based on the roles users could perform on the platform and what their rights and obligations actually entailed. To do so, attention was paid to the essential contractual framework documents such as the Seller's and Buyer's Guides, but also the Silk Road Charter, which were all sources of rules created within the Silk Road community by its administrator, Ross Ulbricht.

These rules were further contextualized by addressing the setup of the identity management system used by Ross Ulbricht, critically analyzed from two perspectives stemming out of the lack of procedural rules to systematically enforce the private regulatory framework: arbitrariness as the main pitfall arising out of the private governance of identity systems and the code-as-procedure view complementing existing narratives surrounding the regulatory nature of the code deployed in cyberspace.

Overall, virtual worlds such as the Silk Road—especially given their use of the Bitcoin blockchain—are a source of untapped potential in exploring further questions relating to how more contemporary blockchain ecosystems can be profiled in terms of community and transactional dynamics.

Whether for social media platforms or for contemporary cryptocommunities, general procedures are vital in establishing consistent operations. Even where specific rules have not yet been developed, procedural clarifications can play a crucial role in dealing with the policy discretion that may be inherent to legal orders coexisting with the state. So far, the legitimacy of these orders has been analyzed primarily through the perspective of substantive rights and obligations and how they may conflict with state law. However, it is equally necessary to explore the notion of procedural law when dealing with the governance of cyberspace, as it may be a much needed ground for convergence between the many legal orders which have emerged between the physical and virtual realities.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://antilop.cc/sr/.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

REFERENCES

AlQahtani, A. A., and El-Alfy, E. M. (2015). Anonymous connections based on onion routing: a review and a visualization tool. *Proc. Comput. Sci.* 52, 121–128. doi: 10.1016/j.procs.2015.05.040

Augot, D., Chabanne, H., Chenevier, T., George, W., and Lambert, L. (2017). "A user-centric system for verified identities on the bitcoin blockchain," in *Data Privacy Management, Cryptocurrencies and Blockchain Technology. DPM 2017, CBT 2017. Lecture Notes in Computer Science*, eds J. Garcia-Alfaro, G. Navarro-Arribas, H. Hartenstein, and J. Herrera-Joancomartí (Cham: Springer).

Bar-Gill, O., and Porat, A. (2017). Disclosure rules in contract law, harvard law school John M. *Paper presented at the Olin Center Discussion Paper No. 907* (Chicago, IL: University of Chicago).

Bartlett, J. (2015). *The Dark Net*. Brooklyn, NY: Melville House.

Biometric Technology Today (2017). Accenture and Microsoft add blockchain tech to biometrics ID platform. *Biometric Technol. Today* 7:12. doi: 10.1016/s0969-4765(17)30141-8

Biswas, K., and Muthukkumarasamy, V. (2016). "Securing smart cities using blockchain technology," in *Proceedings of the IEEE 18th International Conference on High Performance Computing and Communications, IEEE 14th International Conference on Smart City, IEEE 2nd International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Piscataway, NJ.

Can, U., and Alatas, B. (2019). A new direction in social network analysis: online social network analysis problems and applications. *Phys. A Stat. Mech. Appl.* 535:122372. doi: 10.1016/j.physa.2019.122372

Casino, F., Dasaklis, T. K., and Patsakis, C. (2019). A systematic literature review of blockchain-based applications: current status, classification and open issues. *Telematics Inform.* 36, 55–81. doi: 10.1016/j.tele.2018.11.006

Christin, N. (2012). "Traveling the silk road: a measurement analysis of a large anonymous online marketplace," in *Proceedings of the 22nd International Conference on World Wide Web* (Pittsburgh, PA: CyLab), 213–224.

Cleaver, G. M. (2011). *The Idea of Self-Ownership*. Victoria: ORCA.

Cohen, G. A. (1995). *Self-Ownership, Freedom & Equality*. Cambridge, MA: Cambridge University Press.

Constine, J. (2019). *Facebooks New Policy Supreme Court Could Override Zuckerberg*. San Francisco, CA: Techcrunch.

De Filippi, P., and Wright, A. (2018). *Blockchain and the Law*. Cambridge, MA: Harvard University Press.

DiPiero, C. (2017). Deciphering cryptocurrency: shining a light on the deep Dark Web. *Univ. Ill. Law Rev.* 2017:1267.

Dizgovin, F. R. (2016). Foundations of specific performance in investor-state dispute settlements: is it possible and desirable. *Flor. J. Int. Law* 28, 1–62.

Dunphy, P. (2018). "A first look at identity management schemes on the blockchain," in *Proceedings of the IEEE Security and Privacy Magazine special issue on Blockchain Security and Privacy*, Piscataway, NJ.

Eisenberg, M. A. (2003). Disclosure in Contract Law. *Calif. Law Rev.* 91, 1645–1691.

Freund, A. (2017). "Automated, decentralized trust: a path to financial inclusion," in *Handbook of Blockchain, Digital Finance, and Inclusion*, eds D. Lee Kuo Chuen, and R. Deng (Cambridge, MA: Academic Press).

Gebhardt, J. H. (1947). Pacta sunt servanda. *Modern Law Rev.* 10, 159–170.

Goanta, C., and Hopman, M. (2020). Cryptocommunities as legal orders
Greenberg, A. (2012). *This Machine Kills Secrets: How Wikileakers, Cypherpunks, and Hacktivists Aim to Free the Worlds Information*. Boston, MA: Dutton.
Hitch, R. M. (1932). Our debt to roman law. *Loyola Law J.* 13, 66–92.
Holm, E. (2017). The darknet: a new passageway to identity theft. *Int. J. Inform. Sec. Cyber.* 6, 41–50. doi: 10.19107/ijisc.2017.01.04
Hou, H. (2017). "The application of blockchain technology in E-government in China," in *Proceedings of the 26th International Conference on Computer Communications and Networks* (Vancouver, BC: IEEE).
Huang, A. (2015). Reaching within silk road: the need for a new subpoena power that targets illegal bitcoin transactions. *Boston Coll. Law Rev.* 56:10.
Ibba, S., Pinna, A., Seu, M., and Pani, F. E. (2017). "CitySense: blockchain-oriented smart cities," in *Proceedings of the ACM International Conference Proceeding Series*, New York, NY.
Jaffe, C., Mata, C., and Kamvar, S. (2017). "Motivating urban cycling through a blockchain-based financial incentives system," in *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing and 2017 ACM International Symposium on Wearable Computers*, New York, NY.
Jeremy, A. (2000). Pacta sunt servanda the influence of canon law upon the development of contractual obligations. *Law Just.Christ. Law Rev.* 144, 4–17.
Klonick, K. (2018). The new governors: the people, rules, and processes governing online speech. *Harvard Law Rev.* 131:1598.
Koops, E. (2012). Second-Rate citizens: junian latins and the constitutio antoniniana. *Maastricht J. Eur. Comp. Law* 19, 223–239. doi: 10.1177/1023263x1201900202
Kozinski, A. (2015). The two faces of anonymity. *Capital Univ. Law Rev.* 43, 1–17.
Kunkel, W. (1975). *An Introduction to Roman Legal and Constitutional History*. Oxford: Clarendon Press.
La Porta, R., Lopez-de-Silanes, F., and Shleifer, A. (2008). The economic consequences of legal origins. *J. Econ. Literature* 46, 285–332. doi: 10.1257/jel.46.2.285
Langvardt, K. (2018). Regulating online content moderation. *Georget. Law J.* 106:3024739.
Lemieux, V. L. (2016). Trusting records: is blockchain technology the answer? *Rec. Manag. J.* 26, 110–139. doi: 10.1108/rmj-12-2015-0042
Lessig, L. (2000). *Code Is Law: On Liberty in Cyberspace*. Cambridge, MA: Harvard Magazine.
Locke, J. (1821). *Two Treaties of Government*. London: Whitmore and Fenn.
Marsal-Llacuna, M.-L. (2018). Future living framework: is blockchain the next enabling network? *Technol. Forecast. Soc. Chang.* 128, 226–234. doi: 10.1016/j.techfore.2017.12.005
May, T. (1992). *The Crypto Libertarian Manifesto*. Available online at: https://www.activism.net/cypherpunk/crypto-anarchy.html (accessed March 25, 2020).
Mazzacano, P. J. (2011). Force majeure, impossibility, frustration & the like: excuses for non-performance: the historical origins and development of an autonomous commercial norm in the CISG. *Nordic J. Commerc. Law* 11, 1–54.
McMillan, R. (2014). *Hacker Dreams Up Crypto Passport Using the Tech Behind Bitcoin*. San Francisco, CA: WIRED.
Mears, J. (2018). The rise and rise of ID as a service. *Biometric Technol. Today* 2018, 5–8. doi: 10.1016/s0969-4765(18)30023-7
Metjahic, L. (2018). Deconstructing the DAO: the need for legal recognition and the application of securities laws to decentralized organizations. *Cardozo Law Rev.* 39, 1533–1550.
Michaels, L., and Homer, M. (2017). "Regulation and supervision in a digital and inclusive World," in *Handbook of Blockchain, Digital Finance, and Inclusion*, eds D. Lee Kuo Chuen, and R. Deng (Cambridge, MA: Academic Press).
Motoyama, M., McCoy, D., Levchenko, K., Savage, S., and Voelker, G. (2011). "An analysis of underground forums," in *Proceedings of ACM Internet Measurement Conference*, Berlin.
Mudambi, S., and Schuff, D. (2010). What makes a helfpul online review? A study of customer reviews on Amazon.com. *MIS Q.* 34, 185–200.
Nozick, R. (1977). *Anarchy, State & Utopia*. New York, NY: Basic Books.
Ortolani, P. (2016). Self-enforcing online dispute resolution: lessons from Bitcoin. *Oxf. J. Legal Stud.* 36, 595–629. doi: 10.1093/ojls/gqv036
Reijers, W., OBrolcháin, F., and Haynes, P. (2016). Governance in blockchain technologies & social contract theories. *Ledger* 1, 134–151.
Resnick, P., and Zeckhauser, R. (2002). Trust among strangers in internet transactions: empirical analysis of eBays reputation system. *Adv. Appl. Microecon.* 11, 127–157. doi: 10.1016/s0278-0984(02)11030-3
Rivera, R., Robledo, J. G., Larios, V. M., and Avalos, J. M. (2017). "How digital identity on blockchain can contribute in a smart city environment," in *Proceedings of the International Smart Cities Conference* (Wuxi: IEEE).
Rothbard, M. (1978). *For a New Liberty: The Libertarian Manifesto*. New York, NY: Collier Books.
Sharma, P. K., Moon, S. Y., and Park, J. H. (2017). Block-VN: a distributed blockchain based vehicular network architecture in smart city. *J. Inform. Process. Syst.* 13, 184–195.
Sullivan, C., and Burger, E. (2017). E-residency and blockchain. *Comput. Law Sec. Rev.* 33, 470–481. doi: 10.1016/j.clsr.2017.03.016
Van den Bergh, R. (2011). *Communication and Publicity of the Law in Rome*. New York, NY: SUBB Jurisprudentia.
Van Parijs, P. (1995). *Real Freedom For All*. Oxford: Oxford University Press.
Vlavianos, G. (1993). Specific performance in the civil law: mediating between inconsistent principles inherited from a roman-canonical tradition via the french astreinte and the québec injunction. *Rev. Gen. Droit* 24, 469–619.
Walch, A. (2017). The path of the blockchain lexicon (and the Law). *Rev. Bank. Finance Law* 36, 713–767.
Wehberg, H. (1959). Pacta sunt servanda. *Am. J. Int. Law* 53, 775–786.
Witt, A., Suzor, N., and Huggins, A. (2019). The rule of law on instagram: an evaluation of the moderation of images depicting womens bodies. *Univ. N. S. Wales Law J.* 42, 557–596.
Yang, C., Chen, X., and Xiang, Y. (2018). Blockchain-based publicly verifiable data deletion scheme for cloud storage. *J. Netw. Comput. Appl.* 103, 185–193. doi: 10.1016/j.jnca.2017.11.011
Zamfir, V. (2018). *My Intentions for Blockchain Governance Medium*. Available online at: https://medium.com/@Vlad_Zamfir/my-intentions-for-blockchain-governance-801d19d378e5 (accessed March 25, 2020).
Zhang, N., Zhong, S., and Tian, L. (2017). Using blockchain to protect personal privacy in the scenario of Online taxi-hailing. *Int. J. Comput. Commun. Control* 12:886. doi: 10.15837/ijccc.2017.6.2886
Zyskind, G., Nathan, O., and Pentland, A. (2015). "Decentralizing privacy: using blockchain to protect personal data," in *Proceedings of the IEEE Security and Privacy Workshops* (Singapore: SPW), 180–184.

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

ORIGINAL RESEARCH
published: 28 April 2020
doi: 10.3389/fbloc.2020.00017

ID-Based User-Centric Data Usage Auditing Scheme for Distributed Environments

Nesrine Kaaniche[1], *Maryline Laurent*[2,3]* *and Claire Levallois-Barth*[3,4]

[1] Department of Computer Science, University of Sheffield, Sheffield, United Kingdom, [2] SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, Paris, France, [3] Member of the Chair Values and Policies of Personal Information, Paris, France, [4] Télécom Paris, Institut Polytechnique de Paris, Paris, France

Recent years have witnessed the trend of increasingly relying on remote and distributed infrastructures, mainly owned and managed by third parties. This increased the number of reported incidents of security breaches compromising users' personal data, where involved entities may massively collect and process massive amounts of such data. Toward these challenges, this paper combines hierarchical Identity Based Cryptographic (IBC) mechanisms with emerging blockchain technologies and introduces a blockchain-based data usage auditing architecture ensuring availability and accountability in a personal data-preserving fashion. The proposed approach relies on smart auditable contracts deployed in blockchain infrastructures. Thus, it offers transparent and controlled data access, sharing and processing, so that unauthorized entities cannot process data without data subjects' consent. Moreover, thanks to the usage of hierarchical ID-based encryption and signature schemes, the proposed solution protects and ensures the confidentiality of users' personal data shared with multiple data controllers and processors. It also provides auditing capacities with tamper-proof evidences for data usage compliance, supported by the intrinsic properties of the blockchain technology.

Keywords: blockchain, personal data protection, data usage auditing, hierarchical ID-based Cryptography, user-centric, GDPR, accountability

OPEN ACCESS

Edited by:
Oskar Josef Gstrein,
University of Groningen, Netherlands

Reviewed by:
Aneta Poniszewska-Maranda,
Lodz University of Technology, Poland
Jianwei Liu,
Beihang University, China
Qi Xia,
University of Electronic Science and Technology of China, China

***Correspondence:**
Maryline Laurent
maryline.laurent@telecom-sudparis.eu

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *25 September 2019*
Accepted: *31 March 2020*
Published: *28 April 2020*

Citation:
Kaaniche N, Laurent M and Levallois-Barth C (2020) ID-Based User-Centric Data Usage Auditing Scheme for Distributed Environments. Front. Blockchain 3:17. doi: 10.3389/fbloc.2020.00017

1. INTRODUCTION

Nowadays, organizations are collecting large amounts of personal and sensitive data about individuals. The most sensitive data are the most valuable for emerging technologies and applications, namely Artificial Intelligence (AI). Medical records, financial statements, location history, or voice transcripts may all be processed by AI algorithms to provide services that improve individuals' daily lives. This raises the question of the transparency of usage and protection of the collected personal data. Indeed, in several settings, users have little or no control over the data collected and stored about them and how they are used.

Several approaches have been introduced in order to address personal data confidentiality issues, from both legislative and technical perspectives. Indeed, strong authentication and authorization mechanisms based on centralized trusted authorities, emerged for protecting the rights and freedom of the citizens especially for ensuring the right of the protection of personal data and privacy.

In 2018, the General Data Protection Regulation (GDPR) came into force for effectively ensuring the protection of the data subject's personal data [Regulation (EU), 2016]. In particular, the regulation clarifies the conditions under which it is compulsory to obtain the consent of the data subject before processing his personal data, especially for sensitive personal data and data relating to minors. The GDPR also introduces the new obligation of accountability for organizations (i.e., data processors and data controllers). Indeed, each entity processing personal data must be able to demonstrate at any times that it is complying with the obligations laid down by the GDPR.

Recently, various accountable technical systems gained an expanding interest, such as *Bitcoin* that permits users to transfer crypto-currencies (i.e., bitcoins) securely without relying on any centralized entities, thanks to a publicly verifiable open ledger, known as *blockchain*. Thanks to their main intrinsic properties, i.e., tamper-proof infrastructure and availability, blockchain technologies are nowadays widely adopted for data accounting and auditing features.

1.1. Contributions

This paper introduces "a blockchain-based scheme for data usage auditing while preserving personal data confidentiality and ensuring continuous data availability. The proposed scheme relies on hierarchical ID-based cryptographic techniques, where a central master authority delegates the process of public/private keys' generation to the different participating entities, based on authentic public elements" (Laurent et al., 2018). ID-based Encryption (IBE) and Signature (IBS) schemes enable the data subject to encrypt sensitive data and sign transactions, relying on a unique—yet un-identifiable—identifier, respectively.

The proposed solution is multi-fold. First, relying on a blockchain infrastructure, it provides a trusted and transparent environment that permits service providers to collect tamper-proof evidence of received users' consent before gathering, storing, and/or processing their personal data. Second, the proposed framework improves the transnational consent secrecy. "That is, every data subject acts as a delegated PKG by computing an ID-based pair of keys to encrypt/sign the data that he intends to share with either a data controller or a data processor. As such, the data access is managed by the data subject. Third, by using a per smart contract ID-based key, we provide a flexible and secure sharing approach" (Laurent et al., 2018). In fact, the distribution of the decrypting keys between the data subject and the authorized data controllers and processors, does not leak the personal data of the data subject. Fourth, compared to closely related techniques, the proposed solution ensures acceptable processing overheads at both the data owner and the service provider sides.

1.2. Paper[1] Organization

Section 2 introduces the problem statement. Section 3 reviews the blockchain-based technology and discusses related work. Section 4 introduces hierarchical ID-based cryptographic techniques. Section 5 discusses design requirements and highlights security and functional concerns. Section 6 details our proposed solution. Section 7 gives a security analysis of our approach and section 8 concludes the paper.

[1]This paper is an extended and revised version of our former conference work accepted in Kaaniche and Laurent (2017b). Some excerpts of previous publications of the authors (Kaaniche and Laurent, 2017b,c; Laurent et al., 2018) are quoted in the paper from time to time.

2. PROBLEM STATEMENT

According to the GDPR, the data subject's consent is given for specific purposes that must be compliant to both the data controller and the data processor. In this context, three main roles are defined. The *data subject* who gives his consent to a *data controller* (i.e., organization, enterprize) for the processing of his personal data, with the possibility to forward them to a *data processor* (i.e., organization, enterprize) that may process data on behalf of the data controller. Here data controllers are responsible for (i) specifying to the data subject the purpose of data collection, (ii) obtaining the data subject's consent, and (iii) processing personal data according to the consented purposes, and not beyond. We note that for ease of presentation, the remainder of the paper refers to the data subject as the data owner and to both the data controller as well as the data processor as the service provider.

From a data owner perspective, there is a need for new security mechanisms that support data accountability and provenance auditing. In a nutshell, these solutions have to ensure that personal data were accessed by data controllers and/or forwarded to data processors. Indeed, it is important to conceive a secure and transparent solution that permits data owners to (i) check that data controllers and processors are correctly using their personal data with respect to the consented purposes, (ii) verify whether data were accessed, processed, or forwarded without their consent, and (iii) withdraw their consent. From a data controller or processor perspective, there is a need to design a trusted and transparent accountability solution that enables them to get a proof of the data owner's given consent prior to gathering, accessing, processing, or storing his personal data.

3. BLOCKCHAIN-BASED TECHNOLOGY

This section reviews blockchain technologies (cf. section 3.1) and discusses related work (cf. section 3.2).

3.1. Background

Blockchain has gained an attractive interest, since 2008, with the bankruptcy of *Lehman Brothers* in the United States. The root source of this economic crisis is the centralized payment system that relies on clearinghouses, acting as intermediate entities between sellers and buyers in an opaque fashion, and adding a significant extra cost to any inter-bank transactions.

"Bitcoin[2] appeared as an innovative technology enabling users to directly transfer cryptocurrencies in between with no intermediaries. It is considered as the first decentralized cryptocurrency transfer system. It relies on cryptographic proofs of work, digital signatures, and peer-to-peer networking to provide a distributed ledger containing transactions, and referred to as a *blockchain*. A blockchain is essentially a public ledger

[2]https://bitcoin.org/en/

of transactions or events recorded and stored in chronologically connected blocks (Swan, 2015; Crosby et al., 2016; Kaaniche and Laurent, 2017b)."

Two approaches, referred to as permissionless blockchains, have emerged to implement decentralized public services and applications.

"The first approach relies on the existing Bitcoin-blockchain and builds a new framework on top of it. The main advantage of this approach is that the Bitcoin blockchain already exists and is adopted by many users, which makes it more secure, transparent and resilient. The disadvantage is that blocks are mined every 10 min, and the Bitcoin scripting language is not Turing-complete (Swan, 2015).

The second approach is to build an alternative blockchain with all the desired features, which promises full decentralization, such as Ethereum[3]. Additionally to functions already supported by other public blockchain platforms such as bitcoin, e.g., mining of the digital currencies and transaction management, Ethereum also provides a contract functionality known as *smart contract*.

Transactions submitted to the Ethereum environment are organized into blocks and chained to each other based on a cryptographic hash function, initially relying on a pre-computed genesis block. Once a block is added to the blockchain, it cannot be modified or removed for two reasons: first, a block modification would lead to wrong verification of the chain of hash values, and second, the block modification would require intensive efforts to change every replicate of the blockchain supposed to be hosted on a large number of independent nodes. The verification of transactions and the addition of new blocks to the blockchain relies on the so called mining process. Indeed, *miners* have to solve a cryptographic challenge and winners are rewarded (Wood, 2014)", Laurent et al. (2018), referred to as PoW.

Recently, permissioned blockchains are gaining an expanding interest across multiple industries. "This concept appeared as a promoting solution for business applications of blockchain technology and distributed ledgers, in which participants do not necessarily have full trust on each, yet requiring some means of identification" (Kaaniche and Laurent, 2017b).

"Unlike permissionless blockchains, there exists a central entity that decides and grants the right to individual peers to participate in the read/write operations. The Hyperledger[4] project is a prominent initiative dedicated to bringing blockchain technologies to business. It provides a modular consensus protocol, such as Byzantine Fault Tolerance (BFT) algorithm, that ensures efficient scalability and performance complexities with thousands of transactions per second (Kaaniche and Laurent, 2017b; Vukolic, 2017)." Hyperledger development and support are today ensured by a large consortium of world-leading companies. Among prominent instances of permissionned blockchains, we can mention IBM Hyperledger Fabric[5] and R3 Corda[6].

[3]https://www.ethereum.org/
[4]https://www.hyperledger.org/
[5]https://www.hyperledger.org/projects/fabric
[6]https://www.corda.net/

3.2. Related Work

Several solutions have been proposed aiming at empowering data owners while giving them -complete- control over their collected and processed personal data. These solutions mainly investigate the mechanisms that help setting-up an informed consent between the data owner and data controllers/processors, for the different provided applications/services (Laurent, 2019; Lee et al., 2019; Morel et al., 2019; Rantos et al., 2019).

Morel et al. (2019) proposed a framework for gathering data owners consents in IoT environments. In their design, the authors assume that each device collecting data has to send an attestation about the collected and generated information. This information includes the device location, the data types, and privacy policies. The data owner receives the information via a smartphone and gives consent if accepts to share his data. Rantos et al. (2019) introduced a cloud-based platform, called ADVOCATE, which allows users to easily access their personal data and manage consents. The gathered data, from IoT devices, are searchable and can only accessed by authorized entities. Data controllers and processors should submit a query when they want to access the data generated by a specific equipment. And, the query has to clearly specify the purpose of access, the purpose of processing, the period of storage, etc. Once a query is negotiated and accepted by the data owner, the consent has to be signed by both the data owner and the data controller, and stored for auditing purposes. The proposed solution suggests the deployment of a blockchain-based infrastructure to ensure that no modification will be made on given consents. In fact, the main intrinsic property of the blockchain is its suitability for data auditing purposes. It has attracted interest of the research community due to its shared and fault-tolerance database (Zyskind et al., 2015; Linn and Koo, 2016; Ouaddah et al., 2016; Fu et al., 2017; Liang et al., 2017; Shetty et al., 2017; Kaaniche and Laurent, 2018; Ramachandran and Kantarcioglu, 2018; Zhang et al., 2018).

"Liang et al. (2017) proposed a blockchain-based data provenance architecture for cloud applications. Their construction records operations' history as provenance data which are hashed into Merkle tree nodes. A list of hashes of provenance data forms a Merkle tree which are attached as a blockchain transaction. As such, it is possible to immutably prove the provenance of data exchanges. Although the proposed approach is novel, it does not cover the definition of advanced policies or contracts regulating the usage of collected data" (Kaaniche and Laurent, 2017b).

Zyskind et al. (2015) presented an hybrid solution for personal-data management. The proposed system combines both blockchain, as an access moderator, and off-blockchain, for data storage purposes. Proposal Zyskind et al. (2015) considers clients as unique owners of their personal data, thus making them aware of all the associated data being collected by service providers and how they are used. That is, when a client subscribes to a service provider, one transaction is created defining the set of access policies and another contains the hashes of data stored in an off-chain database. However, the Zyskind et al. (2015) proposal permits to only define simple permit/deny access policies through white/blacklisting.

Based on Zyskind et al. (2015), "Linn et al. propose an application of the data auditing framework for health scenarios (Linn and Koo, 2016). In their construction, the blockchain is also considered as an access moderator to control the access to outsourced shared data" (Laurent et al., 2018). Fu et al. (2017) introduced a privacy-aware blockchain-based auditing system for shared data in cloud applications. In order to mitigate the power abuse of single tracing authorities, Fu et al. (2017) presents "a threshold approach, where at least t entities have to collaborate to recover the identity of a malicious user, ensuring thus the non-frameability of users. Based on a blockchain architecture, the proposed construction enables group users to trace data changes and recover latest correct data blocks when current data are damaged" (Kaaniche and Laurent, 2017b). Later, "Neisse et al. discussed the design requirements of blockchain-based solutions for data provenance tracking (Neisse et al., 2017), namely client-centric, server-centric, and data-centric approaches. The authors also presented an evaluation of their implementation results, in order to give a comprehensive overview of different defined approaches" (Laurent et al., 2018).

Dorri et al. (2017) introduced an architecture for data access control management in IoT environments, i.e., smart home application. The architecture relies on a central–unique–miner, i.e., local home miner, to mine blocks, and implement the access policy. The proposed solution ensures the confidentiality of data through a predefined policy, however the introduction of a central miner raises the risk of a single point of failure.

Shetty et al. (2017) proposed a blockchain-based auditing framework, called BlockCloud. The proposed solution supports "monitoring user activities in real-time using hooks and listeners which are special classes of event listeners so that every user operation on files will be collected and recorded for generating provenance data. Each piece of provenance information is referred to as transactions that are broadcasted to the core of the blockchain network created by a specific set of validating virtual machines" (Tosh et al., 2018). "The provenance auditor validates provenance data by retrieving transactions from the blockchain network by using blockchain receipt which contains block and transaction information" (Shetty et al., 2017).

Later, Ramachandran et al. introduced SmartProvenance (Ramachandran and Kantarcioglu, 2018), a blockchain-based auditing solution which applies Ethereum smart contracts to support the data provenance feature. Indeed, their proposed solution is based on two smart contracts, namely Document Tracker contract and Vote contract. The Document Tracker smart contract is used to keep track of changes to a given document while the Vote contract implements the voting protocol. That is, "every provenance change event has to be approved through a voting process by the vote contract. The data trails are only logged if they are approved by the Vote contract. The Vote contract implements two types of voting: simple majority voting and threshold voting" (Ramachandran and Kantarcioglu, 2018) where all blockchain network's minors are called to participate in the voting process. SmartProvenance implements a JavaScript client, run on the browser of each of the user's machines. The client, acting as an interface between the user and back-end smart contracts, is in charge of the synchronization with the smart contract, including the storage, the retrieval, and the validity checking of the changes.

Zhang et al. (2018) introduced an auditing scheme for cloud storage environments, based on a blockchain infrastructure. The proposed solution aims at counteracting provenance records' forgery, removal, and modification attacks. That is, each provenance record is anchored into a blockchain-transaction, and all the provenance records associated with a data item constitute a record chain. Thus, if one record entry is corrupted, the chain is broken.

Table 1 provides a comparison between the proposed scheme and most-closely related solutions, in terms of functional and security features and incurred computation overheads at the data owners and processors' sides. The Blockchain technology indicates whether the proposed schemes is relying on a permissionned or permission-less blockchain. The mining process indicates the consensus algorithm. The confidentiality, unlinkability, and anonymity properties (detailed in section 5) refer to the secrecy of personal data, unlinkability between different transactions and anonymity of the data owner, respectively. The performances at the data owner and service providers sides indicate the processing overheads.

4. IDENTITY BASED CRYPTOGRAPHY

Introduced by Shamir (1985), Identity Based Cryptography (IBC) enables the derivation of public and private keys with no need for a certification authority. Indeed, each entity should use one of its—unique—identifiers as its public key. Rather than relying on a Public Key Infrastructure (PKI), IBC assigns the private key generation function to a third party, referred to as a Public Key Generator (PKG). Thus at the request of an entity, the PKG computes the private key associated with his public key.

In order to be able to derive an entity's private key $sk_{\mathcal{E}}$, the PKG must first define a set of *ID-based public elements* (IBC–PE). The PKG generates the groups $\mathbb{G}_1$, $\mathbb{G}_2$, and $\mathbb{G}_T$ and the pairing function $\hat{e}$ from $\mathbb{G}_1 \times \mathbb{G}_2$ in $\mathbb{G}_T$. $\mathbb{G}_1$ and $\mathbb{G}_2$ are additive subgroups of the group of points of an Elliptic Curve (EC). However, $\mathbb{G}_T$ is a multiplicative subgroup of a finite field. $\mathbb{G}_1$, $\mathbb{G}_2$, and $\mathbb{G}_T$ have the same order q. In addition, $\mathbb{G}_1$, $\mathbb{G}_2$, and $\mathbb{G}_T$ are generated by P, Q and the generator $g = \hat{e}(P, Q)$, respectively.

The PKG defines the groups and a set of hash functions in accordance to the selected ID-based encryption and signature schemes. That is, the PKG uses a hash function $Hash_{pub}()$ to transform the entity's identity ($ID_{\mathcal{E}}$) into a public key as follows:

$$pk_{\mathcal{E}} = Hash_{pub}(ID_{\mathcal{E}})$$

Generally, the public key of the entity is computed as a hash of one of his identities.

The PKG generates an entity's private key based on a local secret $s_{PKG} \in \mathbb{Z}_q^*$ and a private key generation function keygen(). The private key is computed as:

$$sk_{\mathcal{E}} = \mathsf{keygen}(s_{PKG}, pk_{\mathcal{E}})$$

The groups $\mathbb{G}_1$ and $\mathbb{G}_2$, the pairing $\hat{e}$, the points P, Q, and $Q_{pub} = s_{PKG}.Q$, and the hash functions form the ID-based public

TABLE 1 | Comparison between the proposed solution and most closely related approaches.

Scheme	Security and functional properties					Performances	
	Blockchain technology	Mining process	Confidentially	Unlinkability	Anonymity	Owner	Service provider
Zyskind et al. (2015)	Public BC	–	✓	✗	✓	$n_P n_S \gamma_{SC}$	$n_U n_S \gamma_{SC}$
Neisse et al. (2017)	Ethereum	POS	✓	✗	✓	$n_P n_S \gamma_{SC}$	$n_U n_S \gamma_{SC}$
Dorri et al. (2017)	Private BC	collaborative	✗	✗	✓	–	–
Kaaniche and Laurent (2018)	Ethereum	POS	✓	✓	✓	$\gamma_{SC} + (3\gamma_E + \gamma_S)$	$\gamma_{SC} + N(6\gamma_E + 3\gamma_S)$
Shetty et al. (2017)	Public BC	– Cloud pricing	✓	✗	✓	–	–
Ramachandran and Kantarcioglu (2018)	Ethereum	POS	✗	✗	✓	$2\ n_P n_S \gamma_{SC}$	$n_U n_S \gamma_{SC}$
Zhang et al. (2018)	Ethereum	POS	✓	✗	✓	$n_P n_S(\gamma_{SC} + N(2\gamma_E + 2\gamma_S))$	$n_U n_S(\gamma_{SC} + N(2\gamma_E + \gamma_S))$
Our solution	Consortium BC HyperLedger Fabric	PBFT	✓	✓	✓	$n_P n_S(\gamma_{SC} + N(\gamma_E + \gamma_S))$	$n_U n_S(\gamma_{SC} + N(\gamma_E + \gamma_S))$

✓and ✗ indicate whether the property is supported or not, respectively. –Indicates that the property is not-applicable or non specified. n_P, n_O, and n_S indicate the number of service providers, owners, and services/applications, respectively. γ_{SC}, γ_E, and γ_S indicate the processing overhead of a smart contract creation, data encryption/decryption, and data signature/verification, respectively. N is the number of transactions associated with a smart contract.

elements as follows:

$$IBC - PE = \{\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, q, \hat{e}, g, P, Q, Q_{pub},$$

$$Hash_{pub}(), H_1(), ..., H_k()\}.$$

ID-based cryptography suffers naturally from the key escrow attack as a PKG knows the owners' private keys and, if untrusted, can realize masquerading attacks. To mitigate the key escrow attack, a hierarchy of PKGs has been proposed in Gentry and Silverberg (2002) and Horwitz and Lynn (2002), along with the Hierarchical Identity Based Cryptography (HIBC) to reduce the workload of the PKG by delegating the private key generation task to lower level entities, i.e., entities who have already obtained their private keys. As such, Hierarchical Identity based Encryption (HIBE) (Horwitz and Lynn, 2002; Gentry and Halevi, 2009; Blazy et al., 2014; Seo and Emura, 2015) and Hierarchical Identity Based Signature (HIBS) (Gentry and Silverberg, 2002; Chow et al., 2004; Tian and Huang, 2014) schemes emerged, as a generalization of IBE and IBS, respectively, to mirror an organizational hierarchy. That is, an identity at level k of the hierarchy tree can issue private keys to its descendant identities, but it cannot decrypt messages intended for other identities.

HIBE schemes rely on four randomized algorithms, namely: `setup`, `keygen`, `encrypt`, and `decrypt`. The `setup` algorithm generates system parameters, denoted by IBC–PE and the master secret key s_{PKG}. Note that the master key corresponds to the private key at depth 0 and note that an IBE system is a HIBE where all identities are at depth 1. The `keygen` algorithm takes as input an identity $ID = \{ID_1, \cdots, ID_k\}$ of depth k and the secret key $sk_{\mathcal{E}|k-1}$ of the parent entity $\{ID_1, \cdots, ID_{k-1}\}$ of depth $k-1$. It outputs the secret key $sk_{\mathcal{E}|k}$. Similarly to IBE, the encrypt algorithm encrypts messages for an identity using IBC-PE and the decrypt algorithm decrypts ciphertexts using the private key.

5. TECHNICAL DESIGN REQUIREMENTS

In this section, we first review different security and functional requirements (cf. section 5.1). Then, we discuss blockchain design directions (cf. section 5.2) and deployment models (cf. section 5.3).

5.1. Security and Functional Requirements

Our blockchain-based data usage auditing solution has to consider a set of security and functional properties, defined as follows:

- confidentiality—the proposed approach has to "prevent unauthorized disclosure of both personal identifying information and shared data between data owners and service providers" (Kaaniche and Laurent, 2017b).
- unlinkability—the proposed solution has to guarantee that created smart contracts by the same data owner can not be linked by public verifiers as well as service providers.
- auditability—the proposed solution has "to enable auditing authorities to lead an investigation and obtain consistent proofs in case of non-compliant activities" (Kaaniche and Laurent, 2017b). The auditing process, carried by authorized authorities, may be both public and private.
- censorship resistance—this property defines the ability to prevent a third-party from pushing false information across the system (Perng et al., 2005). In fact, service providers are interested by data owners' contract creation and joining, thus allowing them to manage and process data. Otherwise, data owners have the right to revoke or modify their initial consent

agreement. As such, service providers will no longer have access to data, with respect to their granted privileges.
- data transparency—each data owner should have a complete transparent view over how data are collected, accessed, and processed.
- computation overhead—the proposed solution should offer acceptable computation costs.

5.2. Blockchain Design Directions and Discussions

"To efficiently propose the contract design for each participating entity, we have to take into consideration (i) the contract lifecycle, (ii) accountability granularity, (iii) required information to be stored in the contract, (iv) access patterns w.r.t. read/modify granted privileges, and (v) blockchain architecture w.r.t. different models of deployments" (Kaaniche and Laurent, 2017b).

For this purpose, as discussed in Kaaniche and Laurent (2017c), we distinguish "three different cases for designing contracts that depend on the number of involved data owners and processors, as follows:

- M_1—*one data owner contract per specific data:* each data owner has to create a contract for each data instance for all data controllers, while specifying access privileges. In a nutshell, the contract should contain the list of authorized data controllers as well as their respective granted privileges.
- M_2—*one data owner contract for each specific data controller/processor:* the data owner creates a contract for each specific controller processing his personal data. Thus, the contract includes shared data, access policy w.r.t. data usage as well as data usage logs that represent the different events carried out by the data controller on the owner's personal data. This contract design model holds the largest number of contracts and is merely adequate for high security levels needed, for instance, for e-health data.
- M_3—*one data controller contract for multiple data owners w.r.t. data usage policy:* for this sub-case, the data controller creates a generic contract that specifies the manner of processing data received from data owners. When a data owner joins the contract established by the data controller, he then accepts the policy usage identified by the controller. This model is of interest when expecting a lower number of different owners."

"Let us note that the first model (M_1) cannot ensure the unlinkability property because the data owner is known to all organizations under the same identity. Other cases (i.e., M_2 and M_3) allow data owners to protect their identities from linkability attacks, thanks to the use of different identifiers (i.e., blockchain addresses). For instance, the first contract design model (M_1) is more scalable, compared to other design models (M_2) and (M_3), thanks to the use of a generic instance for data.

It is worth noticing that the aforementioned models are different in terms of contracts' number and data auditing granularity allowed to data owners, depending on the required security properties" (Kaaniche and Laurent, 2017c).

5.3. Public vs. Semi-public vs. Private Blockchain

As stated above and as discussed in Kaaniche and Laurent (2017c), "it is important to consider the blockchain architecture (public, semi-public, private), as it impacts several security requirements, mainly public verifiability, unlinkability, and censorship resistance properties (Swan, 2015).

For instance, a public blockchain architecture, such as Bitcoin and Ethereum, allows every entity to read, send transactions and participate in the mining process to decide which blocks are added to the chain and determine the current blockchain state. As such, these public distributed ledgers permit to have a fully transparent system, but with a questionable level of data owners' linkability. In addition, transactions' fees have to be paid for the processing of contract invocations, which could make some approaches unfeasible due to the high number of transactions, like for the data owner-centric models (M_1, M_2). This would make necessary that data owners and service providers agree on an adequate business model to avoid owners supporting all the transaction' fees.

In a private blockchain architecture, read access may be public, or restricted, while write privileges have to be granted by a central entity. In fact, the central entity is responsible for verifying transactions' compliance and censorship resistance with respect to data owners' policies. Hence, a private system provides limited transparency, but with significant unlinkability and personal data protection guarantees. Nevertheless, in a private blockchain, processing fees may be significantly reduced, mainly when considering the service provider-centric model M_3.

In a semi-public or consortium blockchain infrastructure, read access may be public or restricted and the consensus process is controlled by a set of organizations. Thus, a consortium blockchain architecture inherits several benefits from a private blockchain technology—efficiency, unlinkability, and personal data protection guarantees—without relying on one single deciding authority. Additionally, these permissioned blockchains often expose some low-level trust assumptions originating from their consensus mechanism to their smart contract applications, hence ensuring transparency and auditability. Apart from these advantages, a consortium blockchain permits to get reduced transactions' fees thanks to the consensus algorithm, which could be considered as a motivating argument for data-owner centric models, namely M_2."

6. A NEW BLOCKCHAIN-BASED DATA USAGE AUDITING ARCHITECTURE

Our architecture considers that one smart contract per data controller/processor is managed by the data owner (M_2 model of section 5.2), and relies on a consortium-blockchain infrastructure, as discussed in section 5.3.

6.1. Overview

Our blockchain-based data usage auditing solution involves the three following entities: a data owner ($\mathcal{O}$), a service provider ($\mathcal{S}$)

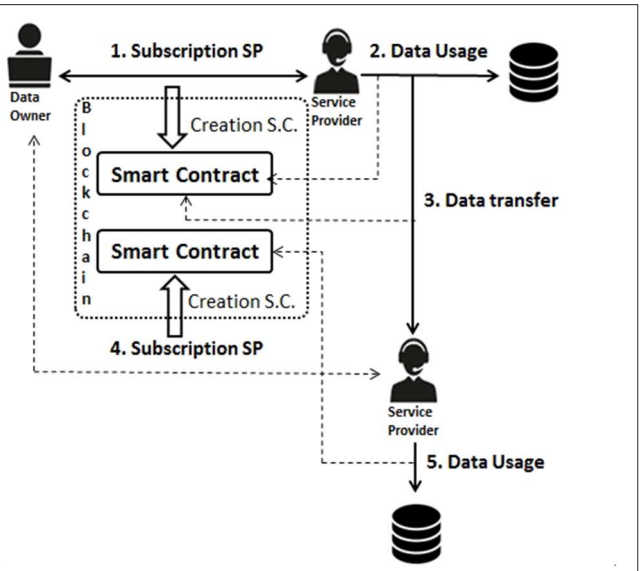


FIGURE 1 | Blockchain-based architecture for data usage auditing schemes.

(acting as a data controller), and a service provider ($\mathcal{P}$) (acting as a data processor) [Regulation (EU), 2016].

Figure 1 depicts the different actors and their interactions. When subscribing to a service provider $\mathcal{S}$, the date owner $\mathcal{O}$ creates first a data usage contract. In this contract, $\mathcal{O}$ provides "the usage policy for both data transferred by $\mathcal{O}$ to $\mathcal{S}$, as well as any other collected data by $\mathcal{S}$ during the interactions with $\mathcal{O}$ (e.g., logging files). When processing $\mathcal{O}$'s data, $\mathcal{S}$ has to comply to the granted usage policy, as registered in the smart contract in the blockchain. In addition, $\mathcal{O}$ needs to identify the list of possible actions that might be performed on behalf of the smart contract, including the smart contract approval (cf. section 6.4).

If authorized by the data usage policy in the smart contract, $\mathcal{C}$ might transmit $\mathcal{O}$'s data to service provider $\mathcal{P}$, by pushing that forwarding information to the blockchain. $\mathcal{O}$ being notified by the forward event, then has to create a new smart contract with $\mathcal{P}$" (Laurent et al., 2018).

As our approach leans on a consortium-blockchain infrastructure, for anyone to be able to read the contract as well as associated transactions, we introduce cryptographic mechanisms—hierarchical ID-based encryption (HIBE) and signature (HIBS) schemes—in order to provide secrecy of exchanged data and preserve data owner's privacy.

Our approach based on hierarchical IBC has several advantages. First, each data owner $\mathcal{O}$, as part of the blockchain, acts as a delegated PKG, thus being enabled to derive an ID-based pair of keys for encrypting and signing all the data intended to be shared with $\mathcal{S}$ or $\mathcal{P}$. It is noteworthy that the data owner $\mathcal{O}$ remains fully in charge of his data access.

Second, using hierarchical identity based schemes introduces the existence of a root PKG entity (i.e., a trusted central entity), which has the responsibility to generate certified public system parameters IBC–PE, for the entities of the system ($\mathcal{O}$, $\mathcal{P}$, $\mathcal{S}$) and to guarantee authentic entities identities within the system. Furthermore, hierarchical identity has the advantage over the client-PKG classical approach—to provide public system parameters common to any entities and not specific to the public identity of the data owner $\mathcal{O}$. "This feature enables $\mathcal{O}$ to get rid of the cumbersome tasks of generating and publishing each $\mathcal{O}$'s own public parameters, and it provides indistinguishability among entities, mostly during the signature verification processes" (Laurent et al., 2018).

Third, having a per smart-contract ID-based key provides our solution with several properties of interest, including indistinguishability among smart contracts at the blockchain level, and unlinkability of smart contracts to data owners. Both properties prevent any reidentification of data owners thanks to search operations over the blockchain.

6.2. Security and Design Assumptions

This section first details the security assumptions in section 6.2.1. Then, it introduces the smart contract design assumptions 6.2.2.

6.2.1. Security Assumptions

Our scheme considers the following security assumptions:

- an off-blockchain secure communication—a secure channel is established between $\mathcal{O}$ and the service provider $\mathcal{P}$ or $\mathcal{S}$, thus enabling $\mathcal{O}$ to securely transmit its personal information along with some data that he expects to share with the service provider. The secure channel supports mutual authentication and data confidentiality and integrity. It can be implemented through the Transport Layer protocol (TLS) (Dierks and Rescorla, 2008), where the data owner can authenticate with a certificate or password. If higher data preservation is requested, then attribute based credential mechanisms, such as Idemix[7] or UProve[8], are good candidates for enabling mutual authentication while revealing only required information to any service provider.
- a robust blockchain and smart contract implementation—any blockchain related operations, including mining and transaction anchoring activities are assumed to be secure and incorruptible, and blockchain to deliver non-tamper proofs of data processing and managing events.
- a trusted PKG entity—the PKG plays a central role, by enrolling the different entities ($\mathcal{O}, \mathcal{S}, \mathcal{P}$) into the system, by deriving their private keys, and by issuing and publishing the security ID-based public elements for the whole ID-based system. In the sequel, in order to strengthen security guarantees and prevent attacks against this central entity, the root PKG functions w.r.t. the derivation of private keys may be distributed among several entities, based on secure multi-party computation mechanisms or threshold techniques (Yu et al., 2010; Prabhakaran and Sahai, 2013; Grumăzescu et al., 2015). Note that, in practice, entities should agree on the hierarchical ID-based encryption and signature schemes which will be used for ciphering and signing messages, respectively. Our proposal does not rely on a specific scheme, however, that choice has a direct impact on the method for generating private keys.

[7]https://www.zurich.ibm.com/identity_mixer/
[8]https://www.microsoft.com/en-us/research/project/u-prove/

6.2.2. Smart Contract Design Assumptions

"For our solution, each data contract involves two parts: *data structure*, referred to as $\mathtt{p}_1$ and *data usage policy*, denoted by $\mathtt{p}_2$. The *data structure* part, includes types and instances of exchanged data between the data owner $\mathcal{O}$ and the service provider $\mathcal{S}$" (Laurent et al., 2018). That is, the data type field contains a list of all data instantiations (i.e., string, integer, date, ...), while data instances provide data values. The *data usage policy* part identifies all the authorized and denied usage actions for both transferred and implicitly collected data. Among data usage actions, we may mention data storage duration granted to the service provider, type of processing permitted over the owner's data, permission to forward data to other service providers or to generate derived data from some exchanged personal data [Regulation (EU), 2016].

Furthermore, for each created smart contract, the data owner $\mathcal{O}$ is required to list the actions that are authorized by the smart contract from $\mathcal{O}$ and service providers ($\mathcal{S}$ and $\mathcal{P}$). Permitted actions from $\mathcal{O}$ include deletion of the smart contract, inactivation of the contract and modification of new data usage policies, and permitted actions from service providers include only approval of the smart contract content.

6.3. System Initialization

We assume that a trusted root PKG entity is in charge of generating and publishing the ID-based public elements, along with the ID-based Signature and Encryption scheme (cf. section 6.2.1), as follows:

$$IBC - PE = \{\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_T, q, \hat{e}, g, P, Q, Q_{pub}, Hash_{pub}(), H_1(), ..., H_k(), IBS_{scheme}, IBE_{scheme}\}$$

Moreover, each entity $\mathcal{E}$ owns a pair of public and private keys $(sk_{\mathcal{E}}, pk_{\mathcal{E}})$, where $\mathcal{E} \in \{\mathcal{O}, \mathcal{S}, \mathcal{P}\}$. Each pair of public and private keys $(sk_{\mathcal{E}}, pk_{\mathcal{E}})$ is mathematically generated by the root PKG in accordance with the selected ID-based cryptographic schemes, as detailed in section 4. Then the root PKG securely transmits the private key $sk_{\mathcal{E}}$ through the secure channel (cf. section 6.2) to the entity $\mathcal{E}$. As such, $\mathcal{E}$ owns his secret $sk_{\mathcal{E}}$ associated to his public key $pk_{\mathcal{E}}$ where $sk_{\mathcal{E}}$ is derived from the unique entity identity of $\mathcal{E}$ (i.e., one of $\mathcal{E}$'s blockchain addresses), based on the PKG's own secret and the public elements IBC–PE. "As such, the pair $(pk_{\mathcal{E}}, sk_{\mathcal{E}})$ is somewhat certified by the root PKG, as the signature generation and data decryption operations require an entity to be equipped with his own secret $sk_{\mathcal{E}}$" (Laurent et al., 2018).

6.4. Smart Contract Creation

Figure 2 gives the full picture of the smart contract creation from the data usage request by $\mathcal{S}$ to $\mathcal{O}$ up-to the approval of the contract by $\mathcal{S}$ and transmission of data to $\mathcal{S}$.

First, $\mathcal{O}$ establishes a direct secure channel with $\mathcal{S}$ where $\mathcal{S}$ and $\mathcal{O}$ authenticate each other thanks to their respective ($sk_{\mathcal{E}}$ and $pk_{\mathcal{E}}$) pair of keys (cf. section 6.3). The transaction is identified with an T_{id} parameter. Through this channel, $\mathcal{S}$ might initiate a data usage request to $\mathcal{O}$, for $\mathcal{O}$ to infer some parameters for the smart contract $\mathcal{C}$ prior to launching the smart contract into the blockchain. The generation of these parameters is presented in section 6.4.1, and details of smart contract creation are given in section 6.4.2.

6.4.1. Generation of Per Smart-Contract Parameters by $\mathcal{O}$

After the initialization phase, as described in section 6.3, $\mathcal{O}$ owns his public and private keys $(sk_{\mathcal{O}}, pk_{\mathcal{O}})$, thus making $\mathcal{O}$ able to derive a per smart-contract $ID_{\mathcal{C}}$, with its associated pair of public and private keys $(sk_{\mathcal{C}}, pk_{\mathcal{C}})$, as follows.

Based on the public elements IBC–PE, and the hierarchical ID-based principle, as detailed in section 4, $\mathcal{O}$ first generates an identifier $ID_{\mathcal{C}}$, that constitutes the hierarchical smart contract identity of depth 2 which relies on both $\mathcal{O}$'s identity $ID_{\mathcal{O}}$ and the smart contract identifier $ID_{\mathcal{C}}$.

Then, thanks to `keygen`, $\mathcal{O}$ derives a per smart contract ID-based public key $pk_{\mathcal{C}}$ based on the general public elements IBC–PE and the concatenation of the data owner's address $ID_{\mathcal{O}}$, the service provider address $ID_{\mathcal{S}}$ and the smart contract identifier $ID_{\mathcal{C}}$.

$$pk_{\mathcal{C}} = Hash_{pub}(ID_{\mathcal{S}}||ID_{\mathcal{C}})$$

Note that the secret key $sk_{\mathcal{C}}$ associated with the public key $pk_{\mathcal{C}}$ is only known by the data owner, at the moment of the smart contract creation.

6.4.2. Creation of the Smart Contract by $\mathcal{O}$

$\mathcal{O}$ approves first the data usage request and then constructs the data section part of the contract by distinguishing parts $\mathtt{p}_1$ and $\mathtt{p}_2$, as presented in section 6.2.2. To avoid publishing personal information, our approach proposes to obfuscate data values in $\mathtt{p}_1$ thanks to the use of the public hash function $\mathcal{H}$ (e.g., SHA-256). To counteract the linkability attacks, we propose to concatenate data values with a symmetric key $k_{\mathcal{C},\mathcal{S}}$ derived from the generated smart contract secret key $sk_{\mathcal{C}}$ before application of the hash function. The derivation of $k_{\mathcal{C},\mathcal{S}}$ has to be unidirectional, through a hashing function for instance. "Then, $\mathcal{O}$ generates a smart contract creation transaction, which will be anchored in the blockchain. As such, we define a smart contract creation transaction T as the tuple $T = [T_{id}, ID_{\mathcal{C}}, U_{\mathcal{C}}, \mathtt{p}_1, \mathtt{p}_2, enc_{pk_{\mathcal{S}}}(k_{\mathcal{C},\mathcal{S}}), \sigma_{\mathcal{C}}]$, where T_{id} is the transaction identifier, $ID_{\mathcal{C}}$ is the smart contract identifier, $U_{\mathcal{C}}$ is a smart contract approval request identifier toward $\mathcal{S}$, $enc_{pk_{\mathcal{S}}}(k_{\mathcal{C},\mathcal{S}})$ is the encrypted symmetric key $k_{\mathcal{C},\mathcal{S}}$, based on the public key of the service provider $\mathcal{S}$ and $\sigma_{\mathcal{C}}$ is the signature of $\mathcal{O}$ of all the previous fields, using the private key $sk_{\mathcal{C}}$, only known to the data owner" (Laurent et al., 2018).

After being notified of the smart contract creation request, the service provider decrypts the enciphered (pairwise) symmetric key $k_{\mathcal{C},\mathcal{S}}$ using the public system parameters IBC–PE, and his own private key $sk_{\mathcal{S}}$. Then, $\mathcal{S}$ verifies the signature of the issuing data owner using the public key $pk_{\mathcal{C}}$. Only if both operations are successful, $\mathcal{S}$ can approve the transaction associated with the smart contract creation thanks to its own key $sk_{\mathcal{S}}$.

Upon receiving the smart contract approval notification, $\mathcal{O}$ transmits his personal data to the service provider, through a secure channel. The compliance between his personal data and the hashed values embedded in $\mathtt{p}_1$ proves his consent agreement.

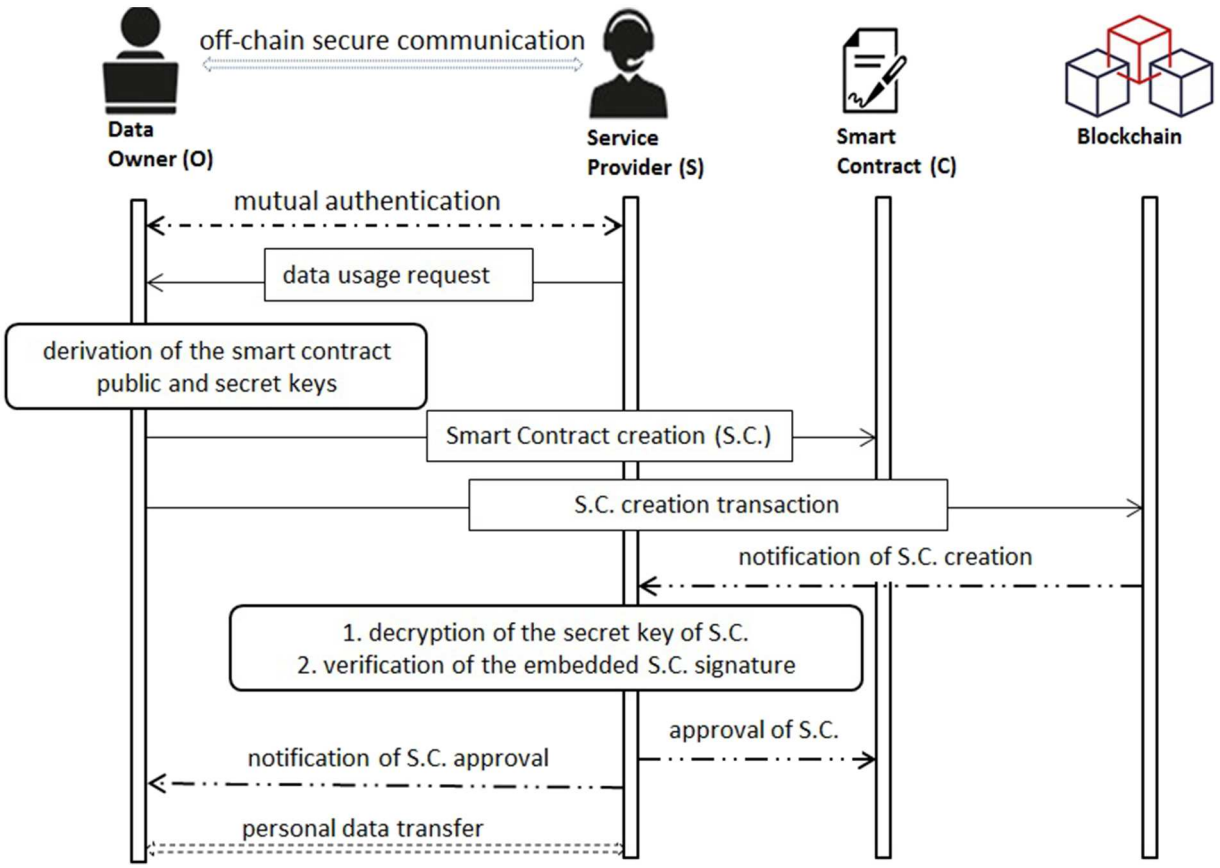


FIGURE 2 | Smart contract creation process.

Then, $\mathcal{S}$ can check the integrity of the received $\mathcal{O}$'s data based on the deciphered pairwise key $k_{\mathcal{C},\mathcal{S}}$ and the received data values.

6.5. Data Usage Transfer

"In case a forward action is authorized on some $\mathcal{O}$'s data by the data usage policy specified in the smart contract with $\mathcal{S}$, then data might be transferred by the service provider $\mathcal{S}$ to the corresponding data processor $\mathcal{P}$ and a transaction has to be pushed to the blockchain" (Laurent et al., 2018).

"For this purpose, $\mathcal{S}$ has to generate a data usage transaction. As such, the data owner is notified about the service provider data transfer activity. We define a data usage transaction T corresponding to a data transfer activity as the tuple $T = [T_{id}, ID_{\mathcal{C}}, ID_{\mathcal{S}}, U_{\mathcal{C}}, enc_{pk_{\mathcal{O}}}(ID_{\mathcal{P}}), \sigma_{\mathcal{S}}]$, where T_{id} is the transaction identifier, $ID_{\mathcal{C}}$ is the smart contract identifier, $ID_{\mathcal{S}}$ is the service provider identity, $U_{\mathcal{C}}$ is a transfer data event identifier, $enc_{pk_{\mathcal{O}}}(ID_{\mathcal{P}})$ is the encrypted identity of $\mathcal{P}$, based on the public key of the data owner $\mathcal{O}$ and $\sigma_{\mathcal{S}}$ is the signature of $\mathcal{S}$ over all the previous fields" (Laurent et al., 2018).

The data owner $\mathcal{O}$ being notified of the data transfer creates a new smart contract with data processor $\mathcal{P}$, following the procedure presented in section 6.4. Note that $\mathcal{P}$ does not know the identifier of the newly agreed contract between $\mathcal{O}$ and $\mathcal{S}$.

6.6. Other Operations Over the Smart Contract

Additionally to the approval of the smart contract by $\mathcal{S}$ or $\mathcal{P}$, the smart contract has to include the following actions:

- deleting the contract from the blockchain by $\mathcal{O}$—In case $\mathcal{O}$ wants to withdraw his data usage consent, the contract must be made inactive, thus leading to denying any subsequent data usages. The only possible remaining activity is the registration of the complete contract history in the blockchain.
- changing data usage—At any moment, $\mathcal{O}$ is able to restrict or expand his data usages.

Any contract policies modifications must be taken into consideration in subsequent data usages. This is made possible thanks to the blockchain notifying any changes w.r.t. the smart contract, as well as data usage policies embedded into p_2, to the involved entities.

6.7. Auditing

The auditing process can be realized either in a private manner by the data owner $\mathcal{O}$ or by a specialized auditing authority; or publicly by anyone.

Public auditing only counts on the transactions' information being readable in the blockchain. As such, a public verifier can detect some non-compliant activities, by comparing any transactions where a specific data usage activity is reported, with the smart contract it refers to through its identifier. Indeed, in case the smart contract does not include the data usage activity reported in the transaction, a compliance issue is revealed. This auditing service is made reliable thanks to the tamper-proof feature of the blockchain, and any transactions being signed by the service provider.

Private auditing enhanced by a dedicated auditing organization, relies on public blockchain information, as well as private data, provided by the claimer. For instance, when a data owner requests a private audit for a misuse of his personal data, he shares the private and public keys associated to the concerned smart contracts $(pk_{\mathcal{C}_i}, sk_{\mathcal{C}_i})_{\{i\in[1,N]\}}$, where N is the number of audited smart contracts. As such, the auditing authority is able to lead an investigation, while crawling blockchain transactions corresponding to the provided smart contracts' identifiers. Having received the smart contracts' private keys $sk_{\mathcal{C}_i}$ where $i \in [1, N]$, the auditing authority is able to interpret the content of blockchain transactions and to detect non-compliant activities. We note that private auditing has to be paid by the audited service provider, if non-compliant activities are reported.

7. SECURITY ANALYSIS

In this section, we first present our threat model. Then, we discuss the resistance of our construction against data confidentiality, unlinkability, and availability attacks.

7.1. Threat Model

For designing a secure blokchain-based auditing scheme, we consider realistic threat models. For instance, our contribution does not take into consideration malicious services that deviate from the protocols as we emphasize that service providers stress about their reputation, in real-world use cases. As such, we consider that "an attacker is able to read, send and drop a transaction addressed to the blockchain. The attacker targets data owners, service providers as well as the blockchain" (Kaaniche and Laurent, 2017b), as follows:

- based on previous data usage sessions, as well as provided blockchain data, "an attacker tries to impersonate a data owner to afford a honest service provider some rights to be logged into the blockchain without the legal data owner's consent" (Kaaniche and Laurent, 2017b). This attack is considered with respect to the confidentiality and auditability requirement.
- an attacker attempts to thwart the unlinkability feature w.r.t. data owners. That is, he attempts to link smart contract identifiers or a smart contract to a specific owner.
- an attacker attempts "to prevent the publication of a legitimate transaction in the blockchain. For example, in order to prevent data transfer notification to the data owner, an attacker may try a DoS attack against a data usage event, or attempt a flooding attack on the blockchain with invalid data usage information. This attack is considered against the auditability, the availability and the censorship resistance requirements" (Kaaniche and Laurent, 2017b).

7.2. Security Discussion

This section discusses the security properties, with respect to the defined threat models. Indeed, the confidentiality (section 7.2.1), auditability (section 7.2.3), unlinkability (section 7.2.2), and availability (section 7.2.4) are analyzed while considering different adversaries.

7.2.1. Confidentiality

Our proposed solution is resistant against data secrecy attacks for several reasons here-below listed:

- "only hashed data values and enciphered information published in the smart contract—the client is in charge of hashing his personal data for fulfilling p_1 of the smart contract, and enciphering a secret with $pk_{\mathcal{S}}$ addressed to server $\mathcal{S}$" (Laurent et al., 2018).
- the enciphering key $k_{\mathcal{C},\mathcal{S}}$ only known by a pair of owner and service provider—the owner $\mathcal{O}$ is the only entity owning secret $sk_{\mathcal{C}}$. As a delegated PKG entity, $\mathcal{O}$ is the only entity able to issue the pair of public and private keys $(pk_{\mathcal{C}}, sk_{\mathcal{C}})$ for a specific smart contract. The derived key $k_{\mathcal{C},\mathcal{S}}$ being securely transmitted to the involved service provider, is thus only known to the owner and the service provider.
- one per smart-contract enciphering key—the pair of keys $(pk_{\mathcal{C}}, sk_{\mathcal{C}})$ generated by the owner is specific to a smart contract, and as such can not leak any information, in case they are compromised, about other per smart-contract keys.

As the public ledger only registers hashed data values or encrypted information, no significant information can be learnt from the blockchain.

7.2.2. Unlinkability

Beyond confidentiality guarantee, the unlinkability property is ensured in our approach thanks to the following technical features:

- one smart contract per service provider—each smart contract is specific to a service provider and has its own identifier $ID_{\mathcal{C}}$ and secret key $sk_{\mathcal{C}}$. The owner using this per smart-contract secret key for signing the contract creation can not be identified. It is even not possible to link two smart contracts provided with two different $ID_{\mathcal{C}}$ to the same owner.
- unique hashed values within smart contracts—linking smart contracts in between as issued by the same owner is not possible, as smart contracts always concatenate data values with a specific information they own (their pairwise key $k_{\mathcal{C},\mathcal{S}}$) before hashing. As such, a search over the blockchain ledger for the same data values (assumed to match the same owner) is inconclusive.

7.2.3. Auditability

The proposed approach ensures the auditability requirement as follows:

- tamper-proof architecture—as emphasized in section 6.2.1, "all blockchain-specific operations, such as transaction anchoring activities, are considered as secure and non-corruptible, thus ensuring non-tamper proofs of data processing and managing events" (Laurent et al., 2018).
- transparent usage—our approach is based on a consortium blockchain infrastructure, that permits public access (i.e., read privilege) the contract and its associated transactions, to anyone. Thus, it provides a transparent view over how data are collected and accessed.
- signed transactions—"our approach relies on signed transactions. That is, both smart contract creation and data usage transactions have to be signed by the data owner $\mathcal{O}$ and the service provider $\mathcal{S}$ and $\mathcal{P}$, respectively. Signed transactions ensure that each activity has been efficiently performed by the holder of the used private key, which is certified by the PKG entity. As such, the resistance of the chosen HIBS scheme against forgery attacks has a direct impact on the fulfillment of the auditability requirement" (Kaaniche and Laurent, 2017a; Laurent et al., 2018).
- approval of smart contracts creation— the service provider is requested to approve each smart contract creation by the data owner $\mathcal{O}$, by using its secret key $sk_{\mathcal{S}}$. More precisely, its secret key enables the service provider to decrypt the pairwise key $k_{\mathcal{C},\mathcal{S}}$ associated to the smart contract, and to prove its authenticity and its legitimacy to have a deciphered access to the shared data.

7.2.4. Availability

The blockchain relies on a highly decentralized infrastructure, thus providing our approach with availability assurance and liveness guarantees of data usage. We also point out the similarity between the well known double spending problem in bitcoin architectures (Karame et al., 2015) and the attack aiming at preventing a valid transaction to be registered in the blockchain. Indeed, both assume that an adversary has control over more than a half of the blockchain nodes, the achievement of which is assumed difficult (Karame et al., 2015).

8. CONCLUSION

Personal data are highly exposed to data leakage and misuse by third parties. As such, users have to own a complete control on their personal data usage without compromising their privacy or limiting service providers to propose personalized services and authorities' ability for auditing activities.

This paper introduces a blockchain-based solution for data usage auditing relying on both hierarchical ID-based encryption and signature mechanisms. Each data owner acts as a delegated PKG, and as such has the technical means to provide consent on his data usage and to control data collection and processing activities based on a per smart-contract approach, while enabling service providers to provide the evidence that any personal data processing was previously subjected to the consent by the data owner. Indeed, based on a consortium blockchain infrastructure, the proposed solution first enables the data owner to grant consent to service providers, specify their data access policy and track data usage flows in a trusted and privacy-preserving manner. Second, it provides a regulatory framework to properly enforce the legislation Regulation (EU) (2016). "For instance, in case of a non-compliance activity (i.e., unauthorized data access) reported by a data owner, authorized authorities may lead an investigation referring to as registered blockchain transactions with respect to concerned entities" (Laurent et al., 2018). Third, it helps in resolving availability concerns as blockchain transactions are replicated a large number of times on independent nodes.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

NK did the main job. ML was closely supervising the research technical aspects. CL-B participated on the law aspects (GDPR).

FUNDING

This work was funded by the chair Policies and Values of Personal Information (cf. https://cvpip.wp.imt.fr/en/).

ACKNOWLEDGMENTS

This paper is an extended and revised version of our former conference work (Kaaniche and Laurent, 2017b) accepted in the 16th International Symposium on Network Computing and Applications (NCA), and is an extended discussion of Kaaniche and Laurent (2017c) with regard to smart contract design models and public vs. private blockchains.

The authors would like to thank the reviewer for their careful reading of the manuscript and their constructive remarks.

REFERENCES

Blazy, O., Kiltz, E., and Pan, J. (2014). "(Hierarchical) identity-based encryption from affine message authentication," in *International Cryptology Conference* (Berlin; Heidelberg: Springer), 408–425.

Chow, S. S., Hui, L. C., Yiu, S.-M., and Chow, K. (2004). "Secure hierarchical identity based signature and its application," in *ICICS, Vol. 4* (Malaga: Springer), 480–494.

Crosby, M., Pattanayak, P., Verma, S., and Kalyanaraman, V. (2016). Blockchain technology: beyond bitcoin. *Appl. Innov.* 2, 6–10.

Dierks, T., and Rescorla, E. (2008). *RFC 5246 - The Transport Layer Security (TLS) Protocol Version 1.2*. Technical Report, IETF.

Dorri, A., Kanhere, S. S., Jurdak, R., and Gauravaram, P. (2017). LSB: A lightweight scalable blockchain for iot security and privacy. *arXiv[preprint]. arXiv:1712.02969*. doi: 10.1016/j.jpdc.2019.08.005

Fu, A., Yu, S., Zhang, Y., Wang, H., and Huang, C. (2017). NPP: A new privacy-aware public auditing sc chem for cloud data sharing with group users. *IEEE Trans. Big Data.* doi: 10.1109/TBDATA.2017.2701347

Gentry, C., and Halevi, S. (2009). "Hierarchical identity based encryption with polynomially many levels," in *Theory of Cryptography Conference* (San Francisco, CA: Springer), 437–456.

Gentry, C., and Silverberg, A. (2002). "Hierarchical ID-based cryptography," in *Advances in Cryptology–ASIACRYPT 2002*, 149–155.

Grumăzescu, C., Pura, M.-L., and Patriciu, V.-V. (2015). "Hybrid distributed-hierarchical identity based cryptographic scheme for wireless sensor networks," in *New Contributions in Information Systems and Technologies* (Azore: Springer), 949–958.

Horwitz, J., and Lynn, B. (2002). "Toward hierarchical identity-based encryption," in *Advances in Cryptology–EUROCRYPT 2002* (Amsterdam: Springer), 466–481.

Kaaniche, N., and Laurent, M. (2017a). "Attribute based encryption for multi-level access control policies," in *14th International Conference on Security and Cryptography, SECRYPT 2017* (Madrid).

Kaaniche, N., and Laurent, M. (2017b). "A blockchain-based data usage auditing architecture with enhanced privacy and availability," in *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)* (Cambridge, MA), 1–5.

Kaaniche, N., and Laurent, M. (2017c). *Blockchain Design Directions.* Internal Telecom SudParis Report.

Kaaniche, N., and Laurent, M. (2018). "BDUA: Blockchain-based data usage auditing," in *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)* (San Francisco, CA), 630–637.

Karame, G., Androulaki, E., Roeschlin, M., Gervais, A., and Capkun, S. (2015). Misbehavior in bitcoin: a study of double-spending and accountability. *ACM Trans. Inform. Syst. Secur.* 18:2. doi: 10.1145/2732196

Laurent, M. (2019). Authenticated and privacy-preserving consent management in the internet of things. *Proc. Comput. Sci.* 151, 256–263. doi: 10.1016/j.procs.2019.04.037

Laurent, M., Kaaniche, N., Le, C., and Vander Plaetse, M. (2018). "A blockchain-based access control scheme," in *15th International Conference on Security and Cryptography (SECRYPT)* (Porto), 168–176.

Lee, G. Y., Cha, K. J., and Kim, H. J. (2019). "Designing the gdpr compliant consent procedure for personal information collection in the iot environment," in *2019 IEEE International Congress on Internet of Things (ICIOT)* (Milan), 79–81.

Liang, X., Shetty, S., Tosh, D., Kamhoua, C., Kwiat, K., and Njilla, L. (2017). "Provchain: a blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability," in *Proceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing* (Madrid: IEEE Press), 468–477.

Linn, L., and Koo, M. (2016). *Blockchain for Health Data and Its Potential Use in Health IT and Health Care Related Research.* Available online at: https://www.healthit.gov/sites/default/files/11-74-ablockchainforhealthcare.pdf

Morel, V., Cunche, M., and Le Métayer, D. (2019). "A generic information and consent framework for the iot," in *2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)* (Rotorua), 366–373.

Neisse, R., Steri, G., and Nai-Fovino, I. (2017). A blockchain-based approach for data accountability and provenance tracking. *arXiv[preprint]. arXiv:1706.04507*. doi: 10.1145/3098954.3098958

Ouaddah, A., Abou Elkalam, A., and Ait Ouahman, A. (2016). Fairaccess: a new blockchain-based access control framework for the internet of things. *Secur. Commun. Netw.* 9, 5943–5964.

Perng, G., Reiter, M. K., and Wang, C. (2005). "Censorship resistance revisited," in *Information Hiding* (Barcelona: Springer), 62–76.

Prabhakaran, M. M., and Sahai, A. (2013). *Secure Multi-Party Computation, Vol. 10*. IOS Press.

Ramachandran, A., and Kantarcioglu, M. (2018). "Smartprovenance: a distributed, blockchain based dataprovenance system," in *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy* (Tempe, AZ: ACM), 35–42.

Rantos, K., Drosatos, G., Kritsas, A., Ilioudis, C., Papanikolaou, A., and Filippidis, A. P. (2019). A blockchain-based platform for consent management of personal data processing in the IoT ecosystem. *Sec. Commn. Netw.* 2019:1431578. doi: 10.1155/2019/1431578

Regulation (EU) (2016). *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, and Repealing Directive 95/46/ec (General Data Protection Regulation), ojeu l 119/1 of 4.05.2016*. Regulation (EU)

Seo, J. H., and Emura, K. (2015). "Revocable hierarchical identity-based encryption: history-free update, security against insiders, and short ciphertexts," in *Cryptographers Track at the RSA Conference* (San Francisco, CA: Springer), 106–123.

Shamir, A. (1985). "Identity-based cryptosystems and signature schemes," in *Proceedings of CRYPTO 84 on Advances in Cryptology* (New York, NY: Springer-Verlag New York, Inc.), 47–53.

Shetty, S., Red, V., Kamhoua, C., Kwiat, K., and Njilla, L. (2017). "Data provenance assurance in the cloud using blockchain," in *Disruptive Technologies in Sensors and Sensor Systems, Vol. 10206* (International Society for Optics and Photonics), 102060I.

Swan, M. (2015). *Blockchain: Blueprint for a New Economy*. O'Reilly Media, Inc, Sebastopol, CA.

Tian, M., and Huang, L. (2014). "Efficient identity-based signature from lattices," in *IFIP International Information Security Conference*, (Marrakesh: Springer), 321–329.

Tosh, D., Shetty, S., Foytik, P., Kamhoua, C., and Njilla, L. (2018). "Cloudpos: a proof-of-stake consensus design for blockchain integrated cloud," in *2018 IEEE 11th International Conference on Cloud Computing (CLOUD)*, (San Francisco, CA).

Vukolic, M. (2017). "Rethinking permissioned blockchains," in *Proceedings of the ACM Workshop on Blockchain, Cryptocurrencies and Contracts, BCC '17* (New York, NY: ACM), 3–7.

Wood, G. (2014). Ethereum: a secure decentralised generalised transaction ledger. *Ethereum Project Yellow Paper*, 151. Available online at: https://ethereum.github.io/yellowpaper/paper.pdf

Yu, F. R., Tang, H., Mason, P. C., and Wang, F. (2010). A hierarchical identity based key management scheme in tactical mobile ad hoc networks. volume 7, pages 258–267. IEEE.

Zhang, Y., Lin, X., and Xu, C. (2018). "Blockchain-based secure data provenance for cloud storage," in *International Conference on Information and Communications Security* (Lille: Springer), 3–19.

Zyskind, G., Nathan, O., and Pentland, A. S. (2015). "Decentralizing privacy: using blockchain to protect personal data," in *Security and Privacy Workshops (SPW), 2015 IEEE* (San Jose, CA), 180–184.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

HYPOTHESIS AND THEORY
published: 28 May 2020
doi: 10.3389/fbloc.2020.00026

Digital Identity and the Blockchain: Universal Identity Management and the Concept of the "Self-Sovereign" Individual

Andrej J. Zwitter[1], *Oskar J. Gstrein*[2]* *and Evan Yap*[3]

[1] University of Groningen, Campus Fryslân Data Research Centre, Professor of Governance and Innovation, Leeuwarden, Netherlands, [2] University of Groningen, Campus Fryslân Data Research Centre, Assistant Professor, Leeuwarden, Netherlands, [3] Lead of Research & Development at Mark Labs, Washington, DC, United States

While "classical" human identity has kept philosophers busy since millennia, "Digital Identity" seems primarily machine related. Telephone numbers, E-Mail inboxes, or Internet Protocol (IP)-addresses are irrelevant to define us as human beings at first glance. However, with the omnipresence of digital space the digital aspects of identity gain importance. In this submission, we aim to put recent developments in context and provide a categorization to frame the landscape as developments proceed rapidly. First, we present selected philosophical perspectives on identity. Secondly, we explore how the legal landscape is approaching identity from a traditional dogmatic perspective both in national and international law. After blending the insights from those sections together in a third step, we will go on to describe and discuss current developments that are driven by the emergence of new tools such as "Distributed Ledger Technology" and "Zero Knowledge Proof." One of our main findings is that the management of digital identity is transforming from a purpose driven necessity toward a self-standing activity that becomes a resource for many digital applications. In other words, whereas traditionally identity is addressed in a predominantly sectoral fashion whenever necessary, new technologies transform digital identity management into a basic infrastructural service, sometimes even a commodity. This coincides with a trend to take the "control" over identity away from governmental institutions and corporate actors to "self-sovereign individuals," who have now the opportunity to manage their digital self autonomously. To make our conceptual statements more relevant, we present several already existing use cases in the public and private sector. Subsequently, we discuss potential risks that should be mitigated in order to create a desirable relationship between the individual, public institutions, and the private sector in a world where self-sovereign identity management has become the norm. We will illustrate these issues along the discussion around privacy, as well as the development of backup mechanisms for digital identities. Despite the undeniable potential for the management of identity, we suggest that particularly at this point in time there is a clear need to make detailed (non-technological) governance decisions impacting the general design and implementation of self-sovereign identity systems.

Keywords: blockchain, digital identity, self-Sovereign identity, governance, innovation, human dignity

OPEN ACCESS

Edited by:
Glenn Parry,
University of Surrey, United Kingdom

Reviewed by:
Kaliya Young,
Merritt College, United States
Richard Tighe,
Oxfam, United Kingdom

***Correspondence:**
Oskar J. Gstrein
o.j.gstrein@rug.nl

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *13 September 2019*
Accepted: *29 April 2020*
Published: *28 May 2020*

Citation:
Zwitter AJ, Gstrein OJ and Yap E (2020) Digital Identity and the Blockchain: Universal Identity Management and the Concept of the "Self-Sovereign" Individual. Front. Blockchain 3:26. doi: 10.3389/fbloc.2020.00026

INTRODUCTION

As mankind continues its journey through the Digital Age our lives are increasingly becoming compositions of our offline and online activities. While the dimensions of "classical" human identity have kept philosophers busy since millennia[1], traditional thinking about "Digital Identity" is primarily machine related. Telephone numbers, E-Mail inboxes, or Internet Protocol (IP)-addresses seem to be irrelevant to define us as human beings at first glance. However, the discussion about surveillance in the digital domain (Council of Europe, 2018), and jurisprudence of the European Court of Justice (ECJ) struggling to clarify under which conditions IP-addresses should be qualified as personal data (ECJ, 2018; Gstrein and Ritsema van Eck, 2018, p. 80–81), show in detail how these technical necessities make it increasingly difficult to distinguish between our offline and online selves. The omnipresence of digital technology and its use to not only control, but also shape society result in the need to reconsider our world and ourselves as beings (Galič et al., 2017). The technological component of the amalgamation that we call "me" has increased considerably in the last decades (Kucklick, 2014, p. 189–235). When looking at these changes from a holistic perspective, it is almost natural to deduct that this digital space as parallel space does not mirror existing governance structures, power relations, human rights, and legal obligations. "Code is Law"—or at least has sometimes normative authority—and as it spreads across our lives new governance decisions are made by those who shape it in an ex- or implicit manner (Lessig, 1996, p. 1–9). Furthermore, it is not only governmental surveillance and "nudging" that shapes our digital identities right from the start, there is also "surveillance capitalism" (Zuboff, 2019, p. 11–12).

In an attempt to create an explicit and universal process, the United Nations (UN) in late 2013 have made a cautious start to address these developments by "[r]ecognizing that the same rights that people have offline must also be protected online." (United Nations, 2014, p. 2). In 2018 they called upon states to "[…] consider developing or maintaining and implementing adequate legislation, in consultation with all relevant stakeholders, including civil society, with effective sanctions and appropriate remedies, that protects individuals against violations and abuses of the right to privacy, namely through the unlawful and arbitrary collection, processing, retention or use of personal data by individuals, governments, business enterprises and private organizations […]" (United Nations, 2018, p. 6). In this spirit and to ensure that specifically digital identities live up to these requirements, the UN have supported the creation of the ID2020 Alliance[2]. As noble as these intentions might be, large corporations such as those assembled under the GAFAM acronym (Google, Apple, Facebook, Amazon, Microsoft) continue their considerable efforts to create a general identity for logins of digital services they themselves or others provide. It is practically impossible to activate the holy grail of expression in the digital age—the modern smartphone—without one or more accounts that are associated with those big players. However, since these actions result in siloed identities tied to proprietary services and applications (Verborgh, 2019), the advent of digital identity systems based on Blockchain and similar Distributed Ledger Technologies (DLT) might offer the opportunity for change.

More than a decade ago, Cameron (2005) formulated seven laws of identity aiming to guide the way from a patchwork of identity one-offs to a universal identity. Cameron's visionary view on the subject of identity, claims and privacy led him to formulate the following principles: (1) user control and consent; (2) minimal disclosure for a constrained use; (3) justifiable parties; (4) directed identity; (5) pluralism of operators and technologies; (6) human integration; (7) consistent experience across contexts. While it would go too far to explain each of them in detail, one can summarize that Cameron laid the foundational principles that many actors in the field of digital identity are aiming for. Only today, first attempts to a universal identity are made, but the reality remains that the individual is composed of a patchwork of identities, logins, usernames, passwords, etc.

The obstacle to an overarching digital-identity is the enforcement of one standard in cyberspace, as the battle over single-log-on's between Google and Facebook illustrates. Interestingly enough, the solution might not be found in the private, but in the public sector. For example, the Netherlands are using a progressive digital identity management system called "DigiD" which allows residents access to public records and governmental services since several years[3]. Furthermore, Georgia and Sweden used blockchain technology to create an immutable record of land titles, identifying individuals as landowners (Nimfuehr, 2018). The World Food Programme pioneered a similar Blockchain guided approach to biometric ID and digital payments with which refugees in a camp in Jordan could reserve funds and buy goods, without needing physical documents or valets (Juskalian, 2018; Wang and De Filippi, 2020, p. 14–17). Furthermore, the European Union (EU) is contemplating digital identity with the ascend of Schengen II and works on the mobility of identity related credentials in its member states through the implementation of the eIDAS Directive (EU Regulation No 910/2014).

Additionally, the concept of "self-sovereign identity" is emerging. While there is currently no universally and legally binding definition of the concept, Allen (2016) has described it as "[…] the next step beyond user-centric identity and that means it begins at the same place: the user must be central to the administration of identity." He goes on to propose ten principles that aim at describing the main characteristics of self-sovereign identity. Wang and De Filippi (2020, p. 9–11) discuss his proposal in-depth, and together with the relationship to so-called decentralized identifiers (DIDs), which are a form of URLs (e.g., unique web addresses) that resolve to a DID

[1]For illustrative purposes a bridge can be built from Aristotle who framed the human as "political animal," to Rene Descartes "rational self" in the context of "cogito ergo sum," further to Martin Heidegger's existence or "Dasein": Dieses Seiende, das wir selbst je sind und das unter anderem die Seinsmöglichkeit des Fragens hat, fassen wir terminologisch als Dasein.

[2]Available online at: https://id2020.org (accessed March 4, 2020).

[3]Available online at: https://www.digid.nl/en/what-is-digid (accessed March 4, 2020).

document (Wang and De Filippi, 2020, p. 9)[4]. Wagner et al. (2018, p. 27) have proposed to define self-sovereign identity as "a model of digital identity where individuals and entities alike are in full control over central aspects of their digital identity, including their underlying encryption keys; creation, registration, and use of their decentralized identifiers [. . .] The architecture gives individuals and entities the power to directly control and manage their digital identity without the need to rely on external authorities." In other words, DLT becomes an infrastructure for the creation of verifiable credentials. It allows to verify claims with (increased) independence of governments or powerful intermediaries. Ultimately, this new approach could replace conventional "legal identities."

Furthermore, if such self-sovereign identity management includes verification of claims based on "zero-knowledge proof" (ZKP) technology, the need to exchange and register "raw" personal data on platforms could be limited significantly. Wagner et al. (2018, p. 48) describe ZKP by stating that its "[. . .] use allows two different actors, the "prover" and the "verifier" to exchange the ownership of a piece of data, without actually revealing the data. The math, probability and cryptography behind this technology makes their application useful in for example allowing the verifier to prove the ownership of a credential to the verifier, such as a driving license [sic!] without revealing the identifier it has been initially issued to."

However, all of these attempts do not consider how identity is constructed from a philosophical perspective, what the legal consequences attached to identity look like, and what a universal digital identity should mean. This general lack of understanding of philosophical and legal underpinnings of identity remains prevalent. Nevertheless, upon closer scrutiny this is not surprising. As illustrated below, the concept of identity is ill-defined in the legal domain, and an attempt to define it spans many different fields with relevant norms in national and international legislation, as well as consideration of corresponding jurisprudence. In order to shed light onto what identity is composed of in practice, this paper presents:

1. An overview of selected philosophical perspectives on identity;
2. An overview of legal aspects of identity;
3. Expressions of identity and ways of identification;
4. A categorization of digital identity;
5. A discussion of opportunities and risks around digital identity.

With this primer on digital identity that includes perspectives on self-sovereign identity, we respectfully submit that before we are even able to seek a sustainable technological solution, identity as a concept should be carefully considered, taking into account the development of cyberspace from a philosophical, legal, and ethical perspective. We acknowledge that major strides toward digital identity are currently being undertaken by the DLT community. Furthermore, it is not our aim to replicate and review the significant number of purely technology-related blogs and white papers written on the subject. Rather, we aim to contribute to the concept of digital identity from a philosophical and legal perspective. Thereby, this paper aims to counteract the current trend to immediately jump on technology on the peak of the "hype-cycle" and present it as panacea for all problems.

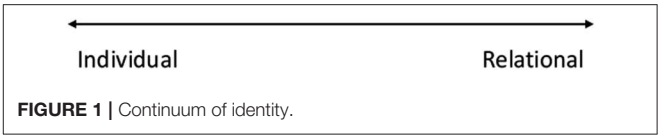


FIGURE 1 | Continuum of identity.

Another important disclaimer is that this paper does not treat digital identity as belonging to DLT exclusively. DLT is just one of many tools that are currently being deployed. Hence other, more centralized systems will also be discussed in their own right. Once these elements are profoundly understood in a larger context and from a legal and philosophical perspective, they can function as determinants of what we need technology to do for digital identities to truly enrich society and make human existence more dignified.

PHILOSOPHICAL PERSPECTIVES ON IDENTITY

The concept of identity in philosophy has many different aspects. Identity plays a central role in logics and metaphysics, in existentialism and other areas. To frame this discussion, we deem it useful to distinguish between two extreme positions at either side of an identity continuum: The naturalist world view ("identity is whole and distinguishable") and the constructivist world view ("identity is compartmentalized and shared"; see **Figure 1**). Roughly, the naturalist argues that identity is tied to the properties of an object or a person. In contrast, the constructivist sees identity as a whole constructed out of the relationship between objects and subjects. As we will show, both result in different degrees to which a universal digital identity can be realized and to what extent such a universal identity is limited by the identity and rights of others.

Naturalist World View

In a simplified manner, the naturalist worldview assumes that everything that resides inside the physical body or is more permanently connected with it forms its identity. It is thus the nature of the physical body and its delineation from other physical bodies that make it unique and distinguishable. From a metaphysical perspective, every physical object has unique properties, be it the position in space, its texture etc. Two objects are non-identical if they differ in at least one of their conditions (e.g., texture, color, composition etc.). Prominent proponents of this world view in recent times include John Dewey, Ernest Nagel, Sidney Hook and Roy Wood Sellars (Papineau, 2016). A common problem around identity is whether an object can be distinguished from another, if it does not differ in any of the conditions that define it as an object. This has already quite profound consequences for digital identity and corresponding problems in data protection, particularly when considering the

[4]Available online at: https://w3c.github.io/did-core/#introduction (accessed March 4, 2020).

use of biometric data for identification purposes (Jasserand, 2018, p. 155). Since digital identity by necessity is a digitalized and reduced reflection of what one voluntarily and involuntarily (e.g., think "data exhaust" or "data trail") projects into the digital sphere, any identity must be sufficiently distinguishable from other identities. **Uniqueness** is thus one criterion that derives from this naturalist world view.

Further assuming that both mind and soul reside in the physical body, this allows us to draw further relevant conclusions on identity. For example, Pythagorean and Platonic transmigration theories ("the wandering soul") raise a rather problematic aspect about whether the physical delineations are correct (Luchte, 2012, p. 174–177). It seems that the naturalist world view becomes already somewhat inconsistent. Does identity actually reside within the boundaries of the body, the soul, or a combination of it? If the answer is Yes, then the physical body is mostly a vessel with sufficient distinctive features for identification. The problem becomes even more pressing as researchers try to use personal profiles on social media and other "data exhaust" of persons to make them digitally "immortal" through the use of artificial intelligence, which also raises the interesting aspect of post-death autonomy or "post mortem privacy" (Harbinja, 2017). If the body has to be considered merely as vehicle for the mind, this means for digital identity and identification that we need to postulate a **Priority Thesis** of mind over matter. This is certainly the case when it comes to questions of uniqueness and of authentication. If the material body is subject to drastic change that can make it a sufficient representation for identification (biometrics), then priority has to be given to the mind for identification (knowledge, passwords, relationships). This leaves us with the following elements for identification:

- Physical body (nature): color of eyes, height, hair color, facial structure, iris, finger, and palm print etc. Essentially, everything that can be used to create biometrical data. Any of these features are subject to change and ultimately serve only as a proxy.
- Non-physical body (nurture): everything that was trained, what one has absorbed into character, education, training, behavior to the extent specific to my identity and sufficiently distinguishable from others. Expressions of the mind in this category are passwords and phrases (security) questions, and answers that only a particular user can know, as well as certain knowledge or abilities.

As argued, two principles of identity and authentication can be derived from the naturalist worldview: Uniqueness ("being sufficiently distinguishable from any other entity") and the Priority Thesis ("expressions of the mind have priority over the physical body"). They, however, raise questions themselves. For example, how distinct from others is an individual if crucial parts of its identity have been formed by nurture such as education, culture etc. which are ultimately derived from other people, and which can also be accessed by other people. This leads us to the other extreme on the identity continuum, constructivist identity.

Constructivist Identity

The extreme position of this view proclaims that identity is wholly shaped by social structure, the relationship with others, the norms and rules that shape the environment, institutions that testify to the existence of an identity (e.g., the government issuing an identity card), and other external factors with which a person interacts ("look at the person with the funny hat" -> becomes "person known for wearing a funny hat"). Social networks, such as LinkedIn and Facebook, very much take this stance: you are who you are connected with. Your social network and the nature of the individual relations to others define your uniqueness. This immediately raises a less discussed subject in the literature around constructivism, namely its relation to network theory. A detailed discussion would go beyond the scope of this primer, but it suffices to say that one important aspect of identity, namely power or the ability to influence others in network theory is a function of centrality within the network. This is helpful when trying to understand how Instagram and Youtube "stars" are able to leverage their identity for fame and financial gains (Rueb, 2019).

Constructivism in a less extreme form would at least argue that identity is relational (**Relationality**). In other words, who you are is a matter of who you relate to and the nature of this relationship. It is therefore not only the relation of identity X to Y (as existent or not existent), but also the relation of X to Y as a father to a son, as brothers, as spouses etc. This view also ties in with the philosophy of Jean-Paul Sartre as he emphasizes that individual identity is shaped by the interaction with "the other person" who is difficult to define in detail. While the environment seems entirely open and explorable to us, the other person is difficult to grasp, yet exists as an undeniable fact (Sartre, 2007, p. 412–413). One more philosophical reference that comes to mind is Georg Wilhelm Friedrich Hegel's dialectic of master and slave (or Lord and Bondsman more precisely). Here, identity is also defined through the role of individuals in society and the relationship they have with each other (Lichtenthähler, 2019, p. 104–123).

Quite commonly, such relationships and family relationships, are registered by governments and expressed through documentation of relational identity such as birth certificates and the entry of someone's marital status into public record. These relational identity features are affecting the legal status of a person. They also define contractual rights and obligations, and other societal obligations such as tax payment. Hence, a public record of these identity features is crucial for society to work. Furthermore, the change in someone's identity can also result in changes for the identity of another person (e.g., with the death of a spouse the other becomes a widower). If identity is relational, information about this identity (often referred to as Personally Identifiable Information or "PII") can also be information about another identity. This aspect connects to the discussion on "data ownership" which is another topic of utmost importance for the future of cyberspace and its governance as it takes more and more control of physical infrastructure. It seems much more complicated than the common discourse suggests, because data about one person can also be data about another person. The information that X is friends with Y is a ready

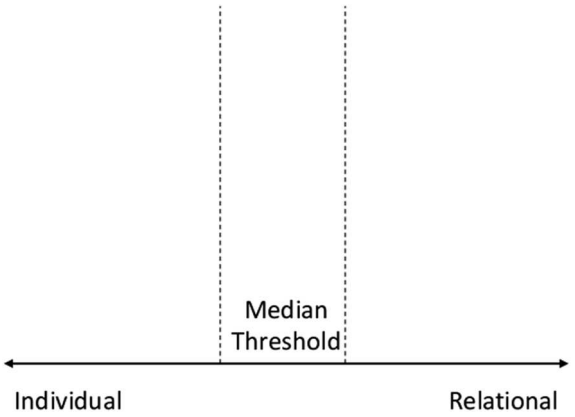


FIGURE 2 | Median threshold between individual and relational identity.

example of this conundrum as to who owns that information. Other relational criteria could be education (e.g., relationship with an educational institution—educational certificate), work (e.g., employer/employee—contract), culture and religion (e.g., membership with a religious community), hobbies and memberships in clubs, as well as membership in political parties. The crucial question is to what extent an extreme position of constructivist identity is feasible. After all, such a stance would be in an uneasy relationship with the principle of uniqueness.

As a result of the discussion on philosophical aspects of identity, we can postulate the existence of a **Median Threshold** that acts as a balance between the naturalist and the constructivist perspective (see **Figure 2**). The Median Threshold at the very least proposes the assumption that either extremes are untenable positions whether in the theory around identity or in the practice of digital identity. The truth and technologically most feasible solutions are probably found somewhere in the middle. For identity in general and digital identity specifically, one aspect can however be deduced. Whether identity is derived from either of these perspectives, they are always expressed in every individual in the whole of their combinations. In other words, every individual will contain and express all of the determinant factors of identity as a whole in itself. "I am a son, a father etc." remains also valid in absence of the other person"s presence to confirm such statements. In summary, this last aspect highlights a principle of identity without which the idea of universal identity is not workable: **Wholeness**.

LEGAL ASPECTS OF IDENTITY

In law, the concept of identity is rather badly developed and dispersed amongst several categories. It can be linked to the concept of personhood which is difficult to define in detail as well (Foster and Herring, 2017). We have already mentioned identity features to which legal consequences are attached, such as life and death. However, "legal identity" (or personhood) is also a relevant precondition to be able to sign a contract, being recognized as child or parent, as well as being entitled to vote or applying for a public post. However, identity as such is not at the center of those legal transactions and legal personhood could be described as a "technical device" (Brozek, 2017, p.12). On an international and national level, our first association might be identity cards and passports, which are attached to citizenship. While many will perceive "their citizenship" as natural and not further noteworthy, this seemingly evident concept turns out to be much vaguer and more complex upon closer inspection (Kochenov, 2009, p. 175–181).

International Law

International law, national law and individual identity are closely connected in the field of human rights (Mutua, 2016, p. 172–174). The state decides over the citizenship of a person and has the right to issue passports and identity cards as an intrinsic property of statehood and sovereignty. Associated with citizenship is the international legal prohibition of statelessness. No person should be without citizenship, because without it one has no legal recourse (Kochenov, 2009, p. 175–181). Legal personhood as a universal human right is enshrined in Article 16 of the 1966 International Covenant on Civil and Political Rights of the UN (ICCPR) and entails the right to be recognized as a person before the law (Blitz and Sawyer, 2011, p. 3–4). In short, recognition as a person with rights and duties is a fundamental aspect of identity, because it enables a person to enjoy associated elements that determine daily life and individuality such as:

- The right to life and to personal integrity as enshrined in Article 6 paragraph 1 ICCPR, as well as prevention from arbitrary arrest as enshrined in Article 9 paragraph 1 ICCPR: the naturalist world view is clearly expressed in this right, as the physical person is protected from any arbitrary interference into its biological workings and into its liberty of movement[5].
- The right to privacy and family life that includes the protection of honor and confidential correspondence as enshrined in Article 17 ICCPR (Cannataci, 2017, p. 36–41).
- The freedom of thought, conscience and religion as enshrined in Article 18 paragraph 1 of the ICCPR: this fundamental right attaches to the non-physical identity. We already discussed that the Priority Thesis postulates that the individual predominantly is an individual because of its mind rather than its body. In this composite human right, different aspects of non-physical identity come to the fore: the right to think what one wants, the right to build ones world view and the right to adopt any religion one wants (with the caveat that it is usually up to the state which religions are officially recognized).
- The freedom of expression as enshrined in Article 19 paragraph 1 and 2 ICCPR: as an extension of someone's non-physical identity, freedom of expression allows the external projection of identity, and can be seen as closely connected to identity.
- Furthermore, minority rights as individual rights to practice culture and religion are granted to persons who belong to a certain group identity as enshrined in Article 27 ICCPR.

[5]As for most human rights there are limitations and derogations possible, as for protection of the public order and for state of emergencies.

- Finally, in democratic countries legal identity also extends to political identity and includes the right to vote and to be elected. For example, this right is enshrined in Article 3 of the protocol of the European Convention on Human Rights of the Council of Europe from 1952.

All of the above shows that specifically within the field of human rights law, that has its heritage in humanism and individualism developed particularly during the period of enlightenment (Morsink, 2012, p. 1–13), individual identity finds several important elements and protection mechanisms. In addition, human rights law already determines which aspects of individual and shared identity are to be recorded to ensure their protection. We can summarize this as **Legal Determinacy** of identity. Some identity aspects are simply a legal necessity, specifically legal personhood, as they associate with the expression of individual identity in so many other areas.

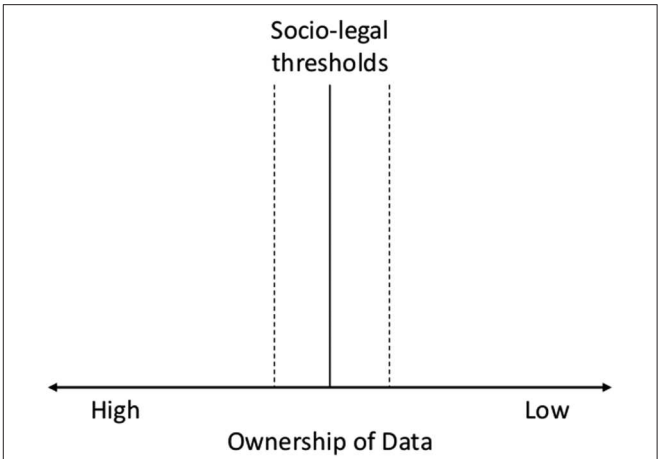


FIGURE 3 | Socio-legal thresholds regarding the individual ownership of data.

National Law

As already stated above many identity aspects that pertain to human rights are simultaneously relevant internationally and nationally. This is particularly visible in the case of European Union citizenship which has a hybrid status between international and national identity. It adds additional features for member state citizens while also producing effects for those who are not citizens of one of the member states of the union (Kochenov, 2009, p. 234–237). Nevertheless, in addition to the aforementioned rights states have developed a host of other identity related norms. As they are being dealt with quite differently across different legal traditions, this section will treat only a select few that one quite commonly encounters:

- Property rights and associated duties: ownership is an important identity characteristic particularly in market economies. Landownership is a specific subset of property rights as it is not only governed by contract law, but also by public law through land registries. This peculiarity stems from two considerations: First, with land being a high-stake property landownership deserves an increased protection by the law. Secondly, since land is placed on the territory of a state, the state reserves itself the right to govern this property title and in some states in cases of necessity, e.g., for the creation of public services, even to disown landowners (mostly under very strict conditions). Property is associated with the duty to pay taxes, which is why financial records and individual taxation form part of individual identity as a citizen or resident.
- Intellectual property rights: as expression of ownership of the products of one's individual mind, these include the rights to exclusively profit from artistic, technical, and scientific output. This element is very interesting in today's data economy as discourse on data ownership is emerging (Tjong Tijn Tai, 2018). Intellectual property rights pertain to a specific kind of data one is producing and that protects property of data of a certain quality with regards to artistic and intellectual qualities. As the value of personal data shifts, and as personal data has become a commodity much value, the threshold of artistic and intellectual quality of data needs to shift in favor of the person producing the data.

In summary, both international and national law impose a minimum set of data that by law is associated with the individual in most countries. In answer to the question of whether the individual would own a lot of data or very little, the legal requirements already indicated that no ownership of data around the identity as a citizen is not plausible. At the same time, as already indicated in the section on constructivist identity, much of what concerns identity data is relational data (e.g., being a mother relates to one's children, being a teacher to one's school and pupil etc.). This relational data is per definition data that concerns more than one entity. This means, the rights of one person to such relational identity data end where another person's rights start. This naturally leads to a socio-legal threshold of a minimum and a maximum ownership of data (see **Figure 3**).

PHILOSOPHICAL AND LEGAL THRESHOLDS OF DIGITAL IDENTITY

Summarizing what has been stated above in the sections on philosophical and legal aspects of identity, we observed that identity is generally regarded as a mixture of individual determination and relational aspects. While the **naturalist world view** establishes identity as a concept that hinges on the concept of **uniqueness** of any identity, it also evokes questions on the **priority thesis** or the dependence and interaction of an individual with its environment and society. Proponents of a constructivist identity emphasize **relationality** while questions of identity as a complete individual entity (**wholeness**) remain. As we went further to consider the legal domain, we observed that particularly in the human rights space identity is determined by several individual rights that states are obliged to grant to individuals (**legal determinacy**). Furthermore, aspects around the ownership of material and immaterial goods ultimately

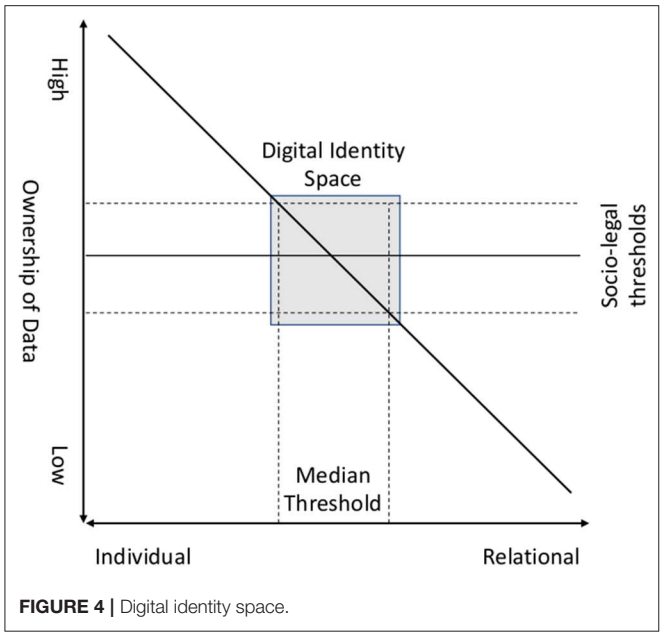


FIGURE 4 | Digital identity space.

highlight the issue of **data ownership** which could be essential to manifest fundamental and simple rights in the digital domain.

Bringing all the above together, allows us to hypothesis a space of digital identity that is both a compromise between the socio-legal threshold, as well as the median threshold between individual and relational identity conceptions. **Figure 4** visualizes this complex triangulation of philosophical and socio-legal conceptions about identity and highlights requirements which the digitization of identity—including DLT—might have to live up to:

DIGITAL IDENTITY CASES

While Blockchain and other DLTs (e.g., Ethereum, Hyperledger Indy, Veres One, IOTA) enable new paradigms of identity management (Blockchain Bundesverband, 2018), we strive to develop an initial categorization of digital identity schemes. As we continue to explore digital identity by portraying case studies of pilots in the public and private sector, including such implementing DLT, "self-sovereign identity," and ZKP, we will first briefly consider the drivers and trends that create the demand and opportunity for the digitization of identity management.

Digitization of Identity Management

Our digital and physical lives are becoming increasingly interlinked through apps and services. Our digital representation is embodied in the many versions and usages of digital identity through which we interact. While there are many reasons for these developments, we wish to highlight three trends that create demand for enhanced digital identities:

1. Necessity to improve process management and information security to protect against cybercrime and identity theft.
2. The feature of decentralization as an enabling factor for the individual.
3. The desire for increased participation in, as well as increased access and efficiency of social service provision and access to the economy.

To discuss the first trend, we should start with establishing that traditionally identity management encompasses a broad spectrum of instrumental identities, which serve the purpose of identification within a specific system or network. In a 2018 report on identity in a digital world of the World Economic Forum three archetypes are described:

- **Centralized** identity systems, where a single organization establishes and manages identity. This is typical for the direct relationship between a state and the individual. Examples include a government electoral roll, or a land registry system, but also the relationship with private actors such as a bank.
- **Federated** identity systems, where different public and private institutions establish stand-alone systems. These systems are subsequently linked through agreements or regulation and each of the managing institutions of the systems becomes a trust-anchor. This allows for some re-purposing of identity credentials, yet the activity remains driven by the initial purpose. For example, a driving license has the primary purpose to show the ability of someone to drive a vehicle, while it might also be used occasionally to prove age in other societal contexts.
- **Decentralized** identity systems, where the individual is at the center and institutions or private corporations just add (verified) credentials to a central "identity hub", "application", or "vault" that is controlled by the individual. In such a system, digital identity is initially purpose-free and becomes a resource or an asset as credentials are acquired (World Economic Forum, 2018b, p. 13–16).

As we will show below when discussing case studies, this last approach of decentralized identity management is becoming more and more popular. As the digital sphere emerges, one of the biggest challenges of siloed approaches of central and federated systems is that they overburden users with identity management. Many different accounts to access their digital identities for each use case are necessary. Already in 2015, it was estimated that in the United States an average e-mail address is associated with 130 digital accounts (Le Bras, 2015). Unsurprisingly, "identity theft" has become one of the major cyberthreats and most common cybercrimes that enables numerous fraudulent activities (Wall, 2013, p. 438–440). Furthermore, the weakness of such systems lies not only on the side of the individual. Data leaks have become a massive problem, as two data protection related fines of the United Kingdom's Information Commissioner's Office (ICO) of 2019 illustrate. As a result of an investigation, the international hotel group Marriott was fined almost £100 m after hackers stole the records of 339 million guests (Sweney, 2019b). In another case that led the ICO to levy the largest fine to date in applying the European Union General Data Protection Regulation (GDPR), the airline British Airways was required to pay £183 m in compensation after it was found that an extensive

amount of data (including login, payment card, name, address, and travel booking) of 500.000 users was stolen (Sweney, 2019a). Hence, there is a clear necessity to improve process management and information security in order to be able to use the digital sphere for meaningful exchanges of information and services.

However, this is not the only way to describe the ongoing transformation. Another important factor in this development is the question who or which entity is controlling/"owning" identity related information. Traditionally, the ownership of such identities lies with the issuer of the credentials, since this is also the institution that is guaranteeing for validity and thereby establishing trust. Hence, artifacts that prove identity such as passports need to be paid for, or the information associated with identity is monetized by private corporations in one of the manifold schemes of "surveillance capitalism" (Zuboff, 2019, p. 10–29). With decentralization however, this pattern might be changing. When focusing on this governance layer of digital identities, three types can be distinguished (Gstrein and Kochenov, 2020):

- **Centralized Top-Down-Approach**, such as for example applied in the world's largest digital identity program "Aadhaar" that is administered by the Unique Identification Authority of India since 2009[6]. This is a centralized system with more than 1.2 billion enrolled users that is not DLT based, but hinges heavily on biometrics to identify users (Rao, 2019), which is also discussed in the context of digital identity systems supported by DLT. The main purpose of Aadhaar is to improve social service provision, while critics fear that it is applied in inappropriate contexts as well. Such "mission creep" might affect the potential for development of groups negatively, and disproportionately limit individual privacy (Privacy International, 2018).
- **Individual Incentive Programs**, such as for example the E-residence scheme of Estonia. Essentially, individuals become virtual residents of Estonia which gives them a platform to operate from regardless of where they originate from. With this, a country tries to get more attractive for investment, or individuals who would like to setup a business. It can also be tied to other policy objectives such as emphasizing certain characteristics of a government, as well as creating a national brand (Poleshchuk, 2016).
- **Community-Based Bottom-Up Approach**, such as the decentralized identity platform Forus in the Netherlands[7]. We discuss it in more detail in the case studies below. Such systems are decentralized by design and entirely user focused. Platforms like Forus develop features incrementally as they grow from concrete use scenarios within communities to regional and potentially global relevance.

Another promise of putting the individual "in charge" of its own identity management has to do with aspects of social participation and the third trend creating demand for enhanced digital identities. Although it is often a non-issue for individuals living in developed countries, many people across the world are excluded from the social ecosystem as they cannot prove their identity. The World Bank Group estimated in 2018 that globally ∼1 billion people face challenges when proving their identity (World Bank Group, 2018, p. 3). The UN acknowledged this pressing need in Sustainable Development Goal (SDG) 16.9, which enshrines a right to legal identity for all, including birth certificates. However, this requires a complete overhaul of the way identities are managed on a national and global scale and will take years if not decades to change. Here, DLT based digital identities may become a catalyst for change, but it remains to be seen how countries which are at times struggling with basic infrastructural needs will be able to leapfrog toward fully decentralized digital identity. Nevertheless, according to an estimation by the World Economic Forum published in late 2018, there will be 150 million people with blockchain based identities by 2022 (World Economic Forum, 2018a, p. 19).

Case Studies in the Public Sector

Public-based digital identity solutions revolve around citizenship and the usage in the interaction with public and private institutions. Governments provide individuals with a variety of different services which are becoming increasingly available online (Schou and Hjelholt, 2018, p. 112–115). The digitization of governmental service includes the need for a safe, portable and easily accessible digital identity. Currently, the only "top-down" implemented use case of a (partly) blockchain-based national identity is Estonia which has established one of the most technologically advanced national ID-card systems. The mandatory card allows access to all secure e-services (Sullivan and Burger, 2017), including travel within the EU, national health insurance, access to bank accounts, e-voting, the administration of medical records, and even tax claims [8].

The physical card is protected with 384-bit ECC public key encryption and can also be used within a digital environment for verification. It utilizes blockchain technology to ensure the validity of the personal information, whilst allowing full control and portability. During a brief period in November 2017 the system could not be used due to a security problem resulting from a design vulnerability of the chip on the card. Estonia's administration reacted quickly, but the incident raised the question how countries with more than ∼1.5 million residents could address such a crisis that entails going back to traditional administrative methods, as well as addressing the security issue by potentially having to replace all cards in use (Asmae, 2017). Nevertheless, the implementation of the system is generally regarded as a success. Although the identity is sometimes deemed "self-sovereign", since the flow of information is fully controlled by the identity owner, there are restrictions in terms of usage. Hence, it could be argued that this system does not represent a pure embodiment of the self-sovereign identity concept and should not be considered as such. Furthermore, it has to be mentioned that this system focusing on Estonian citizens is different from the E-residence scheme of Estonia that

[6]Available online at: https://uidai.gov.in/about-uidai/unique-identification-authority-of-india/about.html (accessed September 13, 2019).
[7]Available online at: https://forus.io (accessed March 4, 2020).
[8]Available online at: https://e-estonia.com/solutions/e-identity/id-card *(accessed March 4, 2020).*

has been described in the governance-focused categorization as an example for individual incentive programs. E-residence is oriented toward foreigners which wish to use Estonia as a business hub. Besides Estonia, such programs seem particularly interesting for countries with a track record in investment migration (Surak, 2016, p. 8–13).

However, the most successful solutions in the public sector aiming at the implementation of decentralization and DLT enabled self-sovereign identity management are based on the Community-based bottom-up model. One example for such a platform is being built in the north of the Netherlands by the foundation Forus[9]. Together with the community of Zuidhorn/Westerkwartier in the province of Groningen a solution to digitize social service provision was developed and implemented from 2017 to March 2018 (Velthuijs, 2018). The main aim of the system is to support children from non-privileged families ("Kindpakket"). During a process of intensive collaboration with the municipality and other partners from the private sector, Forus built an easy to use system for administrative purposes. It allows parents to efficiently and quickly receive funding for children who require financial aid (Van der Beek, 2018). The platform uses DLT and particularly self-sovereign identity as core design elements, as well as ZKP mechanisms to limit the exchange of raw personal data. It is designed to use cryptocurrency (typically Ethereum) as much as possible. However, it can also be used with traditional currency (Euro) when linked to a Dutch bank account. The basic design of interaction entails four roles: the applicant who uses a digital identity application ("Me"), merchants and service providers that register for the platform and also have digital identities, sponsors like public or private institutions that provide funds for which the users can apply, as well as validating parties that confirm credentials if needed. The open-source platform has in the meanwhile also been used for other projects in communities in the Netherlands[10], and continues to be developed adding new features and use cases.[11]

Similar to Forus, Kiva[12] is a decentralized digital identity platform that is being developed and implemented since 2018 in Sierra Leone (Wang and De Filippi, 2020, p. 11-14). Although the system has the support of the central government, it is currently rather limited in reach and is iteratively developed with a small community of users (Wang and De Filippi, 2020, p. 11). Hence, this development can also be considered as a community-based approach that evolves step-by-step. Kiva has the main objective of facilitating the creation of credit history. The system tries to make credit markets accessible for "unbanked people" via the creation of a secure digital identity and stimulates economic development through microloans. It is based on Hyperledger Indy as underlying blockchain layer. While the first implementation steps of Kiva seem promising, current practice is still reliant on guardian- and custodianship. While technically designed to enable identity management along the paradigms of self-sovereign identity with the use of ZKP, lacking technological infrastructure and limited technical knowledge among users highlight that it still will take some time until such solutions can be implemented without friction. At least during this transition period there is a need for concrete governance frameworks that limit and mitigate risks, particularly relating to privacy and individual autonomy (Wang and De Filippi, 2020, p. 13–14).

While these two projects are very promising, it should not be overlooked that many more projects using the Community-based bottom-up approach exist in other parts of the world. Without trying to provide a enumerative list of pilots at the time of writing, initiatives in Austria (Danube Tech), Canada (British Columbia; TheOrgBook), Spain (Alastria), Switzerland (City of Zug; UPort and ti&m), and the United States of America (Illinois Blockchain initiative; Blockchain Bundesverband, 2018, p. 54–56) can be mentioned. Potentially, the relative success of the Community-based bottom-up approach can be explained when considering the aspect that a lot of the difficult questions relating to philosophical and legal aspects of identity are not as pressing on a community level, or are relatively easy to remedy with pragmatic workarounds facilitated by smaller institutional settings.

The digital identity systems described in this section usually use a publicly verifiable digital signature showing governmental approval and are therefore recognized. Hence, these identities could also be used to communicate with other (semi-)public institutions (e.g., hospitals), as well as civil society organizations (CSOs). In this area DLT-based electronic patient records are among the most prominent use cases for pilots. Many organizations such as Medicalchain, MedRec, and MediBloc are focusing on building decentralized record management systems for electronic health records which comply with the regulatory framework (Ekblaw et al., 2016). Nevertheless, medical records are highly sensitive and require utmost care when sharing with other institutions. Still, interoperability and more patient control would be highly desirable as most hospitals work with siloed databases, which can lead to fragmented information regarding the patient and its medical needs. Additionally, inappropriate regulations such as those in the United States have halted the implementation of such solutions (Salzman, 2018).

The digitization of identity management is certainly also appealing in a cross-border context. For example, Dubai partnered with a startup from the United Kingdom named ObjectTech to ensure blockchain-based security measures for the airport. ObjectTech has been working alongside the Dubai Immigration and Visas department to develop digital passports which allow for the elimination of manual checks. At the time of writing it is being stated that Dubai will also introduce a fully developed self-sovereign identity system. However, use cases outside traveling scenarios are not available. The proposed solution is a marriage between a blockchain-based authentication system and biometric verification of an individual (Zhao, 2017). The role of DLT remains limited to the management and validation of the data once it has been registered. Hence, DLT does not aid in ensuring the validity of the submitted data.

[9]https://forus.io (accessed March 1, 2020).

[10]The webshop of the community of Nijmegen with an explanation of the scheme is available via. Available online at: https://nijmegen.forus.io (accessed March 4, 2020).

[11]Available online at: https://github.com/teamforus (accessed March 4, 2020).

[12]Available online at: https://www.kiva.org/protocol (accessed March 1, 2020).

Currently, emerging technologies capable of addressing this issue hinge on the use of biometric identification, which in turn has its own legal and ethical constraints as we know from other fields of application such as surveillance (Jasserand, 2018, p. 154–165). We will further elaborate on this aspect in the discussion.

To conclude this section, there are also cases where governments are unable to provide citizens with humanitarian aid they require due to natural or man-made disasters. As we have already briefly mentioned in the introduction of this article, decentralization, DLT, and self-sovereign identity could potentially remedy the failure of public institutions. Several use cases have already been recorded in the humanitarian domain, where the Word Food Programme and other agencies have successfully implemented digital identity solutions (Zwitter and Boisse-Despiaux, 2018; Wang and De Filippi, 2020, p. 14–16).

In the process of humanitarian intervention and aid, beneficiaries are registered to ensure an organized process of aid delivery. This enables aid organizations to track the amount of people they helped and in what manner, e.g., number of vouchers for providing food or shelter. Corporations such as Tykn and Aid:Tech are prominent examples in this field, as they both have developed and implemented solutions in such environments. Aid:Tech has collaborated with the UN World Food Programme in 2018 and built a wallet with biometric verification for refugees (Juskalian, 2018). With this, highly vulnerable individuals are able to buy items at stores in the camp without requiring a physical wallet (Zambrano et al., 2018). Tykn built a solution in collaboration with the 510-data team of the Dutch Red Cross that has the aim to provide humanitarian aid through data and digital products[13]. Following the destruction by Hurricane Irma in 2017, a pilot was tested on the island of St. Martin to use DLT for to manage disaster relief funds. After setting up a "digital wallet", users would be able to receive digital vouchers, which they could use purchase food, water and other relief goods (Erjula, 2018). Both systems hinge on the self-sovereign identity concept where beneficiaries can manage their own information and have full agency.

Case Studies in the Private Sector

In many ways, digital identity systems using DLT, self-sovereign identity and ZKP can be understood as a response to the commodification of digital identity by powerful private Internet corporations such as Facebook, or Google. After all, the original intention of Blockchain was to get independent of centralized institutions and trust anchors (Nakamoto, 2008, p. 1). DLT-based identity solutions for the private sector are very similar to those described in the previous section, with the focus on leveraging blockchain for (cross-border, cross-system) interoperability, data agency, and potentially also compliance with new regulations such as the EU General Data Protection Regulation (GDPR) (Finck, 2018b). The possibilities for DLT-based identity solutions in the private sector are frequently intertwined with those in the public sector as some require verified credentials by governmental institutions.

When considering concrete applications in the private sector, know-your-customer (KYC) related solutions are a prominent example. Due to insufficient human resources and volume of regulatory change many companies struggle with the escalating costs and complexity in their KYC process. For the year 2016 it was estimated that financial institutions and their corporate customers spent ∼$500 million annually for KYC processes worldwide (Harrop and Mairs, 2016). A partnership between PwC, Onfido and uPort has tried to address this issue with a more efficient solution, potentially used for consumer identities in UK financial services. The ConsenSys-backed uPort showcases the self-sovereign aspect as users can manage the transfer of their own identifiable information, keys and data through their personal device (Wood, 2019).

Another application lies in the area of verification of governmental identity to drive innovations such as e-mobility. In a partnership between Deutsche Telekom, Riddle&Code, Bundesdruckerei and Jolocom a digital identity solution was developed. First, users are verified in person by Bundesdruckerei. Once their identity is confirmed, they can upload their traditional German identity cards into a smart wallet to get access to e-scooters (Habel, 2019). This pilot seems particularly interesting, since such systems could remove the requirement to create a new account when signing up to a new mobility service. However, from the description of the pilot it is not clear whether and how payment could be handled.

A closely related example is age limit control. For example, when age verification is needed (e.g., buying alcoholic beverages, entrance permission to a venue), the age of the client needs to be verified. In many cases this is done by the provider checking the customers identity card. However, this typically means the customer has to share more personal data than required, as identity documents such as driving licenses, identity cards, or passports include much more information than needed for this purpose. Hence, it seems useful to look into the feasibility of ZKP. In a collaboration between Budweiser and Civic at the 2018 Consensus conference they aimed to showcase the value of ZKP for age verification to buy a Budweiser (Capilnean, 2018). For the pilot, Civic provided a digital wallet where the governmental identity document was stored and digitally signed for verification purposes. The wallet is run in a digital device such as a smartphone, which is also used to scan a QR code on a vending machine that provides beer. The vending machine is instructed to only dispense beer to those over the age of 21. While the trial worked as such, some participants had to wait for more than 20 min to be served by the machine. Maybe this might not be untypical in current real-life circumstances when waiting for a beer in a busy bar, but it highlights that ZKPs might be working slower than expected, and are potentially limited in terms of scalability. Nevertheless, it seems not unlikely that the technology will continue to improve and increase efficiency over time.

DISCUSSION

Throughout this piece we are aspiring at describing and framing the changing landscape in the management of digital identity.

[13] Available online at: https://www.510.global (accessed February 22, 2020).

As we have illustrated several of the most promising use cases in the public and private sector, we are now turning to discuss the potential implications of these innovations on the power balance between the individual, government, and corporations. Since this is a subject of immense breadth and where many societal disciplines and fields play a role (Finck, 2018a, p. 182–209), we choose to do so through the application of the lenses of individual and group privacy, as well as considerations on suitable backup-mechanisms for DLT based digital identity systems.

The discussion of privacy is useful, since the topic is a proxy for the division between the sphere of the individual and the sphere of the public (Cannataci, 2017, p. 36–41). We also highlighted this aspect in the section focusing on philosophical perspectives about identity, particularly when considering the constructivist perspective. It must be emphasized however, that privacy should not only be interpreted as a defensive right. As the UN has acknowledged throughout its work on privacy in the digital age, privacy also has an enabling character allowing the individual to develop its views on the world and itself (United Nations, 2016, p. 2). While the concept of self-sovereign identity has many elements that strengthen the individual in its position against governments and corporations, only final products and concrete applications will show whether this promise materializes. For example, when analyzing how responsibilities of controllers could be applied for operating Bitcoin in the regulatory framework of the EU GDPR, there remain uncertainties whether the collective—as partnership—is responsible within the meaning of Article 4 paragraph 7 GDPR, or its individual members are joint controllers under Article 26 GDPR (Buocz et al., 2019, p. 196). In other words, it is impossible to claim individual rights, if it is unclear who precisely has a responsibility of respecting, protecting and promoting them.

Furthermore, potential tensions should be considered on questions such as the private and public nature of data, enforcement of concrete individual rights (e.g., amendment, access, erasure/"right to be forgotten", etc.), data protection by design and default, and other requirements of state of the art data protection law (Finck, 2018b, p. 26–32). While proponents of DLT might question the relevance of high data protection standards for the operation of innovative digital identity systems, since the underlying assumptions of data protection might seem outdated as such to them, it is also fair to remain a believer in the core principles enshrined in regulations such as GDPR. Currently, GDPR and international agreements such as Convention 108+ of the Council of Europe represent two of the few effective safeguards preventing the Facebook, or "Googlization" of Everything (Vaidhyanathan, 2012). This is even more relevant in a time in which whole groups are unaware of the fact that the deployment and use of omnipresent digital technology is significantly affecting—if not eradicating—their opportunity for informational self-determination (Taylor et al., 2017, p. 226–234). As digitization is on the verge of defining what human identity is, and should be worth, these aspects become even more important to address. Hence, if DLT is about to take over identity management in the digital age, for which there are many good practical reasons, and if such identities should be "good identities" enabling a dignified co-existence (World Economic Forum, 2018a, p. 17), the technology also needs to be designed in a way in which classical privacy and data protection safeguards, as well as individual remedies are embedded by default.

However, regardless of how much identity will be digitized, one aspect that probably will have to remain tied to the physical domain is the backup-mechanism for DLT based digital identities. If a device containing a self-sovereign identity gets lost, stolen, or broken, or if the user forgets its access key, there must be a way to restore agency over such elemental information. Many options currently discussed circle around the use of biometrics to generate and potentially restore access to digital wallets, or identity hubs. While biometrics have the advantage that they are relatively persistent and cannot get lost, these characteristics are also the basis for why their omnipresent use can be dangerous.

Once biometrical characteristics of a person are registered, it is possible to paint incredibly detailed pictures of someone's life and interactions. This is often attached to completely unintended consequences, as an example from humanitarian aid shows. In 2019 the UN World Food Programme demanded from Houthi officials in Yemen to allow for the deployment of biometric technologies like iris scans and digital fingerprints to monitor suspected fraud during food distribution (Lontero, 2019). The refusal by the Houthi officials to deploy biometrical recognition over surveillance concerns lead to the cancellation of the aid efforts. This development was met with criticism on the initial demand, claiming that the UN's action was disproportionate and resulting in harm for the weakest (Martin and Taylor, 2019). However, this incident is not the only indication that the widespread, pervasive, and under-considered use of biometrics for identification purposes results in negative outcomes.

The Indian Aadhaar system hinges heavily on the use of biometrics. While the use of Aadhaar by private corporations has been limited by the Indian Supreme Court in a high-profile judgment from 26 September 2018 (Indian Supreme Court, 2018), the appropriateness of the dependency on biometrics was acknowledged as such. The judges came to the conclusion that the information security regarding the management of the biometric data of more than 1.2 billion people stored in a centralized system can be guaranteed by the government. Without speculating about the threat of the government itself abusing this power, only time will tell if the finding is true and whether the data can be kept safe from attackers. As stated in the dissenting judgment from Justice Chandrachud: "The invisible threads of a society networked on biometric data have grave portents for the future. Unless the law mandates an effective data protection framework, the quest for liberty and dignity would be as ephemeral as the wind." (Indian Supreme Court, 2018, p. 337) Therefore, the use of biometrics as backup mechanism for digital identities requires at least well thought through and detailed oversight and review procedures, coupled with the possibility to demand review of decisions and management practice on the basis of individual

request. Ultimately however, it might be desirable to consider other backup-mechanisms which are safer and have less potential for undesirable and dangerous side-effects. The use of biometric data might be part of the solution to this problem, but not the solution as such (Wang and De Filippi, 2020, p. 8–9).

Furthermore, a crucial question of power balance remains. Even if the individual might have full control over his/her credentials and the information contained in a decentralized identity management system, the control over the network and its design remain in the hands of those developing and maintaining the underpinning technological infrastructure. Furthermore, the choice of whether or not to have a digital identity in the future will equal the choice of whether or not to use applications like Facebook or WeChat, that have become omnipresent quasi-standards in many societies. The same network effect that makes it convenient to use these tools creates social pressure, particularly for those who refuse to use them. It should not be taken for granted that advanced digital identities fix these issues. They might as well enable an era of "neo-feudalism" and increased social division, especially if their implementation is done naïvely purely focusing on questions of technical feasibility (Gstrein and Kochenov, 2020).

Bringing the empirical cases together with the theoretical elaborations yields quite interesting results. On the meta-level, it becomes clear that the practice of digital identity management and the theoretical conceptions of uniqueness, relationality and legal determinacy remain relatively disconnected. The concept of legal determinacy is best represented as both private and public sector projects in general are aware and try to abide by legal frameworks such as the GDPR. Almost exclusively, however, most projects focus on the individual as a bearer of identity rights; relational aspects of identity and the problems that will emerge around data protection and data ownership in these cases seem to have no priority for stakeholders in both sectors.

CONCLUSION

The emergence of DLT and the abundance of private and public sector initiatives, as well as the emergent debate around digital identity, make it profusely clear how important it is to gain a profound understanding of the legal and philosophical conceptions and norms that govern identity in general. A sound framework of digital identity management needs to take into account questions of privacy, relationality of identity data, and data ownership. Ultimately, specifically the direct relation between the philosophical conception of identity and its socio-legal foundations, as elaborated in this article, can serve as a foundation toward defining the "self-sovereign" individual with its rights, obligations and limitations.

At the same time, a digital identity management framework is not pre-determined by certain ideas around DLT such as decentralization and immutability. If anything, DLT has enriched the governance toolkit. Public and private sector actors can select among a range of top-down to bottom-up management approaches. Even if decentralization is "en vogue" at the moment in both, the governance debate as well as amongst blockchain advocates, it is by no means a panacea for all old ailments. DLT can be a part of useful solutions, but only if it can incorporate socio-legal and philosophical necessities that digital identity brings with it. Once translated well into practice, DLT has the capacity to strengthen the rights of the individual by providing access to tools that enhance the individual's agency as self-sovereign actor.

AUTHOR CONTRIBUTIONS

AZ, OG, and EY contributed to the conception and design of the study and wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

REFERENCES

Allen, C. (2016). *The Path to Self-Sovereign Identity. Life With Alacrity April 26, 2016.* Available online at: http://www.lifewithalacrity.com/2016/04/the-path-to-self-sovereign-identity.html (accessed February 5, 2020).

Asmae, K. (2017). *Estonia's ID Card Crisis: How E-State's Poster Child Got Into and Out of Trouble.* ZDNet November 13, 2017. Available online at: https://www.zdnet.com/article/estonias-id-card-scrisis-how-e-states-poster-child-got-into-and-out-of-trouble/ (accessed September 13, 2019).

Blitz, B. K., and Sawyer, C. (2011). *Stateless in the European Union*. Cambridge, UK: Cambridge University Press.

Blockchain Bundesverband (2018). *Self-Sovereign Identity: A Position Paper on Blockchain Enabled Identity and the Road Ahead.* Available online at: https://bundesblock.de/groups/digital-identity/blog/new-position-paper-self-sovereign-identity-defined (accessed September 13, 2019).

Brozek, B. (2017). "Legal personhood: animals, artificial intelligence and the unborn," in *The Troublesome 'Person'*, eds A. J. Visa, T. P. Kurki (London, UK: Springer), 2017. doi: 10.1007/978-3-319-53462-6_1

Buocz, T., Ehrke-Rabel, T., Hödl, E., Eisenberger, I. (2019). Bitcoin and the GDPR: allocating responsibility in distributed networks. *Comput. Law Secur. Rev.* 35, 182–198. doi: 10.1016/j.clsr.2018.12.003

Cameron, K. (2005). The laws of identity. *Kim Cameron's Identity Weblog (blog),* May 11, 2005. Available online at: https://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf (accessed September 13, 2019).

Cannataci, A. (2017). "Data protection and privacy under pressure transatlantic tensions, EU surveillance, and big data," in *Games People Play: Unvarnished Insights About Privacy at the Global Level*, eds. G. Vermeulen, E. Lievens, (Antwerp: Maklu Uitgevers), 13–47.

Capilnean, T. (2018). First ever anonymous age verifying beer vending machine in partnership with anheuser-busch inbev. *Civic Blog May 12, 2018.* Available at: https://www.civic.com/blog/first-ever-anonymous-age-verifying-beer-vending-machine-in-partnership-with-anheuser-busch-inbev/ (accessed July 19, 2019).

Council of Europe (2018). *Cybercrime Convention Committee (T-CY), Preparation of a 2nd Additional Protocol to the Budapest Convention on Cybercrime, Conditions for Obtaining Subscriber Information.* Available at: https://rm.coe.int/t-cy-2018-26-ip-addresses-v6/16808ea472 (accessed September 13, 2019).

ECJ. (2018). *Patrick Breyer v Bundesrepublik Deutschland, Case C-582/14.* EU:C:2016:779.

Ekblaw, A., Azaria, A., Halamka, J. D., Lippman, A. (2016). *A Case Study for Blockchain in Healthcare: "MedRec" Prototype for Electronic Health Records and Medical Research Data. Whitepaper.* Available online

at: https://www.healthit.gov/sites/default/files/5-56-onc_blockchainchallenge_mitwhitepaper.pdf (accessed September 13, 2019).

Erjula, T. (2018). *510 is Exploring the Use of Blockchain In Humanitarian Aid With* Tykn. tech. 510 x Tykn Press. Available online at: https://tykn.tech/510-x-tykn-press-release/ (accessed September 13, 2019).

Finck, M. (2018a). *Blockchain Regulation and Governance in Europe*. Cambridge, UK: Cambridge University Press. doi: 10.1017/9781108609708

Finck, M. (2018b). Blockchains and data protection in the European union. *Eur. Data Protect. Law* 2018, 17–35. doi: 10.21552/edpl/2018/1/6

Foster, C., and Herring, J. (2017). "Summary and Conclusion," in *Identity, Personhood and the Law, SpringerBriefs in Law* (Champ: Springer), 57–70. doi: 10.1007/978-3-319-53459-6_5

Galič, M., Timan, T., and Koops, B. J. (2017). Bentham, deleuze and beyond: an overview of surveillance theories from the panopticon to participation, philos. *Technology* 30, 9–37. doi: 10.1007/s13347-016-0219-1

Gstrein, O. J., and Kochenov, D. (2020). Digital identity and distributed ledger technology: paving the way to a neo-feudal brave new world? *Front. Blockchain*. 3, 1–8. doi: 10.3389/fbloc.2020.00010

Gstrein, O. J., and Ritsema van Eck, G. (2018). Mobile devices as stigmatizing security sensors: the GDPR and a future of crowdsourced 'broken windows'. *Int. Data Privacy Law* 8, 69–85. doi: 10.1093/idpl/ipx024

Habel, P. (2019). *Xride: First-of-Its-Kind, Blockchain-Based E-Mobility Project*. Telekom Media. Available online at: https://www.telekom.com/en/media/media-information/archive/xride-first-of-its-kind-blockchain-based-e-mobility-project-580934 (accessed March 1, 2020).

Harbinja, E. (2017). Post-mortem privacy 2.0: theory, law, and technology. *Int. Rev. Law Comput. Technol.* 31, 26–42. doi: 10.1080/13600869.2017.1275116

Harrop, M. D., and Mairs, B. (2016). *Thomson Reuters 2016 Know Your Customer Surveys Reveal Escalating Costs and Complexity*. Press Release May 9, 2016. Available online at: https://www.thomsonreuters.com/en/press-releases/2016/may/thomson-reuters-2016-know-your-customer-surveys.html (accessed July 18, 2019).

Indian Supreme Court (2018). *Justice K.S. Puttaswamy (Retd) vs Union of India (2017)*. Writ Petition (Civil) W.P. (C) No.-000494-000494/2012.

Jasserand, C. (2018). Law enforcement access to personal data originally collected by private parties: missing data subjects' safeguards in directive 2016/680? *Comput. Secur. Rev.* 34, 154–165. doi: 10.1016/j.clsr.2017.08.002

Juskalian, R. (2018). Inside the Jordan refugee camp that runs on blockchain. *MIT Technology Review April 12, 2018*. Available online at: https://www.technologyreview.com/s/610806/inside-the-jordan-refugee-camp-that-runs-on-blockchain/ (accessed September 13, 2019).

Kochenov, D. (2009). Ius tractum of many faces: European citizenship and the difficult relationship between status and rights. *Colum. J. Eur. Law* 15, 169–237.

Kucklick, C. (2014). *Die Granulare Gesellschaft, Der Granulare Mensch Oder Wie Wir Uns Neu Erfinden*. Berlin: Ullstein.

Le Bras, T. (2015). *Online Overload - It's Worse Than You Thought*. Dashlane Blog July 21, 2015. Available online at: https://blog.dashlane.com/infographic-online-overload-its-worse-than-you-thought/ (accessed September 13, 2019).

Lessig, L. (1996). *Code ver 2.0*. New York, NY: Basic Books.

Lichtenthähler, S. (2019). Intersubjektive konstitution des selbstbewusstseins? *Arch. für Rechts- und Sozialphilosophie* 105, 104–123. doi: 10.25162/arsp-2019-0006

Lontero, M. (2019). *Stop Surveillance Humanitarianism*. New York Times July 11, 2019. Available online at: https://www.nytimes.com/2019/07/11/opinion/data-humanitarian-aid.html (accessed July 19, 2019).

Luchte, J. (2012). "Pythagoras and transmigration: wandering souls," in *Epilog: The Fate of the Doctrine of Transmigration,* ed. J. Luchte (New York, NY: Bloomsbury Publishing Plc), 169–173.

Martin, A., and Taylor, L. (2019). *Biometric Ultimata - What the Yemen Conflict Can Tell Us About the Politics of Digital ID Systems*. Global Data Justice June 21 2019. Available online at: https://globaldatajustice.org/2019-06-21-biometrics-WFP/ (accessed July 19, 2019).

Morsink, J. (2012). *Inherent Human Rights: Philosophical Roots of the Universal Declaration*, Philadelphia, PE: University of Pennsylvania Press.

Mutua, M. (2016). *Human Rights Standards: Hegemony, Law, and Politics*, Albany: SUNY Press.

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. Bitcoin.org. Available at: https://bitcoin.org/bitcoin.pdf (accessed September 13, 2019).

Nimfuehr, M. (2018). *Blockchain Application Land Register: Georgia and Sweden Leading*. Medium, August 18, 2018. Available online at: https://medium.com/bitcoinblase/blockchain-application-land-register-georgia-and-sweden-leading-e7fa9800170c (accessed September 13, 2019).

Papineau, D. (2016). *Naturalism. The Stanford Encyclopedia of Philosophy*. Winter 2016 Edition. Available online at: https://plato.stanford.edu/archives/win2016/entries/naturalism/ (accessed September 13, 2019).

Poleshchuk, V. (2016). *'Making Estonia Bigger': What E-Residency in E-Estonia Can Do for You, What It Can Do for Estonia*. IMC Policy Briefs 2016/1. Available at: https://investmentmigration.org/download/making-estonia-bigger-e-residency-e-estonia-can-can-estonia/ (accessed September 13, 2019).

Privacy International (2018). *Initial Analysis of Indian Supreme Court Decision on Aadhaar*. Privacy International September 26, 2018. Available online at: https://privacyinternational.org/long-read/2299/initial-analysis-indian-supreme-court-decision-aadhaar (accessed September 13, 2019).

Rao, U. (2019). Population meets database: aligning personal, documentary and digital identity in aadhaar-enabled India. *South Asia J. South Asian Stud.* 42, 537–553. doi: 10.1080/00856401.2019.1594065

Rueb, E. S. (2019). *Your Instagram Feed Is About to Have More Ads From Influencers*. New York Times June 4, 2019. Available online at: https://www.nytimes.com/2019/06/04/technology/instagram-ads-influencers.html (accessed September 13, 2019).

Salzman, S. (2018). *Electronic Medical Records: Holy Grail for Blockchain*. MedPage Today August 22, 2018. Available online at: https://www.medpagetoday.com/practicemanagement/informationtechnology/74695 (accessed July 17, 2019).

Sartre, J. P. (2007). *Das Sein und das Nichts - Versuch einer phänomenologischen Ontologie, 13th Edn.* ed. T. König (Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag), 412–413.

Schou, J., and Hjelholt, M. (2018). *Digitalization and Public Sector Transformations*. Cham: Palgrave Macmillan. doi: 10.1007/978-3-319-76291-3_6

Sullivan, C., and Burger, E. (2017). E-residency and blockchain. *Comput. Law Secur. Rev.* 33, 470–481. doi: 10.1016/j.clsr.2017.03.016

Surak, K. (2016). *Global Citizenship 2.0 - the Growth of Citizenship by Investment Programs*. IMC-RP 2016/3. Available online at: https://investmentmigration.org/download/global-citizenship-2-0-growth-citizenship-investment-programs/ (accessed September 13, 2019).

Sweney, M. (2019a). *BA Faces £183m Fine Over Passenger Data Breach*. The Guardian July 8, 2019. Available online at: https://www.theguardian.com/business/2019/jul/08/ba-fine-customer-data-breach-british-airways (accessed September 13, 2019).

Sweney, M. (2019b). *Marriott to be Fined Nearly £100m Over GDPR Breach*. The Guardian July 9, 2019. Available online at: https://www.theguardian.com/business/2019/jul/09/marriott-fined-over-gdpr-breach-ico (accessed September 13, 2019).

Taylor, L., van der Sloot, B., Floridi, L. (2017). Conclusion: what do we know about group privacy? *Group Privacy, Philosoph. Studies Series* 126, 225–237. doi: 10.1007/978-3-319-46608-8_12

Tjong Tijn Tai, E. (2018). Data ownership and consumer protection. *J. Eur. Consum. Market Law* 7, 136–140. doi: 10.2139/ssrn.3172725

United Nations (2014). *General Assembly, The Right to Privacy in the Digital Age*. Resolution 68/167, January 21, 2014.

United Nations (2016). *General Assembly, Resolution A/HRC/32/L.20*. June 27, 2016.

United Nations (2018). *General Assembly, The Right to Privacy in the Digital Age*. A/C.3/73/L.49/Rev.1, November 14, 2018.

Vaidhyanathan, S. (2012). *The Googlization of Everything*. Oakland, CA: University of California Press.

Van der Beek, P. (2018). *Blockchain Kindpakket Zuidhorn Wint Prijs*. Computable March 30 2018. Available online at: https://www.computable.nl/artikel/nieuws/digital-transformation/6329958/250449/blockchain-kindpakket-zuidhorn-wint-prijs.html (accessed September 30, 2019).

Velthuijs, M. (2018). *Platform Forus Richt Gemeentelijke Dienstverlening Echt Anders in*. SBIR Gegevenslandschap eindrapportage FASE I. Available online at: https://www.berenschot.nl/publish/pages/6150/sb1bg17020_openbare_samenvatting_2.pdf (accessed September 13, 2019).

Verborgh, R. (2019). *Re-Decentralizing the Web, for Good This Time*. Weblog January 11, 2019. Available at: https://ruben.verborgh.org/articles/redecentralizing-the-web/ (accessed September 13, 2019).

Wagner, K., Némethi, B., Renieris, E., Lang, P., Brunet, E., and Holst, E. (2018). *'Self-Sovereign Identity' Position Paper*. Blockchain Bundesverband. Available online at: https://www.bundesblock.de/wp-content/uploads/2018/10/ssi-paper.pdf (accessed August 7, 2019).

Wall, D. S. (2013). Policing identity crimes. *Polic. Soc.* 23, 437–460. doi: 10.1080/10439463.2013.780224

Wang, F., and De Filippi, P. (2020). Self-sovereign identity in a globablized world: credentials-based identity systems as a driver for economic inclusion. *Front. Blockchain* 28, 26–42. doi: 10.3389/fbloc.2019.00028

Wood, M. (2019). *PwC, Onfido Join Blockchain Identity Platform uPort September, 2019*. Available at: https://www.ledgerinsights.com/pwc-onfido-blockchain-identity-platform-uport/ (accessed February 22, 2020).

World Bank Group (2018). *ID4D Annual Report*. Available at: https://id4d.worldbank.org/sites/id4d.worldbank.org/files/2018_ID4D_Annual_Report.pdf (accessed September 13, 2019).

World Economic Forum (2018a). *Our Shared Digital Future - Building an Inclusive, Trustworthy and Sustainable Digital Society*. Available at: https://www.weforum.org/reports/our-shared-digital-future-building-an-inclusive-trustworthy-and-sustainable-digital-society (accessed September 13, 2019).

World Economic Forum (2018b). *Identity in a Digital World - A New Chapter in the Social Contract*. Accessible at: http://www3.weforum.org/docs/WEF_INSIGHT_REPORT_Digital%20Identity.pdf (accessed September 13, 2019).

Zambrano, R., Young, A., and Verhulst, S. (2018). *Connecting Refugees to Aid Through Blockchain Enabled Id Management: World Food Programme's Building Blocks*. GovLab October 2018. Available online at: https://blockchan.ge/blockchange-resource-provision.pdf (accessed September 13, 2019).

Zhao, W. (2017). *Dubai Plans Digital Passports Using Blockchain Tech*. Coindesk June 9, 2017. Available online at: https://www.coindesk.com/dubai-plans-gate-less-airport-security-using-blockchain-tech (accessed September 13, 2019).

Zuboff, S. (2019). Surveillance capitalism and the challenge of collective action. *New Labor Forum* 28, 10–29. doi: 10.1177/1095796018819461

Zwitter, A., and Boisse-Despiaux, M. (2018). Blockchain for humanitarian action and development aid. *J. Int. Humanit. Action* 3:16. doi: 10.1186/s41018-018-0044-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Blockchain

ORIGINAL RESEARCH
published: 20 May 2021
doi: 10.3389/fbloc.2021.627641

The Formal, Financial and Fraught Route to Global Digital Identity Governance

Malcolm Campbell-Verduyn[1,2†]* *and Moritz Hütten*[3,4†]

[1]*Käte Hamburger Kolleg Centre for Global Cooperation Research University of Duisburg-Essen, Duisburg, Germany,* [2]*Department of International Relations and International Organization, University of Groningen, Groningen, Netherlands,* [3]*Department of Economics and Center for Sustainable Economic and Corporate Policy, Darmstadt University of Applied Sciences, Darmstadt, Germany,* [4]*Department of Political Science, Amsterdam Institute for Social Science Research, University of Amsterdam, Amsterdam, Netherlands*

How can we understand the progressive, piecemeal emergence of global digital identity governance? Examining the activities of the Financial Action Task Force (FATF) - an intergovernmental organization at the center of global anti-money laundering and counter-the-financing of terrorism governance-this paper advances a two-fold argument. First, the FATF shapes how, where and who is involved in developing key standards of acceptability underpinning digital identity governance in blockchain activities. While not itself *directly* involved in the actual coding of blockchain protocols, the FATF influences the location and type of centralized modes of control over digital identity governance. Drawing on the notion of protocological control from media studies, we illustrate how centralized control emerging in global digital identity governance emanates from the global governance *of* financial flows long considered by international organizations like the FATF. Second, we suggest that governance *by* blockchains persistently shapes the ability of the FATF to stem illicit international financial flows. In highlighting both the influence of FATF on blockchain governance and blockchain governance on the FATF, we draw together two strands of literature that have been considered separately in an analysis of the formal, financial and fraught route to global digital identity governance.

Keywords: control, identity, finance, money laundering and financing terrorism, financial action task force, global governance

OPEN ACCESS

Edited by:
Andrej Zwitter,
University of Groningen, Netherlands

Reviewed by:
Jolien Ubacht,
Delft University of Technology, Netherlands
Zeynep Gurguc,
Imperial College Business School, United Kingdom

***Correspondence:**
Malcolm Campbell-Verduyn
campbell@gcr21.uni-due.de

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Blockchain for Good, a section of the journal Frontiers in Blockchain

Received: *09 November 2020*
Accepted: *04 May 2021*
Published: *20 May 2021*

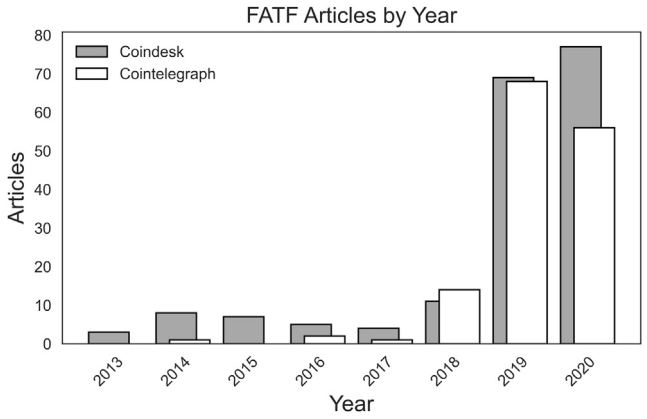
Citation:
Campbell-Verduyn M and Hütten M (2021) The Formal, Financial and Fraught Route to Global Digital Identity Governance. Front. Blockchain 4:627641. doi: 10.3389/fbloc.2021.627641

INTRODUCTION

How can we understand the on-going emergence of global digital identity governance? The seemingly ever progressing digitalization of human activities, accelerated by the Covid-19 pandemic, is not a smooth, linear and all-encompassing affair. Rather, it remains patchy and tension-filled. While activities like digital payments flourish (Boakye-Adjei, 2020; Frazier, 2020), others remain marked by longstanding conflicts. The progressive and piecemeal digitalization of identities exemplifies these broad tensions including, amongst others, user privacy and the informational needs of regulators charged with prevent exploitation, abuse and illicit activities. Blockchains and other novel technologies are continually emerging to square the circle of privacy and surveillance. Yet, their applications often merely shift the location and form of such tensions, rather than resolving them.

Analysis of emerging blockchain-based attempts to resolve these and other longstanding tensions in contemporary governance generally considers governance *by* and *of* blockchain systems

TABLE 1 | Mentions of the FATF in leading industry news outlets.



Source: Authors based on articles collected from cointelegraph.com and coindesk.com.

(Campbell-Verduyn, 2018b; Atzori, 2017; de Filipi, 2018; Herian, 2018; Hooper and Holtbrügge, 2020; Jones, 2019; Reijers et al., 2016).[1] The former stress how blockchain applications *themselves* govern an organization or process while the latter emphasize how blockchains are governed by a range of state and non-state organizations. While generating increasingly nuanced understanding, this growing literature has granted surprisingly little attention to the interplay between governance *of* and *by* blockchains. In particular, little attention has been granted to relations between informal and formal forms of blockchain governance beyond passing mentions to the likes of the International Monetary Fund (IMF) and Organization for Economic Cooperation and Development (OECD).

This article contributes to the filling of this gap by examining relations between evolving forms of governance *by* blockchains and the governance *of* blockchain emanating from the Financial Action Task Force (FATF). Attracting more industry attention than in academic studies of blockchain (**Table 1**),[2] this Paris-based intergovernmental organization is responsible for setting global standards for anti-money laundering and counter-the-financing of terrorism governance (AML/CFT). In tracing both 1) the underappreciated role of this formal organization in shaping the emergence of global digital identity governance and 2) the implications of blockchain activities on its attempts to stem illicit financial flows, this paper draws together of analysis of governance *by* and *of* blockchain. To do so, we harness and extend the notion of protocological control. Developed by media studies scholar Alexander Galloway (2004): 6–7, who built on insights from French philosophers Michel Foucault and Gilles Deleuze, the notion of protocological control helps illustrate how the embedding of specific standards of behaviour into computing protocols provide the key "standards governing the implementation of specific technologies." We show how protocols serve as key forms of governance themselves while also drawing out how the location and form of protocological control is itself shaped. In other words, we clarify the *who*, *where* and *how* of protocological control by pointing to the influence of the FATF on the location and form of protocological control in blockchain-based activities. In doing so, we show that, despite claiming to *distribute* power across the nodes in novel digital networks, applications of blockchains instead *frequently shift the location and type of actors exercising centralized contro*l.

Two central contributions are made in this article. First, we illustrate how the FATF shapes how, where and who is involved in developing key standards of acceptability underpinning digital identities. While not itself *directly* involved in the actual coding of protocols, the FATF influences *the location and type of centralized modes of control*. We stress how protocological control emerging in global digital identity governance emanates from the global governance *of* financial flows long considered by intergovernmental organizations like the FATF. In elaborating the role of this organization, we extend studies identifying the financial roots of digital identity governance beyond informal interactions between the public sector and financial technology industry at the national level (Eaton et al., 2018; Faria, 2021). Second, we suggest that governance *by* blockchains persistently shapes the ability of the FATF to stem illicit financial flows. In highlighting tensions between both the influence of FATF on blockchain governance and blockchain governance on the FATF, we draw two strands literature together in identifying both the formal and financial, as well as the fraught route to global digital identity governance.

We elaborate these arguments over three further sections drawing on primary documents, including guidance and reports of the FATF,[3] as well as secondary documents from

[1]For up-to-date overview of this fast growing literature see https://www.blockchainresearchnetwork.org/docs/blockchain-governance/.

[2]Exceptions include Campbell-Verduyn, 2018a; Naheem, 2019; Pavlidis, 2020.

[3]We manually extracted 17 documents as belonging to the topic "blockchain" from the official website of the FATF spanning the years 2013 to 2020.

TABLE 2 | Overview of the FATF 40 + 9 Recommendations (as of February 15, 2021).

Number	
	A—AML/CFT POLICIES AND COORDINATION
1	Assessing risks and applying a risk-based approach
2	National cooperation and coordination
	B—MONEY LAUNDERING AND CONFISCATION
3	Money laundering offence
4	Confiscation and provisional measure
	C—TERRORIST FINANCING AND FINANCING OF PROLIFERATION
5	Terrorist financing offence
6	Targeted financial sanctions related to terrorism and terrorist financing
7	Targeted financial sanctions related to proliferation
8	Non-profit organisations
	D—PREVENTIVE MEASURES
9	Financial institution secrecy laws
	Customer due diligence and record keeping
10	Customer due diligence
11	Record keeping
	Additional measures for specific customers and activities
12	Politically exposed persons
13	Correspondent banking
14	Money or value transfer services
15	New technologies
16	Wire transfers
	Reliance, Controls and Financial Groups
17	Reliance on third parties
18	Internal controls and foreign branches and subsidiaries
19	Higher-risk countries
	Reporting of suspicious transactions
20	Reporting of suspicious transactions
21	Tipping-off and confidentiality
	Designated non-financial Businesses and Professions (DNFBPs)
22	DNFBPs: Customer due diligence
23	DNFBPs: Other measures
	E—TRANSPARENCY AND BENEFICIAL OWNERSHIP OF LEGAL PERSONS AND ARRANGEMENT
24	Transparency and beneficial ownership of legal persons
25	Transparency and beneficial ownership of legal arrangements
	F—POWERS AND RESPONSIBILITIES OF COMPETENT AUTHORITIES AND OTHER INSTITUTIONAL MEASURE
	Regulation and Supervision
26	Regulation and supervision of financial institutions
27	Powers of supervisors
28	Regulation and supervision of DNFBPs
	Operational and Law Enforcement
29	Financial intelligence units
30	Responsibilities of law enforcement and investigative authorities
31	Powers of law enforcement and investigative authorities
32	Cash couriers
	General Requirements
33	Statistics
34	Guidance and feedback
	Sanctions
35	Sanctions
	G—INTERNATIONAL COOPERATION
36	International instruments
37	Mutual legal assistance
38	Mutual legal assistance: freezing and confiscation
39	Extradition
40	Other forms of international cooperation

Source: Adapted from FATF (2019).

blockchain industry news sites. The following section analyzes how the FATF shapes the exercise of protocological control in regards to governance *of* blockchains generally and digital identity governance specifically. A third section highlights how forms of governance *by* blockchains shaped by the FATF paradoxically undermine the objectives of this organization of reducing illicit international financial flows. A final section summarizes and offers directions for future research.

HOW SOFT INTERNATIONAL LAW SHAPES THE *LOCATION* OF HARD CODE

The FATF was established in 1989 as part of inter-state efforts by the Group of 7 (G7) countries to stem global money laundering. It promulgated an initial 40 recommendations for supporting global anti-money laundering efforts (AML) that were supplemented with 9 special counter-the-financing-of-terrorism (CFT) recommendations following the 11 September 2001 attacks (**Table 2**). The task force issues official reports and guidance on the implementation of these 40 + 9 recommendations for countering the financing of the proliferation of weapons of mass destruction (FATF, 2018) and illicit wildlife trade (FATF, 2020c), as well as extending its recommendations to virtual currencies (FATF, 2015), virtual asset service providers (FATF, 2019) and digital identities (FATF, 2020e).

Scholarly literature has long debated the origins and impacts of FATF's activities (Gutterman and Roberge, 2019: 462; Tsingou, 2010; Hülsse, 2008; Hülsse and Kerwer, 2007; Truman and Reuter, 2004). On the one hand, are critiques of its symbolic "security theatre" as providing weak attempts to show member states that it is "doing something" about international money laundering and the financing of terrorism. On the other hand, FATF activities are regarded as successfully motivating a range of state and non-state actors to prioritize AML/CFT efforts while setting the requirements for the proper monitoring of identity systems in finance. These latter accounts stress how enforcement of the FATF's non-binding, voluntary standards relies on periodic monitoring of compliance with its 40 + 9 recommendations. When what the FATF calls "strategic deficiencies in their regimes to counter money laundering, terrorist financing, and financing of proliferation" is identified, the Task Force enhances its monitoring.[4] However, it lacks direct enforcement mechanisms itself. Instead, the FATF issues warnings to exercise caution to its global network of 39 official state members, as well as non-members in its wider network of some 170 associate and observer members, and related regional bodies around the world. These warnings caution state and non-state actors globally about interacting with "Jurisdictions under Increased Monitoring" (the FATF's unofficial "grey list")[5] and countries on its unofficial FATF "blacklists"[6] The effectiveness of the FATF ultimately relies on peer pressure for its members and non-members alike to sanction jurisdictions on its unofficial lists. The FATF's power is thus *indirect*: it is a standard-setter and monitor rather than an enforcer. It *shapes* global regulatory responses, but it relies on others to develop and enforce them, including its member states, who have in turn tended to "deputize" banks and other financial market actors as AML/CFT enforcers to develop and undertake Know Your Customer (KYC) procedures (Amicelle, 2011; see more generally; Avant, 2005). Such enforcement-by-proxy entails a chain of enforcement in which the FATF relies on member states who in turn rely on market actors in their jurisdictions to implement the intergovernmental organization's guidance.

In this section, we build on and extend insights into the FATF's exercise of indirect power. We show how this IO shapes the location of protocological control by tracing the financial and formal lineage of global digital identity governance. The FATF, we argue, shapes the hard code of the computer protocols underpinning blockchain-based activities through soft international law issuance of guidance and recommendations. A first sub-section considers the impacts of the FATF, 2015 guidance on virtual currencies before a second examines the 2019 guidance on virtual asset service providers. Both of these "risk-based" guidances, we argue, shaped the *market-based* location of protocological control over blockchain technology. The FATF enabled private actors to take charge of monitoring the flow of blockchain transactions and the identities of the entities undertaking them. It has done so by guiding member-states towards letting market actors exercise protocological control in the emerging governance of digital identities. Although becoming more explicit, this steering towards private-led governance is in line with this IO's longerstanding risk-based approach that, as we elaborate, attempts to weigh the costs and benefits of greater public involvement in rapidly evolving technological changes. It is also in line with the wider approach towards innovation and the knowledge economy promoted by leading international organizations like the OECD, at whose Paris headquarters the FATF's secretariat is housed (Hasselbalch, 2018; Campbell-Verduyn and Hütten, 2019).

Guiding Protocological Control by Market Forces

While always a consideration in AML/CFT discussions (see for instance FATF, 2013), technology came to the forefront of FATF activities in the past half decade as financial technologies ("FinTech") and regulatory technologies ("RegTech") gained attention globally. The FATF launched a FinTech/RegTech Forum in 2017 for stimulating more effective monitoring and compliance with its 40 + 9 recommendations. The FATF's engagement with blockchain applications began earlier, with a 2014 report weighing the potential benefits and risks from virtual currencies, which included cryptocurrencies based on blockchain technology. The report specifically highlighted identity topics. On the one hand, the FATF (2014) flagged concerns about the anonymity provided by the technology, the limited possibilities for identification and verification of network participants, as well as a lack of clarity for formal regulatory responsibilities. On the other hand, the FATF (2014) also identified legitimate benefits such as lower transaction costs and possibilities for enhancing financial inclusion. Based on this initial risk assessment formal guidance on how its members should apply its 40 + 9 AML/CFT recommendations to virtual currencies was issued in 2015.[7]

[4]http://www.fatf-gafi.org/publications/high-risk-and-other-monitored-jurisdictions/.

[5]See for example https://www.fatf-gafi.org/publications/high-risk-and-other-monitored-jurisdictions/documents/increased-monitoring-june-2020.html.

[6]http://www.fatf-gafi.org/countries/d-i/iran/documents/call-for-action-june-2020.html.

[7]https://www.fatf-gafi.org/fintech-regtech/fatfonfintechregtech/?hf=10&b=0&s=desc(fatf_releasedate).

The 2015 FATF guidance shaped the location and form of protocological control across emerging blockchain-based activities in two interrelated ways. First, it downplayed the growing calls for *public* authorities to apply *direct* control. Instead, it recommended letting market actors develop appropriate protocols for ensuring AML/CFT controls. This recommendation emerged at a time when formal laws were emerging to restrict blockchain-based activities in member countries like Russia and China. As outright bans on the leading application of the technology to cryptocurrencies were being imposed in prominent jurisdictions, the FATF called for looser, "light touch" regulations. It waded off calls for strong "hands on" public control by not recommending formal regulation of blockchain applications. This was despite the growing gulf between AML/CFT identity requirements and the quasi-anonymity of many cryptocurrencies. The 2015 FATF guidance did call for close monitoring of cryptocurrency exchanges by member states. Yet FATF members and non-members alike were also encouraged to *avoid* formal bans and other actions that could lead blockchain activities shifting to less regulated jurisdictions. The guidance instead called for members to "take into account, among other things, the impact a prohibition would have on the local and global level of ML/TF risks, including whether prohibiting VC [virtual currency] payments activities could drive them underground, where they will continue to operate without AML/CFT controls or oversight" (FATF, 2015: 8–9).

In toning down the growing chorus of calls for "stricter" state regulation of blockchain-based activities, the 2015 FATF guidance on virtual currencies re-enforced the longstanding roles of private actors as enforcers of AML/CFT specifically, and the market-based location of protocological control generally. The 2015 FATF guidance extended the private-led development of standards of information communication in and between blockchain activities. Rather than state authorities, a range of competing start-up technology firms like Mastercoin, Counterparty and Interledger proposed manners of connecting together various protocols building on the Bitcoin protocol. Protocological control was equally left to market actors in governing the kinds of "forks" from the Bitcoin blockchain. The market-based competition led to what was dubbed a civil war in the 2017 "hard fork" of the original computer protocol that became Bitcoin Core (BTC) and Bitcoin Cash (BCH) (Coin Idol, 2019). The latter maintained the features and transaction history of the former protocol, while also introducing a fundamental change in acceptable standards of behavior: the ability to spin-off new protocols or "forks" from an existing protocol. The new BCH protocol then itself split in two as debates over the appropriate block size for recording verified transactions on the shared ledger led to the creation of both Bitcoin Cash Satoshi Version (SV) and Bitcoin Cash Adjustable Blocksize Cap (ABC) in late 2018. Whereas the development of these multiple, overlapping protocols was left to market forces, protocological control in the Ethereum blockchain was centralized in its Foundation and founder, Vitalik Buterin. A major flaw in the protocol of The DAO, a utopian experiment with automatic management of crowdsourced funds, led to a hack and withdrawal of the equivalent to $120 million raised in 2016 before informal centralized control was exercised to repair the underlying code (Hütten, 2019). A year later, the centralized group of "core" Ethereum developers formally adopted a previously informal set of rules in standardizing interactions between the disparate applications on this blockchain (Buntix, 2017). The adoption of what is still known as the "Ethereum Request for Comment" (ERC) number 22 further illustrated how protocological control was left to be exercised by non-state actors. This episode also highlighted the relevance of the identities of the programmers shaping these protocols. Departing from the substantial efforts Satoshi Nakamoto took to remain anonymous, developers behind blockchain protocols became increasingly public figures exercising protocological control over quasi-anonymous payment systems.

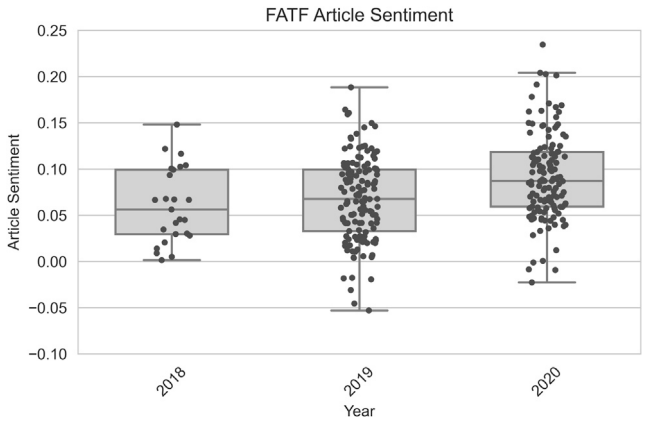
What the 2015 FATF guidance contributed to then was a *taming* of growing worldwide expectations that direct state control could, should and would be exercised over blockchain protocols in quarrels over questions of identity requirements. Key actors at the intersection of cryptocurrencies and fiat currency exchange became increasingly monitored. Yet, protocological control remained exercised by non-state actors. In the Bitcoin protocol debates of 2017 those users most able to harness computing power–the large "pools" of miners–exercised decision-making power in "forking" the original protocol. Similarly, the 2016 hack of The DAO saw a distributed community of users rally around the creator of Ethereum, then 24-year Russian-Canadian Vitalik Buterin, who undertook centralized amendment of this protocol. Both of these instances revealed the degree to which power and control remained market-based and how the FATF guidance did not *alter* non-state control but *extended* it, just as it would do again 4 years later.

Extending Protocological Control to New Markets

The 2019 FATF Guidance assembled fiat-to-cryptocurrency exchanges together with other actors linking real-world identities with the quasi-anonymous payments facilitated by blockchain protocols into a category called "virtual asset service providers" (VASPs). The FATF's guidance on extending its 40 + 9 recommendations to VASPs contained a controversial amendment to its 16th recommendation stipulating that financial institutions should collect and share customer information amongst one another.[8] The FATF specified that by June 2020 VASPs also implement the "travel rules" on customer information adhered to by other financial actors, like banks. The guidance specified that the following identity

[8]The recommendation that "financial institutions include required and accurate originator information, and required beneficial information, on wire transfers and related messages, and that the information remains with the wire transfer or related message throughout the payment chain". The Travel Rule's origins lie in a more than two-decades-old United States requirement that banks store and obtain customer information related to transactions above $3,000.

TABLE 3 | Industry reception of FATF guidance.



Note: Outcomes in boxplots differentiated by year illustrate general sentiments by blockchain industry actors. Sentiments are averaged of 184 articles from Coindesk and 142 articles collected from Cointelegraph. The scale ranges from -1 (most negative) to +1 (most positive) with 0 being neutral.

attributes should "travel" along the chain of transactions exceeding $1,000:

– "(i) originator's name (i.e., sending customer)
– (ii) originator's account number where such an account is used to process the transaction (e.g., the Virtual Asset wallet)
– (iii) originator's physical (geographical) address, or national identity number, or customer identification number (i.e., not a transaction number) that uniquely identifies the originator to the ordering institution, or date and place of birth.
– (iv) beneficiary's name
– (v) beneficiary account number where such an account is used to process the transaction (e.g., the Virtual Asset wallet)" (FATF, 2019: 29)

The collection and transfer of such identity attributes stood in tension with the quasi-anonymous identity standards underpinning digital transactions in and across many blockchain protocols. As one article in the leading cryptocurrency news website *CoinDesk* put it, the extension of the Travel Rule to VASPs "goes against the grain to shoehorn an identity layer onto a technology specifically designed to be pseudonymous" (Allison, 2020a). The FATF guidance and its recommendation to extend the Travel Rule specifically was perceived by many industry actors as "excessively onerous to manage" and decried for the possibility that it "could drive the entire ecosystem back into the dark ages" (Weinberg in Hochstein et al., 2019).

Contrary to these views and critiques of the FATF's exercise of "draconian" power (Hamacher, 2019), however, the task force once again left protocological control to *markets* rather than state-controlled bodies. Notably, the FATF did not call for *public* entities to develop or enforce any set of uniform standards for identity information sharing between VASPs. Instead, the 2019 FATF guidance spurred an intense "race" between *market* players seeking to develop key standards for behavior underpinning digital identity systems that could enable VASPs comply with the Travel Rule and AML/CFT recommendations (De, 2019). Moreover, prior to the publication of the 2019 guidance, the FATF had engaged in a multi-year formal regulatory dialogue with industry actors. It gave dozens of so-called "identity start-up" firms opportunities to develop and test protocols for squaring the circle of, on the one hand, enabling VASPs to collect and transfer data on users, while on the other hand ensuring that user anonymity would remain protected (Henry et al., 2018). What we call a *protocol dialogue* involved industry-FATF deliberations on how protocols can and should be developed and applied *by blockchain start-ups and other technology companies*. The development of the FATF, 2019 guidance was summed up by the FATF Secretariat in an interview stating how "[w]e didn't want FATF to sit down and tell technical details of exactly how companies should comply with it because that would quickly become out of date." (Oki, 2019). Once again, the FATF steered protocological control towards the market rather than calling on member states themselves to develop key standards. The FATF governance *of* blockchain relied closely on governance *by* blockchain developers. **Table 3** provides an indication of how the aggregate reception of the 2019 FATF guidance grew more *positive* as fears of its "draconian" actions subsided.[9]

[9]We used automatized webscrapping with Python to collect articles mentioning the FATF from the leading blockchain media platforms CoinDesk and Cointelegraph utilizing the search feature of the respective sites. Documents were coded manually using the qualitative data analysis software NVivo 12 Plus treating the coding itself as part of the analysis (Basit, 2003). This means that we treated coding as a heuristic, more akin to an exploratory problem-solving technique (Saldana, 2009), starting with an *in vivo* approach that coded sections with a word or short phrase taken from each document. The sentiment analysis of the total 326 articles using Python TextBlob module to compare change over the three years containing the most attention to the FATF activities, 2018–2020. TextBlob assigns polarity values between −1 and 1 to certain words and word combinations in each article indicating if a sentence is more positive or more negative terms. Scores per word or word combination are predefined. For example, the word "great" receives a score of 0.8, the word "bad" scores −0.7, but a negation like "not bad" scores a 0.35. TextBlob then averages the all together for longer text and returns a total polarity value for each article (Schumacher, 2015). While a machine learning approach may yield better results, we used the data predominantly in an explorative fashion, limiting our approach to simple text processing.

A trio of caveats are necessary to clarify our central argument thus far regarding how FATF governance *of* blockchain activity through its formal guidance on virtual currency and virtual asset service providers impacted the *location* of protocological control. First, the FATF *itself* did not exercise protocological control but rather shaped the *location* where such control would be exercised. The task force did so by avoiding recommending public approaches encouraging top-down implementation of its AML/CFT recommendations. Instead, the task force sought to ensure that protocol development, implementation and control remained a more bottom-up affair with "identity" start-ups competing with one another. Second, guidance towards market- rather than government-led development of key standards of behavior to novel blockchain activities is an *extension* of the FATF's longstanding risk-based approach. The approach attempts to weigh the challenges and opportunities involved with implementing the 40 + 9 AML/CFT recommendations, recognizing that "harsher" clamp downs and even bans on certain activities may merely send illicit activities to other jurisdictions while undermining possible benefits of technological innovation. In the context of blockchains, the risk-based approach is one that weighs the risks of illicit activities with cryptocurrencies with the promises of financial surveillance offered by its underlying distributed ledger technology. Third, public actors and official policymakers were not absent, but actively encouraged private-sector standard-setting for squaring the circle of privacy and surveillance in blockchain activities. At the so called Virtual 20 (V20) event, held in parallel to the 2019 Group of 20 meeting in Japan, policymakers including ex-FATF President Roger Wilkins, Japanese Congressman Naokazu Takemoto and Taiwanese Congressman Jason Hsu were present in the signing of the national VASP industry agreement to co-develop standards for digital identities (Zmudzinski, 2019). Representatives from the United States Department of Homeland Security and Treasury Department's Financial Crimes Enforcement Network (FinCEN) were all present at the November 2019 Travel Rule Compliance Conference and Hackathon in San Francisco, California, where the Travel Rule Information Sharing Alliance (TRISA)- a private sector grouping consisting of some 50 blockchain firms and non-profits- pledged to develop "key technical solutions that include a directory of validated VASPs as well as a Certificate Authority (CA) model to ensure the public key cryptography".[10]

In summary, formal FATF guidance influenced the *location* where key standards of information communication between VASPs are developed: in the market rather than (international) state bureaucracies. The FAFT did *not* encourage either top-down or draconian enforcement of its legally non-binding standards. Rather its official guidance has recommended that key protocols and identity standards be persistently set by bottom-up *market* activities. The persistent stress on protocological control by market actors is in line with the wider spate of FATF activities and the organization's longstanding openness to private sector influence (Favarel-Garrigues et al., 2009; Amicelle, 2011; Liss and Sharman, 2015). Indeed, it has been argued that "the private sector—in particular the financial services industry and its high-level representatives—is becoming a "non-great power influencer in FATF" (Nance, 2018: 118). At the same time, former FATF personnel have joined efforts to develop "Travel Rule solutions", such as those offered by the Barbados-based Shyft Network (Allison, 2020b). What we identify as "protocol dialogue" was both present in the development of the 2019 guidance and its on-going implementation. Limits on the effective *form* of protocological control that the FATF helped steer in turn shape the intergovernmental organization's goal of preventing money laundering and the financing of terrorism.

FORM OVER FUNCTION: THE FRAUGHT EXERCISE OF PROTOCOLOGICAL CONTROL

In this section we highlight tensions between governance *of* blockchains and governance *by* blockchain. Specifically, we illustrate how the market-based *form* of protocological control the FATF has promoted fails to overcome the "pitfalls of private governance of identity" (Goanta, 2020; see more generally; Ronit and Schneider, 1999) and undermines the objectives reducing illicit finance in blockchain-based activities. While this argument can only be confirmed through analysis of events unfold over the coming years, we mobilize initial support for our position across two subsections. First, we point to the growing divide between standards of behavior in two spheres of blockchain-based activity, noting the development of *dualling identity protocols*. Second, we examine the 2020 FATF guidance on digital identities where we note a doubling down on the existing form of market-led protocological control. These trends, we contend, contribute to the fraught route towards global digital identity governance, one in which the reducing illicit activities appears increasingly unattainable.

Dualling Identity Protocols

Protocological control by market actors in blockchain activities has taken on a *dual* form undermining rather than addressing the goals of the FATF of reducing international illicit financial flows. Highly fragmented and split standards of behavior emerged for blockchain activities governed by market forces. On the one hand are protocols integrating the identity requirements of the Travel Rule. On the other hand, are protocols disregarding FATF recommendations and seeking to maintain the anonymity of their users. While both sets of protocols pledge to maintain user privacy, only the former incorporate blockchain-based activities into the identity requirements of the existing global AML/CFT regime. The latter protocols, meanwhile, push blockchain-based activities further out of the reach of formal remit of AML/CFT enforcement. This leads the very illicit activity the FATF is charged with reducing and stamping out to be progressively driven further into, rather than out of, the shadows of the "dark net". In elaborating this argument, we first detail the "dualling identity protocols" before situating their importance in the emergence of global digital identity governance.

[10]https://trisa.io/trisa-momentum-announcement/.

Protocological control is exercised by some market actors in ways that closely accord with FATF guidance. Here user identity information is collected and exchanged between and beyond VASPs in ways that closely resemble the more established forms of centralized governance that blockchains originally arose to bypass and counter. Centralized messaging platforms for "VASPs to share encrypted transmittal information with each other securely and privately" are provided by firms like Taiwan-based Sygna.[11] Other start-ups such as Coinfirm, Netki, Shyft and KYC Chain all provide similar "solutions" and are based on private or permissioned blockchain protocols with centralized gatekeepers akin to those of traditional digital systems. Even purportedly "decentralized" solutions offered by blockchain alliances and associations take-on degrees of centralized control. A prominent example is the Travel Rule Information Sharing Alliance, an association of more than 50 entities "focused on security and interoperability between the travel rule standards and protocols".[12] In December 2019, this alliance developed the Intervasp Messaging Standard 101 (IVMS-101) standard (Allison, 2020d), described as "a universal common language for communication of required originator and beneficiary information between VASP"[13]. In May 2020, InterVASP was launched as a technical standard providing a common language for communication between originator and beneficiary VASPs.[14] Such private sector-self governance closely emulates longstanding types of global associations of highly centralized financial exchanges like the World Federation of Exchanges (McKeen-Edwards, 2010).

Further steps towards "decentralized" peer-to-peer solutions being developed also contain persistent elements of centralization. For example, certificates holding transacting users' Personally Identifiable Information are maintained by centralized authorities. TransactID is overseen by California-based Netki, while the free open-source peer-to-peer VASP Address Confirmation Protocol is developed by California-based CipherTrace,[15] which sells the above type of "forensic tool" for the United States Department of Homeland Security.[16] These "blockchain forensics tools" developed to extend CFT/AML standards clearly recentralize in collaborating not only with traditional financial intermediaries, but also with governments (Nelson, 2020). The degree of such collaboration became apparent in a leaked 2019 report provided to the United States Financial Crimes Enforcement Network[17] and other financial regulators by the Cryptocurrency Indicators of Suspicion (CIOS) Working Group, a network of blockchain intelligence firms, exchanges and big banks that detailed dozens of illicit patterns of transactions on blockchain along with a "road map" for tackling them (del Castillo, 2019). Given these connections, it is not inconceivable that these firms enable the sharing of customer information not only between VASPs, but also with law enforcement and intelligence agencies, many of whom are their clients or prospective future clients. Sharing of such information would replicate the kinds of longstanding relationships between such agencies and banks (Amicelle, 2011), the latter of whom are also developing protocols such as Travel Rule Protocol developed between Dutch bank ING, British bank Standard Chartered and United States brokerage firm Fidelity (Allison, 2020e).

A parallel form of protocological control is exercised by market actors *eschewing* customer identification and information sharing requirements and pushing blockchain activity further from official regulatory remit. So-called "privacy protocols" Cashshuffle/Cashfusion,[18] Enigma,[19] Lelantus, MimbleWimble, OpenBazaar and others being tested like Lelantus (Powers, 2020a) provide enhanced standards of anonymity that do not attempt to maintain compliance with either AML/CFT or the Travel Rule customer identification and information exchange requirements. While some protocols here aim for compliance with FATF recommendations and are incorporating blockchain-based activities into formal global AML/CFT governance,[20] most protocols push blockchain-based activities further out of the reach of formal remit of AML/CFT enforcement. The FATF, 2019 guidance has affected what we label the *protocol selection* of VASPs undertaking selective, *ad hoc* compliance with AML/CFT rules. For example, fit-to-cryptocurrency exchanges have delisted cryptocurrencies whose protocols facilitate high standards of anonymity. OKEx Korea confirmed in 2019 that it would halt trading of privacy-coins Monero (XMR), Dash (DASH), Zcash (ZEC), Horizen (ZEN) and Super Bitcoin (SBTC), citing grounds of conflicts with FATF guidelines (Suberg, 2019). Nonetheless, around a third of the top 120 exchanges themselves were found in a survey to have little in the way of AML/CFT controls themselves (Palmer, 2019).

Protocol selection leads to patterns of illicit activity the FATF is charged with reducing and stamping out to be driven deeper into the shadows of the "dark net." Blockchain intelligence firm CipherTrace, for example, reported in 2020 that some 90% of suspicious transactions in cryptocurrencies were being missed by financial institutions (Haig, 2020). The FATF itself lamented these trends in a September 2020 report on "Virtual Assets Red Flag Indicators of Money Laundering and Terrorist Financing." This report was based on more than one hundred

[11]https://www.sygna.io/blog/types-of-fatf-r16-crypto-travel-rule-solutions/.
[12]https://trisa.io/.
[13]https://intervasp.org/.
[14]https://trisa.io/.
[15]https://ciphertrace.com/travel-rule-info-sharing-architecture/.
[16]https://ciphertrace.com/ciphertrace-announces-worlds-first-monero-tracing-capabilities.
[17]Which issued and began immediately enforcing a version of the Travel Rule for United States-based exchanges in 2019.
[18]https://github.com/cashshuffle/spec/blob/master/CASHFUSION.md.
[19]https://enigma.co/.
[20]For example the "trust framework" released by Norwegian start-up Notabene in June 2020 reportedly provides a know-your-customer (KY) through "elements of decentralized identity management to link blockchain addresses to verified profiles" (Allison, 2020c). Switzerland-based OpenVASP, to which Notabene is a member, coordinates the development of a protocol based on Ethereum that "puts privacy of transferred data at the center of its design". It suggests the use of a peer-to-peer messaging system called Whisper which "employs so-called dark routing to obscure message content and sender and receiver details to observers, a bit like anonymous web browsing Tor" (Allison, 2020a). Here identity management is undertaken by a smart contract-based "blockchain public key directory for the VASP and an IBAN-like numbering format: the virtual asset account number" (ibid).

case studies of what it noted are "indications of suspicious activities or possible attempts to evade law enforcement detection" (FATF, 2020b: 5). Meanwhile, FATF's 1 year progress survey of the status of Travel Rule extension to VASPs reported that despite "progress in the development of technological standards for use by different travel rule solutions," there was less implementation of the Travel Rule than other AML/CFT standards (FATF, 2020d: 11). The uneven outcome was blamed on a lack of "sufficient holistic technological solutions for global travel rule implementation that have been established and widely adopted." (FATF, 2020d: 12). Recognizing the "decentralisation ethos that underpins virtual assets, there appears to be a general desire for multiple potential solutions, rather than one centralised travel rule solution." (ibid.). The FATF stressed how the "usage of common standards will assist in ensuring different solutions are interoperable" (ibid.) and called upon "the VASP sector to redouble its efforts towards the swift development of holistic technological solutions encompassing all aspects of the travel rule" driving technology convergence (ibid.). Given the widely reported "struggle to implement" the non-binding "rule" around the world, a 1-year "sunrise period" review extension was granted to VASPs (Bryanov, 2020). By mid-2020, it was reported that authorities in 35 of 54 jurisdictions had implemented Travel Rule standards into domestic legislation and that another 19 had not yet done so (FATF, 2020d). The FATF doubled down on the roles of market actors and emphasized the need for "quick development of technology solutions" (FATF, 2020c: 12).

Governance *of* blockchain by the FATF shaped the *location* of protocological control in ways that allow for the persistent obscuring of identities in blockchain-based activities such as quasi-anonymous payments. Wasabi Wallet, for example, was launched in 2018 to scramble transactions and is based on "secret contracts." In contrast to the smart contracts in Ethereum, secret contracts have nodes capable of calculating data without ever "seeing" them (EC3 Cyber Intelligence Team, 2020). The Secret Network launched a bridge between its privacy-focused blockchain and Ethereum in late 2020 (Powers, 2020b). Europol cited such "privacy-enhanced wallet services" as a "top threat" in its 2020 Internet Organized Crime report.[21] Meanwhile, so-called "decentralized exchanges" (DEXs), developed largely on Ethereum protocols, expanded as increasingly important forums for users to meet and build some semblance of trust in arriving at peer-to-peer agreements to directly exchange cryptocurrency without the use of a formal intermediary or verifying of identities. While still representing a small percentage of overall cryptocurrency trading at the time of writing (around one per cent), the aggregate monthly volumes on DEXs hit records in 2020. British defense and security think-tank RUSI warned that DEXs "have the potential to weaken the role of centralized VASPs and so blunt the effect of governmental regulation" (Moiseienko and Izenman, 2019: viii). The largest DEX by value exchanged, Uniswap, had digital tokens equivalent to more than $1 billion trade in September 2020, yet neither listing rules nor KYC verification procedures (Madeira, 2020). DEXs thus stood at the same crossroad of dualling identity standards as the FATF (2021) proposed in draft guidance published in March 2021to consider them "high-risk" VASPs if they did not implement the Travel Rule standards. The draft guidance also highlighted a number of new "elements of risk," including "[e]xposure to Internet Protocol (IP) anonymizers such as The Onion Router (TOR), the Invisible Internet Project (I2P) and other darknets, capable of further obfuscating transactions or activities and inhibiting a VASP's ability to know its customers and implement effective AML/CFT measures" (FATF, 2021: 15).

In sum, the key risk is that "harsher," "hands-on," state-led restrictions on blockchain activities have the potential to merely *shift* rather than *reduce* illicit activities has emerged in part due to the FATF's shaping of private-sector led exercise of protocological control. While risks pertain to any and all forms of governance, the risks of bottom-up governance strategies are well known now more than a half decade into the FATF's governance *of* blockchain. Calls began to emanate in 2020 for "developing entirely new approaches to manage money laundering and terrorist financing risks" by key industry players (Sian Jones quoted in Allison, 2020f). Without tabling a completely new approach, the FATF (2021) did nevertheless propose some substantial changes in draft guidance published in 2021 that suggested self-regulatory bodies were insufficient for VASP supervision and only "competent authorities" (FATF, 2021: 5) could act as supervisors. The draft guidance proposed in March 2021 also suggested new Principles of Information-Sharing and Co-operation Amongst VASP Supervisor that "[g]iven the pseudonymous, fast-paced, cross-border nature of VAs [Virtual Assets], international co-operation is all the more critical between VASP supervisors." It called for a more "proactive" roles for supervisory authorities rather than self-regulatory industry organizations (FATF, 2021: 94). Even though the six principles the FATF outlined were general and the proposed guidance emphasised in bold that they are non-binding, the draft guidance proposed in March 2021 marked a shift in the FATF's emphasis on closer international cooperation between public supervisors. This shift was emphasized by its guidance, issued just a year earlier, for digital identity (DID) systems. A May 2020 report on how "effective authentication of customer identity for authorizing account access" can enhance "certain elements of customer due diligence (CDD) under FATF Recommendation 10" (FATF, 2020a: 5–7), had still largely called for market-based exercise of protocological control. It recommended member states to leave standard setting to non-state actors, even when using the standards for their own government backed DIDs.[22]

[21]https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment-iocta-2020.

[22]Government authorities should be "supporting the development and implementation of reliable, independent digital ID systems by auditing and certifying them against transparent digital ID assurance frameworks and technical standards, or by approving expert bodies to perform these functions. Where authorities do not audit or provide certification for IDSPs themselves, they are encouraged to support assurance testing and certification by appropriate expert bodies so that trustworthy certification is available in the jurisdiction" (FATF, 2019: 6).

Government authorities were also recommended to remain "flexible" and merely monitor "the rapid evolution of digital ID technology" in order to "help promote responsible innovation and future-proof the regulatory requirements," as well as to support "the development and implementation of reliable, independent digital ID system" along with "assurance testing and certification by appropriate expert bodies."[23] The thrust of the May 2020 guidance persistently focused on ensuring "multi-stakeholder" solutions through constructs such as regulatory "sandboxes" where government authorities monitor private sector trials rather than lead them in any meaningful way. The March 2021 proposed guidance suggesting greater public supervisory cooperation marked a departure from the longstanding emphasis on market-based governance. Future research will have to determine whether the former proposals were mere blips in the longer trend emphasizing the latter.

CONCLUSION: PERSISTENT FORM AND UNACHIEVABLE OUTCOMES?

How can we understand the progressive, piecemeal emergence of global digital identity governance? This paper advanced a two-pronged argument that highlighted the need to consider interactions between governance *of* and *by* blockchains. First, formal governance by the FATF has shaped the "financial route" to global digital identities. Building on its governance *of* financial flows, the FATF has extended its risk-based approach to digital identity. Second, this model of leaving the reins of governance to blockchain developers and start-up firms is fraught with problems. The persistent encouragement of a reliance on market actors in developing blockchain protocols has led to the development of what we identified as dualling identity protocols, or the situation in which some activities are underpinned by standards of activity adhering to AML/CFT rules while others are not at all in accordance with such standards. The persistence of the latter, we argued, undermines FATF goals of reducing rather than just shifting illicit international financial flows. Tensions thus exist and persist between governance *by* and *of* blockchains. Blockchain studies, and emerging literatures on digital identity governance, need to consider the interplay between both forms of governance and how they interact in (un)predictable manners in order to come to a clearer understanding of the roots and evolving forms of digital identity governance.

Future studies should maintain a critical focus on the activities of the FATF and other international organizations, particularly those that have become increasingly vocal about using blockchain to "fight fire with fire," (Lagarde in Wilmoth, 2018) as the former IMF Managing Director Christine Lagarde put it in a 2018 speech. The shaping of protocological control by formal standard-setting organizations is essential to investigate in relation to informal modes of control in developing more nuanced understandings of global digital identity governance. The 2020 "Global Standards Mapping Initiative" of the World Economic Forum and Global Blockchain Business Council, for instance, flagged digital identity as one of the five main areas where overlapping standards have led to gaps in other places (World Economic Forum, 2020). The formal activities of IOs like the International Standards Organization (ISO) require much further attention going forward, especially regarding its various blockchain working groups.[24] Further scholarship would identify whether and how these IOs influence the location and forms of protocological control. They should provide normative assessments of the shifting forms, impacts and limits such forms of control have on actually stemming illicit activity, as well as on socio-economic development more widely. Finally, the extent to which the forms of protocological control shaped by the FATF and other global north rich country clubs can also be effectively contested and challenged by actors in the global south deserves further investigation. In sum, there are promising and pressing research pathways for future studies to explore at the intersection of governance *by* and *of* blockchains.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

MH benefited from funding by the Hans-Böckler-Stiftung, Project number 2017-437-2.

ACKNOWLEDGMENTS

We thank Oskar Gstrein for the invitation to contribute to these debates. We are also grateful for feedback from Gary Robinson, as well as participants in workshops "Anticipatory Global Governance: International Organisations and Political Futures" held at the EISA European Workshop in International Studies, Kraków, June 2019, and "Algorithmic Knowledge in Culture and Media" held at the Open University of Israel, Tel Aviv, October 2019. The usual disclaimers apply.

[23]https://www.fatf-gafi.org/media/fatf/documents/recommendations/pdfs/Guidance-on-Digital-Identity-Executive-Summary.pdf.

[24]https://www.iso.org/committee/6266604.html.

REFERENCES

Allison, I. (2020b). Binance Throws Weight behind Shyft Network in 'Travel Rule' Standards Race. Available at: https://www.coindesk.com/binance-throws-weight-behind-shyft-network-in-travel-rule-standards-race (Accessed November 6, 2020).

Allison, I. (2020d). Crypto Firms Establish Messaging Standard to Deal with FATF Travel Rule. Available at: https://www.coindesk.com/crypto-firms-establish-messaging-standard-to-deal-with-fatf-travel-rule (Accessed November 6, 2020). doi:10.1515/9780822377245

Allison, I. (2020f). *FATF Needs Entirely New Approach to Regulating Crypto, Says V20 Summit*. New York City, NY: Coindesk.

Allison, I. (2020c). Identity Startup Notabene Launches Exchange Tool for FATF Travel Rule Compliance. Available at: https://www.coindesk.com/crypto-identity-startup-notabene-launches-trust-framework-for-fatf-travel-rule (Accessed November 6, 2020). doi:10.1515/9780822377245

Allison, I. (2020e). In Banking First, ING Develops FATF-Friendly Protocol for Tracking Crypto Transfers. Available at: https://www.coindesk.com/in-banking-first-ing-develops-fatf-friendly-protocol-for-tracking-crypto-transfers (Accessed November 6, 2020). doi:10.1515/9780822377245

Allison, I. (2020a). Inside the Standards Race for Implementing FATF's Travel Rule. Available at: https://www.coindesk.com/inside-the-standards-race-for-implementing-fatfs-travel-rule (Accessed November 6, 2020).

Amicelle, A. (2011). Towards a 'new' Political Anatomy of Financial Surveillance. *Secur. Dialog.* 42 (2), 161–178. doi:10.1177/0967010611401472

Atzori, M. (2017). Blockchain Technology and Decentralized Governance: Is the State Still Neccesary? *J. Govern. Regul.* 6, 1–62. doi:10.22495/jgr_v6_i1_p5

Avant, D. (2005). *The Market for Force: The Consequences of Privatizing Security*. Cambridge: Cambridge University Press. doi:10.1017/cbo9780511490866

Basit, T. (2003). Manual or Electronic? the Role of Coding in Qualitative Data Analysis. *Educ. Res.* 45 (2), 143–154. doi:10.1080/0013188032000133548

Boakye-Adjei, N. Y. (2020). Covid-19: Boon and Bane for Digital Payments and Financial Inclusion. Available at: https://www.bis.org/fsi/fsibriefs9.pdf (Accessed November 6, 2020).

Bryanov, K. (2020). Slow but Steady: FATF Review Highlights Crypto Exchanges' Struggle to Meet AML Standards. Available at: https://cointelegraph.com/news/slow-but-steady-fatf-review-highlights-crypto-exchanges-struggle-to-meet-aml-standards (Accessed November 6, 2020).

Buntix, J. P. (2017). ERC20 Token Standard Officially Formalized by Ethereum Developers. Available at: https://themerkle.com/erc20-token-standard-has-now-been-officially-formalized-by-the-ethereum-developers/ (Accessed November 6, 2020).

Campbell-Verduyn, M. (2018b). "Towards a Block Age or Blockages of Global Governance?," in *Bitcoin and beyond: Cryptocurrencies, Blockchains, and Global Governance* (New York: Routledge), 178–197.

Campbell-Verduyn, M. (2018a). Bitcoin, Crypto-Coins, and Global Anti-money Laundering Governance. *Crime Law Soc. Change* 69 (2), 283–305. doi:10.1007/s10611-017-9756-5

Campbell-Verduyn, M., and Hütten, M. (2019). Anticipating Decentralization through Protocological Control: International Organizations and the Standardization of Blockchain Technology within Financial/Security Infrastructures," in Finance/Security Infrastructures workshop, November 13 (New York City, NY: University of Amsterdam).

Campbell-Verduyn, M., and Hütten, M. (2020). Is the Travel Rule Good or Bad for Crypto? Both. Available at: https://www.coindesk.com/is-the-travel-rule-good-or-bad-for-crypto-both (Accessed November 6, 2020). doi:10.1057/s41268-020-00198-5

De Filippi, P. (2018). "Blockchain : a Global Infrastructure for Distributed Governance and Local Manufacturing" in *The Mass Distribution of Almost Everything*. Editor T. Diez (Spain, Institute for Advanced Architecture of Catalonia). doi:10.3917/puf.filip.2018.01

De, N. (2019). CipherTrace Enters Race to Solve Crypto's FATF Compliance Headache. Available at: https://www.coindesk.com/ciphertrace-enters-race-to-solve-cryptos-fatf-compliance-headache (Accessed November 6, 2020).

del Castillo, M. (2019). Crypto's Valachi Papers. Available at: https://www.forbes.com/sites/michaeldelcastillo/2019/12/04/cryptos-valachi-papers/?sh=1a8637dd3117 (Accessed November 6, 2020).

Eaton, B., Hedman, J., and Medaglia, R. (2018). Three Different Ways to Skin a Cat: Financialization in the Emergence of National E-ID Solutions. *J. Inf. Technol.* 33, 1–83. doi:10.1057/s41265-017-0036-8

EC3 Cyber Intelligence Team (2020). Cyber Bits. Available at: https://www.coindesk.com/wp-content/uploads/2020/06/1_5096130532387848318.pdf (Accessed November 6, 2020).

Faria, I. (2021). The Market, the Regulator, and the Government: Making a Blockchain Ecosystem in the Netherlands. Finance and Society, earlyView. doi:10.1109/blockchain.2019.00067

FATF (2020d). 12-Month Review of the Revised FATF Standards on Virtual Assets and Virtual Asset Service Providers. Available at: https://www.fatf-gafi.org/media/fatf/documents/recommendations/12-Month-Review-Revised-FATF-Standards-Virtual-Assets-VASPS.pdf (Accessed November 6, 2020).

FATF (2020e). Digital Identity. Available at: https://www.fatf-gafi.org/media/fatf/documents/recommendations/Guidance-on-Digital-Identity.pdf (Accessed November 6, 2020).

FATF (2019). DRAFT GUIDANCE ON DIGITAL IDENTITY. Available at: https://www.fatf-gafi.org/media/fatf/documents/publicconsultation/Digital%20ID-public-consultation-version.docx (Accessed March 25, 2020).

FATF (2021). Draft Updated Guidance for a Risk-Based Approach to Virtual Assets and VASPs. Available at: https://www.fatf-gafi.org/media/fatf/documents/recommendations/March%202021%20-%20VA%20Guidance%20update%20-%20Sixth%20draft%20-%20Public%20consultation.pdf (Accessed March 24, 2021).

FATF (2018). FATF Guidance on Counter Proliferation Financing. Available at: http://www.fatf-gafi.org/media/fatf/documents/reports/Guidance-Countering-Proliferation-Financing.pdf (Accessed November 6, 2020).

FATF (2015). Guidance for a Risk-Based Approach: Virtual Currencies. Available at: https://www.fatf-gafi.org/media/fatf/documents/reports/Guidance-RBA-Virtual-Currencies.pdf (Accessed November 6, 2020).

FATF (2020a). International Standards on Combating Money Laundering and the Financing of Terrorism & Proliferation. Available at: www.fatf-gafi.org/recommendations.html (Accessed November 6, 2020).

FATF (2020c). Money Laundering and the Illegal Wildlife Trade. Available at: http://www.fatf-gafi.org/media/fatf/documents/Money-laundering-and-illegal-wildlife-trade.pdf ([Accessed November 6, 2020).

FATF (2013). Prepaid Cards, Mobile Payment and Internet-Based Payment Services. Available at: https://www.fatf-gafi.org/media/fatf/documents/recommendations/Guidance-RBA-NPPS.pdf (Accessed November 6, 2020).

FATF (2020b). Virtual Assets Red Flag Indicators of Money Laundering and Terrorist Financing. Available at: http://www.fatf-gafi.org/media/fatf/documents/recommendations/Virtual-Assets-Red-Flag-Indicators.pdf (Accessed November 6, 2020).

FATF (2014). Virtual Currencies - Key Definitions and Potential AML/CFT Risks. Available at: https://www.fatf-gafi.org/media/fatf/documents/reports/Virtual-currency-key-definitions-and-potential-aml-cft-risks.pdf (Accessed November 6, 2020).

Favarel-Garrigues, G., Godefroy, T., and Lascoumes, P. (2009). *Les sentinelles de l'argent sale au quotidian: Les banques aux prises avec l'antiblanchiment*. Paris: La Decouverte. doi:10.3917/dec.favar.2009.01

Frazier, L. (2020). Already Leaning towards Digital Money, Covid-19 Pushes More People towards Contactless Payments. Available at: https://www.forbes.com/sites/lizfrazierpeck/2020/08/21/already-leaning-towards-digital-money-covid-19-pushes-more-people-towards-contactless-payments/?sh=70317e13012a (Accessed November 6, 2020). doi:10.2172/1763531

Galloway, A. R. (2004). *Protocol: How Control Exists after Decentralization*. Cambridge, Massachusetts: MIT press. doi:10.7551/mitpress/5658.001.0001

Goanta, C. (2020). The Private Governance of Identity on the Silk Road. *Front. Front. Blockchain.* 3, 4. doi:10.3389/fbloc.2020.00004

Gutterman, E., and Roberge, I. (2019). "The Financial Action Task Force: Fighting Transnational Organised Crime, Money Laundering, and the Limits of Experimentalist Governance," in *In* Handbook Of Organised Crime And Politics. Editors F. Allum and S. Gilmour (Northampton, Massachusetts: Edward Elgar Publishing Limited)), 455–467. doi:10.4337/9781786434579.00043

Haig, S. (2020). Banks Failing to Identify up to 90% of Suspicious Crypto Transactions. Available at: https://cointelegraph.com/news/banks-failing-to-identify-up-to-90-of-suspicious-crypto-transactions (Accessed November 6, 2020).

Hamacher, A. (2019). CipherTrace and Shyft Unveil a Fix for Draconian FATF Anti-terrorism Rules. Available at: https://news.yahoo.com/ciphertrace-shyft-unveil-fix-draconian-131045661.html (Accessed November 6, 2020). doi:10.1007/978-3-662-59008-9

Hasselbalch, J. A. (2018). Innovation Assessment: Governing through Periods of Disruptive Technological Change. *J. Eur. Publ. Pol.* 25 (12), 1855–1873. doi:10.1080/13501763.2017.1363805

Henry, C. S., Huynh, K. P., and Nicholls, G. (2018). Bitcoin Awareness and Usage in Canada: An Update. Available at: https://www.bankofcanada.ca/wp-content/uploads/2018/07/san2018-23.pdf (Accessed November 6, 2020).

Herian, R. (2018). *Regulating Blockchain: Critical Perspectives in Law and Technology*. London: Routledge. doi:10.4324/9780429489815

Hochstein, M., De, N., and Baydakova, A. (2019). Beyond KYC: Regulators Set to Adopt Tough New Rules for Crypto Exchanges. Available at: https://www.coindesk.com/beyond-kyc-global-regulators-appear-set-to-adopt-tough-new-rules-for-crypto-exchanges (Accessed November 6, 2020).

Hooper, A., and Holtbrügge, D. (2020). Blockchain Technology in International Business: Changing the Agenda for Global Governance. *Ribs* 30 (2), 183–200. doi:10.1108/ribs-06-2019-0078

Hülsse, R. (2008). Even Clubs Can't Do without Legitimacy: Why the Anti-money Laundering Blacklist Was Suspended. *Regulation & Governance* 2 (4), 459–479. doi:10.1111/j.1748-5991.2008.00046.x

Hülsse, R., and Kerwer, D. (2007). Global Standards in Action: Insights from Anti-money Laundering Regulation. *Organization* 14, 5625–5642. doi:10.1177/1350508407080311

Hütten, M. (2019). The Soft Spot of Hard Code: Blockchain Technology, Network Governance and Pitfalls of Technological Utopianism. *Global Network* 19 (3), 329–348. doi:10.1111/glob.12217

Idol, Coin. (2019). Hard Fork: Motivations Fueling Bitcoin Civil War. Available at: https://coinidol.com/bitcoin-civil-war/ (Accessed March 25, 2021).

Jones, K. (2019). Blockchain in or as Governance? Evolutions in Experimentation, Social Impacts, and Prefigurative Practice in the Blockchain and DAO Space. *Ippologia* 24 (4), 469–486. doi:10.3233/ip-190157

Liss, C., and Sharman, J. C. (2015). Global Corporate Crime-Fighters: Private Transnational Responses to Piracy and Money Laundering. *Rev. Int. Polit. Econ.* 22 (4), 693–718. doi:10.1080/09692290.2014.936482

Madeira, A. (2020). The Rise of DEXs: Fueled by DeFi and Ready to Disrupt the Status quoAccessed. Available at: https://cointelegraph.com/news/the-rise-of-dexs-fueled-by-defi-and-ready-to-disrupt-the-status-quo (Accessed November 6, 2020).

McKeen-Edwards, H. (2010). "World Federation of Exchanges" in *Handbook of Transnational Economic Governance Regimes*. Editors C. Tietje and A. Brouder (Leiden: Martinus Nijhoff Publishers), 489–500.

Moiseienko, A., and Izenman, K. (2019). *From Intention to Action - Next Steps in Preventing Criminal Abuse of Cryptocurrency*. London, United Kingdom: RUSI Occasional Paper.

Naheem, M. A. (2019). Exploring the Links between AML, Digital Currencies and Blockchain Technology. *Jmlc* 22 (3), 515–526. doi:10.1108/jmlc-11-2015-0050

Nance, M. T. (2018). The Regime that FATF Built: an Introduction to the Financial Action Task Force. *Crime Law Soc. Change* 69, 109–129. doi:10.1007/s10611-017-9747-6

Nelson, D. (2020). Inside Chainalysis' Multimillion-Dollar Relationship with the US Government. Available at: https://www.coindesk.com/inside-chainalysis-multimillion-dollar-relationship-with-the-us-government (Accessed November 6, 2020).

Oki, H. (2019). 'Not Everyone Is Happy but We Have to Move on,' Some Challenges to the FATF's New Guidance. Available at: https://cointelegraph.com/news/not-everyone-is-happy-but-we-have-to-move-on-some-challenges-to-the-fatfs-new-guidance (Accessed November 6, 2020).

Palmer, D. (2019). A Third of Crypto Exchanges Have Little or No KYC, Says CipherTrace. Available at: https://www.coindesk.com/a-third-of-crypto-exchanges-have-little-or-no-kyc-says-ciphertrace (Accessed March 25, 2021). doi:10.1136/bmjspcare-2019-huknc.60

Pavlidis, G. (2020). International Regulation of Virtual Assets under FATF's New Standards. *J. Invest. Compl.* 21 (1), 1–8. doi:10.1108/JOIC-08-2019-0051

Powers, B. (2020a). *Secret Network Launches Bridge to Bring Transactional Privacy to Ethereum*. New York City, NY: Coindesk.

Powers, B. (2020b). Zcoin Employs Burn-And-Redeem Privacy Model, Offering Alternative to Coinjoins. Available at: https://www.coindesk.com/privacycoin-zcoin-burn-redeem-testnet (Accessed November 6, 2020).

Reijers, W., O'Brolcháin, F., and Haynes, P. (2016). Governance in Blockchain Technologies & Social Contract Theories. *Ledge* 1, 134–151. doi:10.5195/ledger.2016.62

Ronit, K., and Schneider, V. (1999). Global Governance through Private Organizations. *Governance* 12, 243–266. doi:10.1111/0952-1895.00102

Saldana, Johnny. (2009). *The Coding Manual for Qualitative Researchers*. Thousand Oaks, California: Sage Publications ltd.

Schumacher, A. (2015). TextBlob Sentiment: Calculating Polarity and Subjectivity. Available at: https://planspace.org/20150607-textblob_sentiment/ (Accessed November 6, 2020). doi:10.1007/978-3-658-10702-4

Suberg, W. (2019). OKEx Korea Delists Monero, Dash, Privacy-Cryptos over FATF Demands. Available at: https://cointelegraph.com/news/report-okex-delisting-monero-dash-privacy-cryptos-over-fatf-demands (Accessed November 6, 2020).

Truman, E. M., and Reuter, P. (2004). *Chasing Dirty Money: Progress on Anti-money Laundering*. Washington: Institute for International Economics.

Tsingou, E. (2010). Global Financial Governance and the Developing Anti-money Laundering Regime: what Lessons for International Political Economy? *Int. Polit.* 47 (6), 617–637. doi:10.1057/ip.2010.32

Wilmoth, J. (2018). 'Fight Fire with Fire': IMF Chief Lagarde Calls for Blockchain-Powered Bitcoin Regulation. Available at: https://finance.yahoo.com/news/fight-fire-fire-imf-chief-185347015.html (Accessed March 25, 2021).

World Economic Forum (2020). Global Standards Mapping Initiative: An Overview of Blockchain Technical Standards. Available at: http://www3.weforum.org/docs/WEF_GSMI_Technical_Standards_2020.pdf (Accessed November 6, 2020).

Zmudzinski, A. (2019). Group of Digital Asset Trade Associations to Establish Global Cryptocurrency Association. Available at: https://cointelegraph.com/news/group-of-digital-asset-trade-associations-to-establish-global-cryptocurrency-association (Accessed November 6, 2020).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership