Towards trustworthy blockchains
normative reflections on blockchain-enabled virtual institutions

Teng, Yan

DOI
10.1007/s10676-021-09581-3

Publication date
2021

Document Version
Final published version

Published in
Ethics and Information Technology

Citation (APA)
Teng, Y. (2021). Towards trustworthy blockchains: normative reflections on blockchain-enabled virtual institutions. *Ethics and Information Technology*, *23*(3), 385-397. https://doi.org/10.1007/s10676-021-09581-3

Ethics and Information Technology
https://doi.org/10.1007/s10676-021-09581-3

ORIGINAL PAPER

Towards trustworthy blockchains: normative reflections on blockchain-enabled virtual institutions

Yan Teng[1]

Accepted: 18 January 2021

Abstract

This paper proposes a novel way to understand trust in blockchain technology by analogy with trust placed in institutions. In support of the analysis, a detailed investigation of institutional trust is provided, which is then used as the basis for understanding the nature and ethical limits of blockchain trust. Two interrelated arguments are presented. First, given blockchains' capacity for being institution-like entities by inviting expectations similar to those invited by traditional institutions, blockchain trust is argued to be best conceptualized as a specialized form of trust in institutions. Keeping only the core functionality and certain normative ideas of institutions, this technology broadens our understanding of trust by removing the need for third parties while retaining the value of trust for the trustor. Second, the paper argues that blockchains' decentralized nature and the implications and effects of this decentralization on trust issues are double-edged. With the erasure of central points, the systems simultaneously crowd out the pivotal role played by traditional institutions and a cadre of representatives in meeting their assigned obligations and securing the functional systems' trustworthy performances. As such, blockchain is positioned as a technology containing both disruptive features that can be embedded with meaningful normative values and inherent ethical limits that pose a direct challenge to the actual trustworthiness of blockchain implementations. Such limits are proposed to be ameliorated by facilitating a shift of responsibility to the groups of people directly associated with the engendering of trust in the blockchain context.

Keywords Blockchain technology · Trust · Trustworthiness · Trust in institutions · Trust in technology · Blockchain ethics

Introduction

The question of trust is of essential importance to the prominence achieved by blockchain technology. In the whitepaper of Bitcoin, the pseudonymous creator, Satoshi Nakamoto, makes it clear that the primary purpose of creating a decentralized electronic payment system is to remove the need for trusting third-party institutions (e.g., banks) that are often considered necessary for facilitating online transactions between heterogeneous groups of participants (Nakamoto 2008). However, in recent years, increasing research has pointed out that, rather than evaporation of trust, it might be more accurate and less ambiguous to interpret blockchain-enabled "trustlessness" as a shift of trust from centralized authorities to blockchain technology and the associated people, such as developers and miners (Werbach 2018; Sas and Khairuddin 2017; Al-Saqaf and Seidler 2017; Ostern 2018). The trust shifted to blockchain technology is sometimes framed as trust in code (Maurer et al. 2013; Velasco 2017), trust in quasi-entities (Reijers and Coeckelbergh 2018), or trust in algorithmic authority (Lustig and Nardi 2015).[1]

While the above efforts suggest the viability of blockchain trust, little research has explicated the nature and ethical limits of this trust form through the lens of philosophical theories of trust. Conceptualizing blockchain trust in one way or another largely impacts how we understand the role played by this technology in our lives, and more importantly, it can carry different implications shaping what values we want trustworthy blockchains to embody. In philosophical studies of trust, several important accounts have been

✉ Yan Teng
tyan0318@outlook.com

[1] Ethics and Philosophy of Technology, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, The Netherlands

[1] While there are different setups of blockchain technology such as public/private/consortium blockchains, for analysis reason, the blockchain systems discussed in this paper refer to public blockchains such as Bitcoin and Ethereum.

Published online: 28 January 2021

proposed to understand trust in technologies (Taddeo 2010; Coeckelbergh 2012; Nickel 2013). Yet what is intriguing and unique about blockchain trust is that it seems not just a matter of trust in technologies. The blockchain's potential for providing a self-sufficient way to reach consensus facts without third-party authorities indicates the system's mixed role transgressing between a technological system for achieving functional services and an institution-like entity that can organize relatively stable patterns of social practices. Such a combined position explains why this technology has been referred to as both "an institutional technology" and "a technological institution" (Davidson et al. 2018; Reijers et al. 2016). The institutional aspect of blockchain technology emphasizes the importance of understanding and evaluating blockchain trust based on what we know about trust in institutions. Neglecting the richness and moral significance of institutional trust may conceal what we expect from trustworthy institutions, and thus reduce the tasks that should be addressed by blockchain systems as alternatives to traditional institutions to merely technical aspects.

Unlike trusting a particular person, our trust placed in institutions and those who fill institutional roles is often considered abstract, diffuse, and impersonal (Govier 1997; Luhmann 1979; Coeckelbergh 2015). According to Luhmann (1979, p. 48), the aim of this form of trust (or "system trust" in his term) is to reduce the complexity of interacting with different functional systems (e.g., a financial system) usually seen as necessary for individuals to live in a complex modern society. Although Luhmann's account does not delve into the normative aspect of institutional trust, it takes complex social processes, norms, and the functionality of social systems as important sources of trust that govern our shared expectations about the right ordering and stability of the systems. This view seems to provide a good starting point for understanding blockchain trust given the similar striking capacity of this technology for delivering predefined normative values. As many scholars have argued, the original blockchain is not value-neutral; it is the manifestation and reinforcement of particular norms and values over others (De Filippi and Loveluck 2016; Golumbia 2015; De Filippi and Hassan 2018; Ishmaëv 2019). Besides, applications of this technology may further transform social relations in a way that follows the systems' rigid and non-negotiable features (Reijers and Coeckelbergh 2018). The shared capacity between institutions and blockchains for being normative entities indicates the possibility of understanding blockchain trust in terms of the features of institutional trust.

With these considerations, this paper presents a novel and meaningful way of conceiving of trust in blockchain technology by analogy with what we understand of trust in institutions. In support of the analysis, two core issues revolving around blockchain trust are examined. First, by discussing how blockchain trust resembles our predictive and normative expectations towards institutions, the nature of blockchain trust is argued to be best understood as a special type of trust in institutions with trust-inviting elements built into, rather than outside, its technical infrastructure. Second, what we know about institutional trust is further utilized as an analytic tool on which the ethical limits of blockchains' trustworthiness can be reflected. As such, a constructive reflection on blockchain trust as a special form of trust in institutions is provided, with the aim of providing perspectives from which the trustworthiness of blockchain applications could be responsibly improved.

It should be emphasized that such an analysis of blockchain trust touches on two core questions of trust as a relational structure:

1) What constitutes the trustor's trust? This question primarily concerns the trust-establishment phase.
2) What constitutes the trustee's trustworthiness? This question focuses more on the trust-evaluation phase.

By elucidating these two questions in the blockchain context, this paper not only contributes to clarifying what people may expect from specific blockchains and how such institution-like systems should be assessed, but more importantly, it builds a normative conception of blockchain trust that could help proactively shape blockchain applications and their effects. The analysis provided in this paper, thus, provides a way of doing blockchain ethics via a constructive reflection on the most crucial value (i.e., trust) associated with this disruptive technology.

This paper will proceed as follows. It begins by discussing the trust revolution brought about by the technical potential of blockchains for creating various virtual institutions that could replace third-party authorities in promoting trusted interactions. Given the importance and possibility of exploring blockchain trust in terms of trust in institutions, the paper then embarks on a detailed investigation of institutional trust. Next, the institutional trust account proposed is applied to analyse the normative aspects of blockchain trust, allowing blockchain trust to be understood as a plausible and meaningful form of trust resembling institutional trust. Here the ethical limits of blockchains' trustworthiness are discussed as a result of removing central authorities. Finally, the limits articulated are used as perspectives from which blockchain implementations' trustworthiness can be properly improved by facilitating a shift of responsibility to the developers and networks of users.[2]

[2] "User" here refers to both miners who contribute to the operation of the network and a wide range of normal users who only use the network as a way to facilitate interactions.

The trust revolution: blockchain systems as virtual institutions

The following section discusses how blockchain systems disrupt a traditional way of facilitating trusted interactions. First, it briefly clarifies what is meant by the term "trust" in philosophy and the role played by third parties in promoting interactions between people who have no trust in each other. It then looks at the technical potential of blockchain technology for eliminating the need for third parties and thus revolutionizing the way we trust.

Understanding trust and the role played by third-party authorities

As much research into trust would agree, trust is an elusive concept that has multifaced nature (Simon 2013; Baier 1994; Ess 2010). In the most general sense, trust can be regarded as a phenomenon that develops within a relation that requires at least two parties: a trustor and a trustee (McLeod 2020; Coeckelbergh 2012; Taddeo 2010). In trust discourse, scholars have proposed several important accounts that can help tease out the complex nature of the trust notion. Gambetta (1988, p. 217), for example, suggests a rational account by defining trust as a probabilistic assessment of the likely behavior of another. Likewise, Coleman (1990) views trust as a cognitive decision made in line with one's benefit-risk analysis of engaging in some form of cooperation with another. Despite the importance of cognitive reasons for trust, reducing the richness of trust relations to purely cognitive dimension is widely considered narrow and hollow since it does not touch upon the essence of our sense of trust (Hollis 1998; Baier 1986; Hardin 2002).

Unlike reliance, trust is a balance between confidence and vulnerability in that by trusting, one is willing to give up some discretionary power and freedom to the trustee whose behavior one cannot perfectly control or predict (Baier 1986; Werbach 2018). In other words, trust always involves the risk of being letting down that purely rational accounts fail to explain. In most cases, such "giving up" and risk-taking can be explained by an important non-cognitive dimension emphasized by other trust accounts, such as normative accounts that consider trust as reliance on others' responsibility for accomplishing their duties and obligations, e.g., trust in institutional representatives (Hollis 1998; Walker 2006), affective accounts in which emotions and affects play a determinant role for one to develop trust, e.g., children-to-parents trust (Weckert 2005), or motivation-based accounts that highlight the moral significance of the trustee's goodwill towards the trustor, e.g., trust between good friends (Baier 1986). Thus, despite the debate over which dimension is the primary source of trust, human trust is usually thought to be an integrated result of both cognitive and non-cognitive dimensions (Ess 2010; Taddeo 2010), and the question of which particular non-cognitive factor becomes *most* relevant is deeply entwined with the nature of the trust relation in question (Simon 2013).

The cognitive and non-cognitive dimensions of trust are closely engaged with the reasons for trust. As Ferrario et al. (2019) argue, reasons for one to trust another contain two sorts: pragmatic reasons that trusting someone or something can probably improve the trustor's well-being, such as gaining profits, building cooperation, saving time and energy, and preserving moral values, and epistemic reasons that relate to the trustor's belief in the trustee's trustworthiness. This means, on the one side, trust is deeply relational—engaged with a particular person's needs and interests—and highly contextual—impacted by whether there are better alternatives in a specific context. On the other side, for the trustor, the value of trust is achievable insofar as the trustee is in fact trustworthy with respect to the entrusted task (Hardin 2002; Nickel 2015). Both sides show the importance of the trustor's awareness of the trustee's trustworthiness.

Unlike trust, trustworthiness is a quality that indicates to others whether one will act as expected (Taddeo 2010). It allows others to expect the benefit and risk of placing trust reasonably. Yet, arriving at cogent reasons for trust requires the trustor to be familiar with the potential trustee which also explains why trust in tightly-knit groups is widely regarded as the original form of trust (Luhmann 1979). When the two parties are not familiar with-, or do not already trust each other, a credible third-party or middleman who can help them build trusted interactions is often needed. For example, think of Alice and Bob as two teenagers who have no trust in each other but would like to trade stamps, and think of Clark as a credible stamp shop owner in town, who offers the service of facilitating stamp trading for earning a good reputation and small fees. In this case, it is fairly reasonable for Alice and Bob to proceed with the trade through the hand of Clark since, with him, they could trade safely without the need for trusting each other.

This simple way of facilitating trusted interactions between individuals is in fact prevalent in almost all sorts of modern economic activities. For high-stakes decisions and more complex interactions, Clark's role is usually filled by trusted institutions that could provide formal endorsements and indemnity by protecting the participants' vulnerabilities and interests. By placing trust in third-party authorities rather than one another, participants reduce their risk. Such risk-reducing interactions make it reasonable for participants to engage in an activity. From the direct communicative actors, to a credible third-party like Clark, and then to formal institutions, the shift of trust highlights the fact that, for transactions between strangers, the goods of trust for the trustor are not necessarily linked to a particular trustee but can be achieved by alternatives contingent on social and technological development. As will be discussed below, this also explains why blockchain technology is frequently viewed as an alternative to third-party institutions.

The elimination of third parties: blockchains as alternatives

In the context of online transactions, institutions for promoting trusted interactions are mainly banks, firms, markets, exchanges, governments, and the relevant financial and legal systems they collectively furnish. While these institutions provide necessary means for economic activity to be processed recurrently and reliably, dependence on centralized entities not only involves extra costs, risks, and uncertainties but also relies heavily on their integrity and credibility (Nakamoto 2008). Along with the financial crisis of 2008, increasing concerns about the drawbacks and insufficiency of trusting centralized authorities have been expressed. With this background, blockchain technology, as the decentralized solution enabling the Bitcoin project, first came to prominence with realizing the ledger function that used to be provided exclusively by centralized institutions.

A blockchain is a distributed transactional database that enables continuous transitions of system states without the intervention of any intermediary (Glaser 2017). The core quality valued by blockchain start-ups, as Dupont and Maurer (2015) state, is blockchains' potential for being record-keeping devices. A record-keeping device (i.e., a ledger) provides a way to create consensus on the factual recording of the state of an economy, which is considered of pivotal importance for coordinating modern commerce (Davidson et al. 2018). Traditionally, such a ledger is issued and kept exclusively by a central authority that monitors all transactions that have ever taken place. By contrast, the Bitcoin blockchain adopts a decentralized and transparent approach with all valid transactions publicly announced to a large network of computers in chronological order, providing an alternative way to ensure the accuracy of transaction records and prevent double-spending attempts.

In cases where no trusted authority is involved, achieving a factual and shared state of the ledger is the main issue faced by any new solution. The Bitcoin blockchain solves this problem by using a consensus algorithm (i.e., proof-of-work) based on cryptographic tools and a series of consensus rules such as a fixed block format, the longest chain rule, and the incentive mechanism. More specifically, new transactions are collected from the memory pool and grouped into a block with other information required by the block format. Nodes competing for the single power of adding a new block to the chain are called miners. They are incentivized to join the competition by profitable rewards in return for their computation power and electricity. The competition requires them to find a solution to a complex cryptographic puzzle for the issued block as the proof-of-work. After a miner solves the puzzle, the result will be broadcasted to the network and verified by other nodes. And if it is valid, the block will be added to the chain, and that miner can get some coins as rewards. For the blockchain, modifying data in a past block is extremely hard and costly since a malicious user has to assemble a majority of the hash power to redo the proof-of-work of the target block and all blocks after it. Regarding this, the peer-to-peer network is considered robust enough to maintain a single history of order in which all blocks and transactions recorded are valid and immutable (Nakamoto 2008).

In short, the blockchain is designed to facilitate a reliable ledger that could replace those issued by commercial banks, and it has been proven to be secure since no permanent damage has been done to the network since its inception. As transactions processed by the blockchain are validated and verified within the system, the network is able to provide a new basis of trust without relying on a third party (van den Hoven et al. 2019). Considering this third-party-free setting and the fact that Bitcoin meets all criteria of existing legal institutions of digital property, Ishmaëv (2017) further argues that the blockchain can function as a self-sufficient alternative institution of property alongside the traditional structure.

Furthermore, by integrating a fully-fledged, built-in programming language, the Ethereum blockchain introduces another main functionality known as smart contracts to the blockchain industry (Buterin 2013). Essentially, smart contracts are the pieces of code that can be built in a way that only the code determines what will happen once it is triggered (Glaser 2017). Such programmable contracts enable interactive services and market mechanisms to be built on distributed autonomous organizations (DAO) made of software and governed by a network of participants, further releasing this technology's potential for being an institutional technology. As Davidson et al. (2018) put forward, the fact that blockchain technology possesses many elements of market capitalism—such as exchange mechanisms, property rights, code-based law, and financial investments—makes it eligible to create a new mechanism for coordinating market economy. Such a new mechanism has the potential to complement or replace the current mechanism operated collectively by governments, firms, markets, etc. Considering the essential roles played by ledgers and contracts for constituting modernity (Reijers et al. 2016; Dupont and Maurer 2015), it is not surprising to see that ambitious blockchain-based initiatives aiming to create state-like, cloud

Table 1 The conceptual structure of institutional trust

Dimensions of institutional trust	Institutions' qualities	Reasons for institutional trust
Predictive expectations	Functional aspects	Epistemic/pragmatic reasons
Normative expectations	Normative aspects	Pragmatic reasons

communities (e.g., the Bitnation project) are also proposed (Tempelhof et al. 2017).

To sum up, the above discussion on blockchains' potential for *record-keeping* and *contract-enforcement* provides an analysis of how blockchains can function as virtual institutions and facilitate trusted interactions between participants. This means users of the networks can reliably interact with each other without, apparently, the need for trusting any external authority or anybody in particular. For transactions enabled by blockchain systems, instead, everything needed seems to be users' trust in these institution-like entities. However, based on the analysis of the complex nature of trust presented above, it can be argued that more is needed to understand the relationship between "blockchain trust" and intuitional trust in addition to the similar functions provided by the two sorts of entities. In other words, a plausible notion of blockchain trust resembling the essence of trust in institutions should also explicitly refer to the normatively loaded expectations one may hold towards the systems. This requires us to first understand the rich meaning of institutional trust, including an understanding of the normative expectations towards institutions, in order to properly grasp and assess blockchain trust as a special form of trust in institutions.

A conceptual investigation of institutional trust

The functional aspects of blockchains discussed above show the technical potential of the systems for being institution-like entities. The following section works to elaborate on the nature and moral significance of institutional trust, setting the stage for further exploration of blockchain trust.

Beyond prediction: normative expectations of institutions

As mentioned, trust in institutions is diffuse and does not necessarily depend on personal contact. As Alfano and Huijts (2020) put forward, trust in large-scale institutions can be non-partner-relative, meaning that trustingness and trustworthiness can be valid without a predefined partner. Also, institutional trust could be non-thing-specific. In many cases, "we did not rely on X to do A and Y to do B… but rather that we expected reliable, courteous, and orderly service" of that institution (Walker 2006). Based on the non-partner-relative and non-thing-specific structure of institutional trust, when individuals state that they trust an institution, what they are referring to and relying upon is closer to an acceptable and stable service state of that institution; i.e., they trust that it will do, in a general and abstract sense, what it is institutionalized to do. In this paper, this institutional trust account is named *the normative account of institutional trust*. Through the lens of this account, the establishment of institutional trust is not exclusively grounded in our predictive expectations about the functions that an institution will provide, but more importantly, it relies on our normative expectations of that institution and individuals who fill the institutional roles to do what they are supposed to do. Such normative attitude links trust to the relevant trustees' responsibility for complying with their duties and obligations assigned by their institutional roles, capturing the non-cognitive dimension of institutional trust.

These two sorts of interrelated expectations echo the two dimensions of trust discussed earlier and are closely related to the trustor's reasons for trust. More specifically, the predictive expectations towards the relevant trustees are commonly grounded in the trustor's epistemic and pragmatic reasons for trust, i.e., whether one believes that the trustees are trustworthy enough to provide specific functions that can satisfy a particular end of the trustor. Alternatively, the normative expectations towards an institution seem to engage with the trustor's pragmatic reasons for trust, depending on whether particular values and norms one favours are inherent in a given institution and can be delivered by its representatives and overall performances. For the sake of clarity, a sketch of the conceptual structure of institutional trust proposed is shown in Table 1.

For understanding the essence of the normative account of institutional trust, it is important to discuss how people's normative expectations towards others or institutions are generated and why such expectations are essential for building trust. As Hollis (1998, p. 34) argues, normative expectations, under either moral or social headings, are not congruent with merely predictive expectations we hold towards functions of objects. Instead, they are grounded in the shared moral understanding that people will act as they should, and the anticipation of others' responsibility for complying with standards and behave responsively (Walker 2006). In other words, these two bases allow us to expect *of* others that they will act as what the standards require while holding them responsible for meeting those standards (Jones 2004). For instance, when we trust a taxi driver, a dentist, and a delivery person whom we do not know well, we expect of them that they will do their jobs correctly and meet their obligations, promises, and professional standards in a responsible way without assuming their particular concern or regard for us.

When it comes to an institution trustee, normative expectations are often grounded in our shared belief about the normative values stably tied to an institution. According to Turner (1997), institutions are a complex of norms, values, roles, and positions embedded in specific kinds of social structures that can organize fairly enduring patterns of social practices. In other words, institutions can be understood as entities carrying predefined normative qualities, such as moral, social, and legal norms. Such norms, as Lewis (2002) argues, can be interpreted as promises and commitments, providing signals for people to form their beliefs about the actions that should be followed in order to fulfil those promises. Likewise, Bicchieri (2006) depicts norms as "collectively shared scripts" that can guide common anticipation of the corresponding actions that are considered consistent and appropriate under such norms. In this regard, it can be said that normative values inherent in institutions play a significant role in shaping and guiding what one expects from institutions and their representatives.

Considering this, institutions could be viewed as viable entities of trust in the sense that people can rely on and evaluate them in a normatively loaded way. Such expectations might be thinner than those relevant to interpersonal trust, but they are still natural and comprehensible given our everyday experience with some institutions. For example, in trusting the value of money, one presumes that the economic system and the relevant people will perform in the right way that is considered normatively desirable and established as practically trustworthy (Jalava 2006). Such trust is developed via continual, affirmative experience in using money. It supports one to believe that the system can facilitate the desirable characteristics embedded in fiat money (e.g., acceptability, durability, portability, etc.) stably and recurrently while requiring no specific guarantees. During constant interactions with different sorts of institutions and social systems, such expectations often become a default that is not necessarily assessed every time before interaction (Luhmann 1979, p. 50). As such, a positive feedback loop of trust between humans and the monetary system can be built via the normative qualities inherent in the system and our daily experience that help confirm the usefulness of the relevant expectations.

Responsible actors as the way to secure institutions' qualities

Similar to interpersonal trust, our shared understanding of the relevant human actors' responsibility for complying with their assigned obligations provides us a way to believe that our expectations invited by institutions' built-in qualities are secured, as we see from Walker (2006, p. 84), I give the bank teller my deposit, but I rely on the institution's competence and fiduciary responsibility, and the system of regulation that ensures and enforces its compliance, and the responsibility of whoever ultimately oversees that system to see to it that my money goes and stays where it is supposed to.

In this example, responsibility is ascribed to the bank, the associated legal systems, and the human agents who fill the relevant institutional roles. When people interact with the bank teller, on the one hand, they tend to place a default trust in the whole functional system, supposing that the system will work effectively. On the other hand, they presume that the bank teller and other representatives have some sort of legal and moral responsibility for complying with obligations assigned by their institutional roles and are to be held accountable if trust is violated after the fact. As such, people take themselves to be entitled to the right order of particular services of the system and the generally responsive and trustworthy behavior of those representatives.

Although the trustor's premise that individual representatives of institutions will and should be responsible for their obligations generally remains tacit, unreflective, and nonspecific, this premise seems to be crucial for our understanding of institutional roles. As Demolombe and Louis (2006) clarify, an institutional role refers to a set of implicit and explicit rights and obligations in relation to some individuals' position or legal status in an institution. People who fill such a role, accordingly, can be understood as individuals to whom the predefined set of norms and status functions apply (Searle and Willis 1995). According to Demolombe and Louis, an institutional role contains two sorts of properties—i.e., descriptive and normative properties—that both give specifications of the role and direct our expectations of their performances. For example, the role of bank teller is characterized by descriptive properties: to have specific professional skills and experience, and by normative properties: to have obligations to assist customers with all relevant bank services. In this case, if anyone is in fact a bank teller, it is reasonable for a customer to presume that she is competent in handling particular tasks and has responsibility for doing whatever obligations assigned by the role of bank teller.

In particular, knowing that someone will and should be responsible for doing what they ought to do provides the trustor extra confidence in institutions' trustworthiness in three types of situations. Firstly, since any trust contains the risk of being violated, such a premise makes the trustor reasonably expect that, were things to go wrong, they would ultimately identify someone to be held accountable for the wrong things and get them changed to the right way. Secondly, in ambiguous and flexible situations, the premise that some human actors can finally be found allows trustors to hope that there is some space for negotiation that could benefit themselves. Thirdly, such a premise also drives one to believe that, apart from what is required by the representatives' institutional roles, these individuals are prone to perform in a trust-responsive way since people are inherently reputation-seeking and have the desire to be well regarded (Pettit 1995).

To sum up, the analysis above proposes to understand trust in institutions as predictive and normative expectations towards institutions' performances, with a particular consideration of how the responsibility of institutions and their representatives shapes our expectations of institutions. Accordingly, institutions' trustworthiness is mainly influenced by the functional and normative aspects of their performances, as well as the responsibility of the relevant individuals for securing the realization of institutions' built-in qualities. This is in a nutshell the conceptual structure of institutional trust proposed by this paper. This structure, on the one hand, gives form and direction to examine the extent to which blockchain trust can be regarded as a type of trust in institutions. On the other hand, and relatedly, it paves the way for a broader reflection on blockchain applications' actual trustworthiness.

Applying the institutional trust account to blockchain trust

The above institutional trust account shows the importance of the normative values built into institutions for the generation of trust and the importance of responsible actors for the realization of institutions' built-in qualities. Applying this account to trust issues related to blockchains, thus, requires an understanding of: (1) whether blockchains contain the capacity for delivering norms that could *invite* the corresponding expectations similar to those invited by counterpart institutions, and (2) whether the systems can provide a way to *secure* the realization and maintaining of predefined functional and normative requirements. While the former question ultimately determines the plausibility of conceptualizing blockchain trust as a meaningful form of trust in institutions, the latter question directly leads us to reflect on blockchain applications' actual trustworthiness.

The normative relevance of blockchain trust

It is often thought that blockchain technology will be eliminating the need for trust. One important claim of this paper is that the removal of third parties does not eliminate the need for trust, or more specifically, its non-cognitive dimension. It rather shifts the trust to blockchains. First, many empirical studies on technology trust have shown that trust as a value predisposition or a mental shortcut significantly impacts public perceptions and adoption of sophisticated technologies (Ho et al. 2010; Mah et al. 2014). This means that, in a descriptive sense, trust that takes into account non-calculative factors, such as normative and affective sources, could be to some extent seen as a prerequisite for those who lack systematic knowledge and expertise of a technology application to take a "leap of faith" and use the application. This is particularly the case when the application in question is so complicated that reaching a rational assessment of the entire system's trustworthiness is extraordinarily difficult (Corley and Scheufele 2010; Ishmaëv 2018), or when the trust is about innovative practices that are inherently uncertain (van den Berg and Keymolen 2017).

Second, this function of trust, as related to technology adoption and complexity reduction, can be supported by the essence of trust discussed earlier: trust is a way to allow people to accept the fact that dependence on another person or entity will expose them to the possibility of being harmed (Möllering 2006). Thus, on the one hand, relying on a particular technology implicitly or explicitly requires the need for trust to suspend vulnerabilities and risks involved in the use of that technology. Removing the intervention of third parties, hence, tends to shift trust from a system's human masters to the system itself and the network behind (Werbach 2018). On the other hand, given the different criteria people employ to develop and assess trust, the heterogeneity involved in humans' trust in technologies should be specified on a case-by-case basis (Taddeo 2010). Nevertheless, there are approaches that can interpret some common characteristics of trust in technologies. Coeckelbergh (2012), for example, proposes a phenomenological-social approach that captures trust as an emergent and/or embedded property of social relations. In this way, he argues that, as technologies are already part of our lives, trusting technologies is less under the control of individuals but more like a default that emerges from social relations.

The idea of conceptualizing trust in technologies in terms of institutional trust—which this paper endorses—is essentially an effort to interpret more specific inner connections between humans and technologies. Such an idea is not new, but it has not yet been systematically explored. Nickel (2020; 2013) notes that technologies can be direct objects of trust since they are subscripted by some of the evaluation standards that are used to reach and justify our trust decisions towards institutions. The way we evaluate sophisticated systems, as he argues, is not merely about whether their functions are reliable or not (like a hammer); we also care, in an evaluative sense, whether they are doing things correctly. In this claim, an analogy between institutions and technologies is drawn in virtue of their similar capacity for being entities that can invite normative evaluation of their performances. The view that technologies contain normative aspects can be better explained by Moor's (2006) clarification of the two categories of normative viewpoints. As he argues, technologies are normative entities because they can be evaluated by:

a. Non-moral normative viewpoints, which assess particular technologies' performances in terms of their intended purposes or design norms. Such norms can be interpreted as the principles and objectives guiding technologies' performances that do not necessarily draw on ethical consequences;
b. Moral normative viewpoints, which take moral norms to evaluate those technical performances that are of ethical relevance. This could be the case, for example, when the technical performances can generate ethical consequences or contain built-in moral considerations (Tavani 2015).

On the basis of these two ways of understanding the normative aspects of technologies and the institutional trust account articulated, it can be said that technologies resemble institutions in their design capacity for carrying normative values and inviting relevant expectations about what they are supposed to do. In this regard, it seems that technologies could be plausibly viewed as objects of trust in the sense that they could be relied upon and evaluated in a normatively loaded manner.

The analogy discussed above becomes more striking when the technology trustee in question is the original blockchain. Not only does the system share comparable norm-delivering capacity with institutions, but it is explicitly designed to carry out exactly the same core functions of its counterpart institutions and deliver the set of design norms that are considered desirable in the economic context. For this reason, this paper argues that blockchain trust is not simply a type of trust in technologies (like trust in an autonomous vehicle) that can be framed as institutional trust in a general sense, but blockchain trust is itself a form of trust in institutions. For creating an alternative to the trust model enabled by third-party authorities, as articulated, the most daunting task of the peer-to-peer network is to reach a shared state of the database that can ensure the validity and irreversibility of all transactions, which is also viewed as the core functionality provided by every blockchain system (Glaser 2017). Essentially, the normative purpose of this task is to provide a global source of truth on which the associated values required for empowering the decentralized solution can be reasonably approached. Such values primarily include (a) data integrity, which indicates the completeness and accuracy of the information shared; (b) data transparency that prevents counterfeits and dishonest behavior by improving information symmetry and audit compliance; (c) data authentication, which ensures a reliable process to verify the identity of a person or a single piece of data; and (d) data security that makes sure that records issued by the network are tamper-resistant and risk-tolerant. Following Moor's clarification of normative entities, these values could be viewed as the design norms built into the blockchain's infrastructure, which can readily inspire users to generate the corresponding expectations.

Thus, if we consider institutions as a complex of norms and values folding into particular social structures for delivering relatively stable services, the blockchain can be seen as a further step that attempts to keep only the core functionality and certain normative ideas of its counterpart institutions while eliminating these entities as well as their bureaucratic processes and power holders. Moreover, certain parts of the blockchain incorporate explicit considerations as a resistance to the power dynamics enabled by centralized authorities, bringing about effects and implications that are not just normatively but also morally relevant. According to Tavani (2015), Moor's two kinds of standards for evaluating the impacts of a specific technology (i.e., design norms and moral norms) can lead to different levels of trust, from low to high. In this regard, if one's expectations towards the blockchain are about its morally relevant features, the level of trust the trustor places in the system can be higher than those who bear no such expectations.

A proper understanding of the moral features tied to the original blockchain's performances requires a brief review of the moral significance of cryptography-enabled data decentralization. In 1985, David Chaum proposed the idea of using decentralized solutions based on cryptographic techniques to solve moral issues—such as mass surveillance, erosion of democratic rights, and opinion manipulation—entailed by centralized computer systems (Chaum 1985). As a crucial component of Bitcoin protocol, cryptographic techniques thus provide a good starting point for establishing a global decentralized infrastructure that could dilute the power of monopolies and contribute to protecting moral values such as freedom, autonomy, and privacy (Ishmaëv 2019; Scott 2014). In this respect, the distributed database technology might be considered more praiseworthy than traditional solutions, especially in cases when these values are already at stake. For similar reasons, the blockchain has been depicted as a neoliberal project or a libertarian dream through which the control of nation-states on the economy can be reduced so that "governing without governments" might be achieved (De Filippi and Loveluck 2016). Systems built on blockchain technology, thus, have the capacity for bringing about significant effects on challenging authorities and shaping people's understanding of the power-relation of the society (Reijers and Coeckelbergh 2018).

In this regard, the specific norms and values presented by blockchain implementations are very likely to attract the participation of those actors who favour such normative ideas. Also, the profound norm-delivering capacity of this technology inevitably attracts those who are interested in using such capacity for their own purposes (Ishmaëv 2019). These normative aspects of blockchains, thus, can be valid and plausible reasons to invite users' trust. Such reasons fall into the category of the pragmatic reasons discussed earlier, which are deeply relational and engaged with the trustor's specific interests and needs that might be met by a given trustee's performances. With all these considerations, it seems fair to say that blockchain trust is grounded in and goes beyond our trust placed in institutions. By removing the role of internet aggregators while providing an alternative way to help achieve the value of trust for the trustor, blockchain technology brings about a fundamental change in the way we trust and benefit from the goods of trust. Thus, the normative conception of blockchain trust as a special type of trust in institutions is proposed by analysing the nature of human-to-blockchain relation against the background of theories of trust and institutional trust, which should not be confused with any descriptive claim about trust.

The ethical limits of blockchains' trustworthiness

The above analysis shows the appropriateness and plausibility of conceptualizing blockchain trust based on trust placed in institutions. Nevertheless, the goods of trust only accompany well-grounded trust (McLeod 2020). Thus, it is crucial to distinguish between how trust can be invited and how trust should be evaluated. From the perspective of blockchain ethics, while the former considers the importance of addressing the conceptual vacuum of blockchain trust by understanding its nature, the latter focuses on assessing the implications of blockchain trust with the aim that more well-grounded trust can be achieved.

Despite the advantages produced by blockchains' disruptive features, along with the erasure of central points, blockchain applications' trustworthiness also raises ethical concerns. This paper argues that the decentralized novelty of blockchain technology has dual effects on trust. It eliminates the risk, cost, and complexity related to third parties while simultaneously crowding out the pivotal role of institutions and a cadre of representatives in meeting their assigned obligations and securing the functional systems' reliable performances.

This means that blockchains' decentralized nature carries significant implications and consequences for issues impacting trust that are of ethical relevance. First, individual representatives do play an important role, especially in unexpected situations. Although there is a risk that, after the fact, human actors of institutions are shown to be incompetent and not responsive to their duties and obligations, these people can be held accountable for their misconducts and even facing punitive measures. In comparison, the lack of control over a blockchain's performance and the lack of clear attribution of responsibility in blockchain communities imply that, were things to go wrong (e.g., loopholes and attacks), nobody would be held accountable for the incidents, and the irreversible nature of the system leaves almost no room for recourse. As Reijers and Coeckelbergh (2018) point out, the high level of blockchains' rigidity is achieved at the cost of a reduction in the dynamic understanding of the freedom and responsibility of the actors involved. In this respect, a market economy built on blockchains may put its trustors in a more vulnerable position than the trust model involving centralized authorities, particularly considering those small networks where attacks are easier to occur.

Second, there are risks deriving from unreasonable normative expectations. Although many expectations related to blockchains seem plausible, such as those related to the design norms and moral norms of the original blockchain discussed earlier, it is not at all surprising that some expectations are not evidence-based. Unrealistic normative expectations, as Buechner and Tavani (2011) mention, also exist in human-to-institution trust. Yet, the fundamental difference between those invited by institutions and blockchains is that the relevant qualities of blockchains are often hidden and less guaranteed (Ishmaëv 2019). Think of the Bitnation project that purports to create blockchain-enabled democratic communities online. A fundamental concern of this idea is that democratic communities in civil society are created by negotiation and compromise between members with diverse backgrounds, conflicting interests, and different conceptions of the common good, but not by a homogenous group of participants who can voluntarily join and leave (De Filippi and Hassan 2018). Thus, the intention of transforming territorial associations into blockchain-based communities would be fatal to democratic values as it tends to eclipse other types of moral and political reasoning. A more profound ethical concern is the non-neutrality of blockchain technology itself. As Golumbia (2015) argues, the basic setting of blockchain technology is considered deeply political, with "right-wing, libertarian, and anti-government" ideology embedded. Organizing democratic communities via blockchains, thus, makes democracy vulnerable to the ideological biases inherent in this technology (Dumbrava 2018). In this regard, if advocates normatively expect that the project can safeguard and promote democracy adequately, their expectations will be frustrated due to the deeply flawed assumptions built into the system.

At the very least, institutional norms are usually under constant scrutiny of democratic debates and examined by long-lasting practices (De Filippi and Hassan 2018). In contrast, we could say that, in addition to trustor's lack of investigation, the generation of unrealistic expectations towards blockchain implementations may also be caused by the absence of actors who are formally responsible for explicating, scrutinizing, and updating the set of assumptions inscribed into the systems and monitoring the actual performance of the systems' norm-delivering capacity. As Jones (2012) argues, trustworthiness requires that trustees are willing and able reliably to signal to others the domains that they are competent and will be responsive to others' dependency. Therefore, compared to institutions where the implementation of the relevant normative ideas is secured by a number of human actors and well-established procedurals, blockchains are designed to float merely in the rules of algorithms. This raises ethical concerns over the reliability of the systems' normative qualities.

The above analysis clarifies that blockchains' decentralized nature and the implications and effects of this decentralization on trust are double-edged. Without the backing of credible parties, the systems put more burden and risk on users themselves without proper measures to redeem unexpected situations and guarantee the systems' actual norm-delivering performances. All these claims seem to point to the ever-pressing need of trustors for being vigilant and reflective knowers. To reach well-grounded trust in a digital context, resonated with Simon's (2010) view, not only do trustors have a duty to check the integrity and competence of the trusted entity, but they must also scrutinize their standards for evaluating others' trustworthiness. Simply put, users need to be more responsible for their trust decisions as a result of distrusting third parties.

Towards trustworthy blockchains: a shift of responsibility

Seeking to make trust more well-grounded, nevertheless, is just one side of the coin and restricted by subjectivity-specific differences with respect to users' knowledge, time, and resource. As Keymolen (2019) points out, our ability to establish trust is affected also by the social context in which we are positioned, such as social roles that make each other's actions and expertise more predictable. In the blockchain context, to effectively respond to the challenges faced by the technology, this paper argues that, apart from a wide range of users, more responsibility should be shifted to developers and active network peers (i.e., miners) who are associated with the actual performance of blockchain applications. Clarifying the specific roles played by these groups and reframing their responsibilities accordingly provide a way to improve our abilities to develop trust by addressing a focus on understanding what is at stake for the development of trustworthy blockchains. Such an effort can be used to inform the design and decision-making related to blockchains-based systems, building affordances that foster warranted trust and foreclose affordances that would undermine warranted trust. In this way, the ethical limits of blockchains' infrastructure discussed earlier are used as perspectives from which the trustworthiness of blockchains might be gradually improved.

While blockchain technology is designed to eliminate the need for centralized authorities, it is not designed to remove the reliance on developers who maintain the actual codebase through the workflow and determine the functionality and the main values of the system (Glaser 2017). As Nickel (2013) clarifies, developers are presupposed to have two trust-related tasks: the first is to make the system as reliable as possible, and the second is to identify the system's trustworthiness to people in a position to trust that system. The core issue here, coupled with the two dimensions of blockchain trust discussed above, is to sufficiently show that the disruptive functions of blockchain technology together with the meaningful normative values imparted can be realized in practice.

For the functional aspects, compared to the big promises made by the original blockchain's whitepaper, its current reference implementation, referred to as Bitcoin Core, is facing many intractable technical issues such as low throughput, high latency, and a tremendous waste of electricity, which are especially apparent in comparison with the efficiency of the incumbent payment gateways they tend to replace, e.g., Visa and PayPal (Swan 2015). While it is clear that the development and practical applicability of blockchain implementations are still in their infancy, solving the above issues is the shared responsibility of the developer community inherent to their role in the whole ecosystem.

Moreover, given the risk of unexpected situations harming the basic functions of the network, explicit strategies for self-governance and crisis response within the developer community and the peer-to-peer network are hardly optional tasks. A valuable lesson learned from the most infamous incident that occurs in the Ethereum blockchain (i.e., the DAO hack) is that decentralization should not be either-or.[3] The accident shows that in order to protect the network's overall interests, certain sacrifice of the blockchain's immutability and decentralization is in fact considered appropriate and acceptable for the majority of the community. In this sense, effective self-governance adopted to ensure the proper performance of a blockchain might be as useful as the safeguard provided by centralized institutions. However, the current governance structure of blockchain communities is quite technocratic, and the responses provided are relatively arbitrary, two facts which cause concerns about the fragility of the community's decision-making processes and its capacity for dealing with incidents. (De Filippi and Loveluck 2016). Thus, what is lacking is a generic, well-established

[3] For more information about the hack, see https://medium.com/@ogucluturk/the-dao-hack-explained-unfortunate-take-off-of-smart-contracts-2bd8c8db3562.

governance mechanism ready to be applied to interpret and respond to possible contingencies. A way in which laws and regulators can here truly help, as Werbach (2018) notes, is not to offer specific governance rules for the community but to provide the community with jurisprudential insights into how rules should be formulated and enforced in a formal way.

As discussed, the normative ideas inscribed into blockchains are also crucial sources of trust and important criteria for evaluating trust. However, for plenty of blockchain implementations, these ideas are not transparent and well-scrutinized, which makes them easy to be flawed and generate undesirable effects on users and society at large. Some of the assumptions simply fall into naive technological determinism, just like the case of Bitnation. As professionals who have the direct ability to use technical means to express human values, developers and designers can play a vital role in advancing responsible technological innovations by helping realize these values properly (van den Hoven et al. 2015). Indeed, many current proposals seek to embed particular desirable normative goals into blockchain design, such as Enigma and Zcash that aim to create privacy-preserving blockchains and Datawallet that is designed to facilitate data ownership. What is lacking, based on the discussion provided in the above subsection, is a satisfactory explanation and justification of how these norms are embedded in and embodied by the technical design. In this regard, making the normative goals transparent to the public is just the first step. Constantly scrutinizing and updating the systems' built-in assumptions on a case-by-case basis is of central importance for improving the normative qualities of blockchains (Nickel 2013; Ishmaëv 2019), and these are the aspects that developers, network peers, philosophers, policymakers can all take part in and contribute to approaching more trustworthy blockchains.

Conclusion

This paper has critically discussed blockchain trust by analogy with trust placed in institutions. Doing so provides a close philosophical reflection on the nature and ethical limits of this trust form. As a result of blockchain's double-edged peculiarities, blockchain trust is characterized, on the one side, as a form of trust grounded in- and going beyond institutional trust. By coding the normative values and technical properties into its basic infrastructure, the original design of blockchain technology touches the most intriguing aspect of trust, i.e., we want our trust to be warranted, more than ever, to dispel our anxieties and worries about the discretionary power possessed by third parties with the hope that the vulnerabilities and risks engendered by placing trust can be minimized to the greatest extent possible. On the other side, blockchain-based systems are confronted with challenges to their actual trustworthiness for functioning as an institution-like entity. Reframing the responsibility shifted to the relevant groups of people in the blockchain context is an essential component of a strategy to address the ethical and societal challenges posed by this disruptive technology. As such, the institutional trust concept is used as an analytical tool to disentangle the double-edged effects of blockchain on trust, and informing ways in which the trustworthiness of blockchain applications could be gradually improved.

Acknowledgements I would like to thank Mark Alfano and Filippo Santoni de Sio for their insightful comments on this paper. I am grateful to Ronald Oosting for his comments on an earlier draft of this paper. I am also grateful for "The Society for Philosophy and Technology" conference in Texas, United States, 2019, for offering me an opportunity to present some of this work. And, I wish to thank Anna Melnyk and Emily Sullivan for their useful feedback on my presentation at the "Normative Issues in Blockchain Implementations" workshop in Delft, the Netherlands, 2018. I wish to thank China Scholarship Council for the financial support through which my position is made possible. In addition, I wish to thank the anonymous reviews whose comments have largely improved the quality of this paper.

References

Al-Saqaf, W., & Seidler, N. (2017). Blockchain technology for social impact: Opportunities and challenges ahead. *Journal of Cyber Policy, 2*(3), 338–354.

Alfano, M., & Huijts, N. M. A. (2020). Trust and distrust in institutions and governance. In J. Simon (Ed.), *Handbook of trust and philosophy*. Routledge Taylor & Francis Group.

Baier, A. (1986). Trust and antitrust. *Ethics, 96*(2), 231–260.

Baier, A. (1994). Trust and its vulnerabilities. *Moral prejudices*, 130–151.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. New York: Cambridge University Press.

Buechner, J., & Tavani, H. T. (2011). Trust and multi-agent systems: Applying the "diffuse, default model" of trust to experiments involving artificial agents. *Ethics and Information Technology, 13*(1), 39–51.

Buterin, V. (2013). Ethereum white paper. *GitHub Repository*. https://ethereum.org/en/whitepaper/. Accessed January 5, 2018.

Chaum, D. (1985). Security without identification: Transaction systems to make big brother obsolete. *Communications of the ACM, 28*(10), 1030–1044.

Coeckelbergh, M. (2012). Can we trust robots? *Ethics and information technology, 14*(1), 53–60.

Coeckelbergh, M. (2015). *Money machines: Electronic financial technologies, distancing, and responsibility in global finance*. Farnham: Ashgate.

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge: Harvard University Press.

Corley, E. A., & Scheufele, D. A. (2010). Outreach gone wrong? When we talk nano to the public, we are leaving behind key audiences. *Scientist, 24*(1), 22.

Davidson, S., De Filippi, P., & Potts, J. (2018). Blockchains and the economic institutions of capitalism. *Journal of Institutional Economics*, 1–20.

De Filippi, P., & Hassan, S. (2018). Blockchain technology as a regulatory technology: From code is law to law is code. *arXiv preprint* . arXiv:1801.02507.

De Filippi, P., & Loveluck, B. (2016). The invisible politics of bitcoin: Governance crisis of a decentralized infrastructure. *Internet Policy Review*, 5(4).

Demolombe, R., & Louis, V. (2006). Norms, institutional power and roles: Towards a logical framework. In *International Symposium on Methodologies for Intelligent Systems* (pp. 514–523). Springer, Berlin, Heidelberg.

Dumbrava, C. (2018). Citizenship forecast: Partly cloudy with chances of algorithms. In R. Bauböck (Ed.), *Debating transformations of national citizenship* (pp. 299–303). Cham: Springer.

DuPont, Q., & Maurer, B. (2015). Ledgers and law in the blockchain. *Kings Review*. http://kingsreview.co.uk/articles/ledgers-and-law-in-the-blockchain/. Accessed November 28, 2018.

Ess, C. M. (2010). Trust and new communication technologies: Vicious circles, virtuous circles, possible futures. *Knowledge, Technology & Policy, 23*(3–4), 287–305.

Ferrario, A., Loi, M., & Viganò, E. (2019). In AI we trust incrementally: A multi-layer model of trust to analyze human-artificial intelligence interactions. *Philosophy & Technology*, 1–17.

Gambetta, D. (1988). Can we trust trust? In Gambetta, Diego (ed.), *Trust: Making and breaking cooperative relations*, 213, 214. Oxford: Basil Blackwell.

Glaser, F. (2017). Pervasive decentralization of digital infrastructures: A framework for blockchain enabled system and use case analysis. In *Proceedings of the 50th Hawaii International Conference on System Sciences*, 1543–1552. Hawaii, United States.

Golumbia, D. (2015). Bitcoin as politics: Distributed right-wing extremism. In G. Lovink, N. Tkacz, and P. de vries (Eds), *MoneyLab reader: An intervention in digital economy*. Amsterdam: Institute of Network Cultures.

Govier, T. (1997). *Social trust and human communities*. Montreal: McGill-Queen's University Press.

Hardin, R. (2002). *Trust and trustworthiness*. New York: Russell Sage Foundation.

Ho, S. S., Scheufele, D. A., & Corley, E. A. (2010). Making sense of policy choices: Understanding the roles of value predispositions, mass media, and cognitive processing in public attitudes toward nanotechnology. *Journal of Nanoparticle Research, 12*(8), 2703–2715.

Hollis, M. (1998). *Trust within reason*. Cambridge: Cambridge University Press.

Ishmaëv, G. (2017). Blockchain technology as an institution of property. *Metaphilosophy, 48*(5), 666–686.

Ishmaëv, G. (2018). Rethinking Trust in the Internet of Things. In R. Leenes, R. van Brakel, S. Gutwirth, & P. de Hert (Eds.), *Data protection and privacy: The internet of bodies* (pp. 203–230). Oxford: Hart Publishing.

Ishmaëv, G. (2019). *Open sourcing normative assumptions on privacy and other moral values in blockchain applications* (doctoral dissertation). Delft University of Technology, the Netherlands.

Jalava, J. M. (2006). *Trust as a decision: The problems and functions of trust in Luhmannian systems theory (Niklas Luhmann)*. https://helda.helsinki.fi/bitstream/handle/10138/23348/trustasa.pdf?sequen. Accessed June 6, 2020.

Jones, K. (2004). Trust and terror. In P. DesAutels & M. U. Walker (Eds.), *Moral psychology: Feminist ethics and social theory* (pp. 3–18). Maryland: Rowman & Littlefield.

Jones, K. (2012). Trustworthiness. *Ethics, 123*(1), 61–85.

Keymolen, E. (2019). When cities become smart, is there still place for trust. *European Data Protection Law Review, 5,* 156.

Lewis, D. (2002). *Convention: A philosophical study*. Oxford: Blackwell Publishers Ltd.

Luhmann, N. (1979). *Trust and power*. Chichester: John Wiley.

Lustig, C., & Nardi, B. (2015). Algorithmic authority: The case of Bitcoin. In *48th Hawaii International Conference on System Sciences* (pp. 743–752). Hawaii, United States.

Mah, D. N. Y., Hills, P., & Tao, J. (2014). Risk perception, trust and public engagement in nuclear decision-making in Hong Kong. *Energy Policy, 73,* 368–390.

Maurer, B., Nelms, T. C., & Swartz, L. (2013). "When perhaps the real problem is money itself!": The practical materiality of Bitcoin. *Social Semiotics, 23*(2), 261–277.

McLeod, C. (2020). Trust. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2020/entries/trust/. Accessed September 6, 2020.

Möllering, G. (2006). *Trust: Reason, routine, reflexivity*. Amsterdam: Elsevier.

Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems, 21*(4), 18–21.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. *Bitcoin*. https://bitcoin.org/bitcoin.pdf Accessed July 1, 2016.

Nickel, P. J. (2013). Trust in technological systems. In M. J. de Vries, S. O. Hansson, & A. W. M. Meijers (Eds.), *Norms in technology, philosophy of engineering and technology* (pp. 223–237). Dordrecht: Springer.

Nickel, P. J. (2015). Design for the value of trust. In J. van den Hoven, PE. Vermaas, & I. van de Poel (Eds.), *Handbook of ethics, values, and technological design: Sources, theory, values and application domains*, 551–567. Dordrecht: Springer Netherlands.

Nickel, P. J. (2020). Trust in engineering. In D.P. Michelfelder & N. Doorn, (Eds.), *Routledge companion to philosophy of engineering*.

Ostern, N. (2018). Do you trust a trust-free transaction? Toward a trust framework model for blockchain technology. In *Thirty Ninth International Conference on Information Systems*, San Francisco, United States.

Pettit, P. (1995). The cunning of trust. *Philosophy & Public Affairs, 24*(3), 202–225.

Reijers, W., & Coeckelbergh, M. (2018). The blockchain as a narrative technology: Investigating the social ontology and normative configurations of cryptocurrencies. *Philosophy & Technology, 31*(1), 103–130.

Reijers, W., O'Brolcháin, F., & Haynes, P. (2016). Governance in blockchain technologies & social contract theories. *Ledger, 1,* 134–151.

Sas, C., & Khairuddin, I. E. (2017). Design for trust: An exploration of the challenges and opportunities of bitcoin users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 6499–6510). ACM. Denver, United States.

Scott, B. (2014). Visions of a techno-leviathan: The politics of the Bitcoin blockchain. https://www.e-ir.info/2014/06/01/visio ns-of-a-techno-leviathan-the-politics-of-the-bitcoin-blockchain/. Accessed September 15, 2020.

Searle, J. R., & Willis, S. (1995). *The construction of social reality*. New York: The Free Press.

Simon, J. (2010). The entanglement of trust and knowledge on the Web. *Ethics and Information Technology, 12*(4), 343–355. https://doi.org/10.1007/s10676-010-9243-5.

Simon, J. (2013). Trust. In Pritchard, D. (Ed.): *Oxford bibliographies in philosophy*. New York: Oxford University Press

Swan, M. (2015). *Blockchain: Blueprint for a new economy*. Sebastopol: O'Reilly Media Inc.

Taddeo, M. (2010). Modelling trust in artificial agents, a first step toward the analysis of e-trust. *Minds and machines, 20*(2), 243–257.

Tavani, H. T. (2015). Levels of trust in the context of machine ethics. *Philosophy & Technology, 28*(1), 75–90.

Tempelhof, S. T., Teissonniere, E., Tempelhof, J. F., & Edwards, D. (2017). Bitnation white paper. *GitHub repository*. https://github.com/Bit-Nation/Pangea-Docs. Accessed January 20, 2019.

Turner, J. H. (1997). *The institutional order: Economy, kinship, religion, polity, law, and education in evolutionary and comparative perspective*. New York: Longman Publishing Group.

van den Berg, B., & Keymolen, E. (2017). Regulating security on the Internet: Control versus trust. *International Review of Law, Computers & Technology, 31*(2), 188–205.

van den Hoven, J., Pouwelse, J., Helbing, D., & Klauser, S. (2019). The blockchain age: Awareness, empowerment and coordination. In D. Helbing (Ed.), *Towards digital enlightenment* (pp. 163–166). Cham: Springer.

van den Hoven, J., Vermaas, P. E., & Van de Poel, I. (2015). Design for Values: An Introduction. In J. van den Hoven, PE. Vermaas, & I. van de Poel (Eds.), *Handbook of ethics, values, and technological design: Sources, theory, values and application domains*. Dordrecht: Springer Netherlands.

Velasco, P. R. (2017). Computing ledgers and the political ontology of the blockchain. *Metaphilosophy, 48*(5), 712–726.

Walker, M. U. (2006). *Moral repair*. New York: Cambridge University Press.

Weckert, J. (2005). Trust in cyberspace. In R. J. Cavalier (Ed.), *The impact of the internet on our moral lives* (pp. 95–117). Albany: SUNY Press.

Werbach, K. (2018). *The blockchain and the new architecture of trust*. Cambridge: MIT Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.