What does it mean to trust blockchain technology?

Teng, Yan

DOI
10.1111/meta.12596

Publication date
2022

Document Version
Final published version

Published in
Metaphilosophy

Citation (APA)
Teng, Y. (2022). What does it mean to trust blockchain technology? *Metaphilosophy*, *54*(1), 145-160. https://doi.org/10.1111/meta.12596

DOI: 10.1111/meta.12596

ORIGINAL ARTICLE

What does it mean to trust blockchain technology?

Yan Teng[1,2]

[1]Ethics and Philosophy of Technology Section, Delft University of Technology, The Netherlands

[2]Governance Research Centre, Shanghai Artificial Intelligence Laboratory, China

Correspondence
Yan Teng, Governance Research Centre, Shanghai Artificial Intelligence Laboratory, 37–38 floors, Yunjin Road 701, Xuhui District, Shanghai 200232, China
Email: tyan0318@outlook.com

Abstract

This paper argues that the widespread belief that interactions between blockchains and their users are trust-free is inaccurate and misleading, since this belief not only overlooks the vital role played by trust in the lack of knowledge and control but also conceals the moral and normative relevance of relying on blockchain applications. The paper reaches this argument by providing a close philosophical examination of the concept referred to as trust in blockchain technology, clarifying the trustor group, the structure, and the normatively loaded nature of this trust relation. The paper ends by critically reflecting on two of the most promising values (decentralization and transparency) that can invite users' trust in blockchain technology, arguing that there is a tension between the pressing values that are intended to be achieved by developers and the predicament situations caused by current blockchain implementations.

KEYWORDS

blockchain technology, blockchain trust, ethics and philosophy of blockchain, ethics of trust, trust and trustworthiness, trust technology

1 | INTRODUCTION

If we think of traditional institutions as the trusted intermediaries that help human cooperation progress from direct reciprocity (that is, an eye for an eye) to indirect reciprocity on which more sophisticated mechanisms of cooperation can be built, then blockchain technology appears to be a possible means to establish a new basis of truth and trust without the need for any third party (van den Hoven et al. 2019). Traditionally, online interactions between heterogeneous participants are facilitated by trusted third-party authorities, such as financial

institutions and their legal branches. As the distributed database technology behind Bitcoin (that is, a cryptocurrency), blockchain technology came to prominence as a decentralized solution that relies instead on consensus algorithms and rules to ensure the validity and immutability of transactions processed by the peer-to-peer network (Nakamoto 2008). With this sort of decentralized nature, the original blockchain can perform as a virtual institution that users can directly rely upon and interact with, which may significantly reduce the risk, uncertainty, and cost involved in trusting third parties.

While the blockchain's designed attempt to eliminate the need for trusting third parties is distinct, the ways of characterizing the role of trust played in blockchain-based interactions are fairly controversial in the literature (Jacobs 2020). Although some writers describe blockchain technology as trustless or trust-free (Nakamoto 2008; Glaser 2017), others capture the change of trust enabled by the technology as a shift of trust from third parties to the underlying algorithms (Simser 2015; Velasco 2017; Werbach 2018, 29; van Lier 2017), and yet others depict the change as trust distributed among developers and miners (Kasireddy 2018).[1] While each of the efforts partially captures the idea of how blockchains change the way we trust, they fall short of structuring a relatively complete picture of the blockchain-enabled revolution in trust. The ambiguity involved not only presents difficulties for developers, regulators, users, and the public to reflect on the value, the targets, and the corresponding risk of talking about trust in the context of blockchain technology, it also conceals the moral and normative relevance of blockchain-powered solutions. A systematic analysis teasing out the intertwined relationship between trust and blockchains is needed.

To this end, it is important to first make clear what causes such a divergence in understanding the blockchain-enabled trust revolution. Two reasons appear to be germane when we examine this issue from the perspective of the trust phenomenon: (1) one shortcoming of the current discussion is the absence of a clearly defined group of trustors (the persons who *give* trust). Due to the complex relationship between knowledge and trust, the lack of a clearly defined trustor group can confuse the understanding of the relationship between humans and technologies, since such a relationship might be interpreted differently—for example, with or without the need for trust—across diverse communities having different levels of blockchain knowledge. Thus, clarifying the main trustor group of blockchain technology is this paper's first task and contribution. (2) The other reason for the divergent views on the role played by trust in blockchain-based interactions, as Jacobs (2020) argues, is rooted in different assumptions that scholars make under the term "trust." For less demanding accounts, trust is understood as predictive expectations a rational actor holds concerning the performances (actions) of a trustee (the person who *receives* trust) (Gambetta 1988; Coleman 1990). In trust theories, these accounts are often labeled rational-choice accounts that regard trust as being a result of weighing all risks and benefits of potential options in context, which is used much like the term "reliance" (Simon 2013). Following these accounts, it seems that trust is applicable not just to humans but also to things, since trust does not require the trustee to have any condition other than reliability for developing a trust relation. This enables people to talk about trust in the blockchain context as a shift of trust from traditional third parties to the algorithms, developers, miners, and markets (such as exchanges and online markets) involved in blockchain technology.

In contrast, for more demanding accounts, trust is construed as a rich, non-vernacular concept that differs from reliance in that the generation of trust involves moral, normative, or affective beliefs about the trustee.[2] Meanwhile, this requires the trustee to have a commitment to the trustor's dependency (McLeod 2020). Such a connection is seen as the essential factor that

[1]It should be noted that as blockchain is an umbrella technology that can be implemented in various ways, the focus of this paper is the original setup (that is, the public, permissionless blockchain) that has the decentralization property and does not rely on any central authority to execute the protocol. This is also the only type of blockchain that can sometimes be considered trustless.

[2]Baier 1986; O'Neill 2002; Hollis 1998; Holton 1994; Weckert 2005.

allows one to take a "leap of faith" and form interactions amid the lack of knowledge and control (Möllering 2006). For these accounts, this connection is the answer to why we use the term "trust." As a result, these accounts are often considered not applicable to nonhuman agents that have no will, which directly results in the understanding that blockchain-based systems are trust-free or trustless.

Other research, however, has shown that, apart from physical persons, we also normatively have expectations of professionals (Jones 2004), institutions (Walker 2006, 83), and technological systems (Nickel 2013). From this perspective, it seems that the depiction of blockchain as a trust-free technology might ignore the rich array of expectations invited by the manifold normative values embedded in blockchain's basic infrastructure, and the negative attitudes (such as disappointment, anger, and feelings of being betrayed) when one's trust is frustrated after the fact. Both aspects are seen as important cues for a normatively loaded trust relation going beyond mere reliance. Thus, the second task of this paper is to explore a conception of blockchain trust that takes into account the trustor's normative expectations about the blockchain's performances. Philosophical discussion on this aspect can help people take a step back from merely focusing on the judgment of the system's reliability and begin to think about questions of moral and normative significance and the relevant risks when talking about blockchain trust. The analysis provided here can further be utilized to steer the design and policy making associated with blockchain implementations, with the aim of indicating directions for developing more trustworthy blockchains and reducing the risk and misplacement of trust.

With these considerations, this paper provides a comprehensive analysis of the role and risk of trust in blockchain-enabled interactions, proposing a user-centered, multilayer-structured framework for understanding blockchain trust in a meaningful way. The rest of the paper is organized as follows. Section 2 defines normal users as the main trustor group of blockchain systems, clarifying the reasons trust is needed in interactions between users and blockchains. Section 3 structures blockchain trust by integrating the current ways of characterizing users' trusting attitudes toward blockchains into a blockchain engineering framework, providing a holistic view of how trust is established in accordance with the pivotal elements underlying blockchain-based platforms. Section 4 conceptualizes blockchain trust in line with the distinctive feature of the trust phenomenon, which not only clarifies the normatively loaded nature of blockchain trust but also offers a perspective from which one can scrutinize the appropriateness granted to the normative values built into blockchain applications. Following this idea, section 5 examines two of the most promising values put forward by developers of the original type of blockchain that have the potential to ground rich trust decisions. It argues that more efforts should be undertaken with respect to improving the blockchain's decentralization and privacy-preserving capacity to make the system more trustworthy.

2 | WHY BLOCKCHAIN TRUST IS NEEDED, AND FOR WHOM

Much of the research into trust would agree that trust is risky.[3] By trusting, trustors have to take the risk of being let down, and they may lose whatever is entrusted to trustees. Yet, why do people not stay away from this vulnerable position? The reason may lie in the basic fact that every social being has limited cognitive and practical power; one is not capable of doing everything by oneself (Jones 2012). In everyday life, not only do we need to rely on others to satisfy

[3]Luhmann 1979; Baier 1986; Becker 1996; McLeod 2020.

our fundamental human needs (such as food, water, and shelter), we also need to rely on others in the acquisition of basic facts, scientific knowledge, and practical techniques (Hardwig 1991; Simon 2010). Trust provides a way of coping with our essential finitude by relying on others to help, learn, and cooperate, bringing both pragmatic and epistemic value to people in need.

In relationships underpinned by trust, trustors are optimistic about trustees' commitment and competence in doing certain things, but they understand that their trust might be frustrated and are willing to give discretionary power to the trustee (Baier 1986; Jones 2004). Implicit in this statement is the complex relationship between trust and knowledge. Although we use knowledge to place and withdraw trust (Simon 2010), "it is an important fact about trust that it cannot be given except by those who have only limited knowledge, and usually even less control, over those to whom it is given" (Baier 1992). From this perspective, the need for trust might be considered a sufficient condition for knowing that one lacks knowledge and control over the trustee. In other words, the lack of knowledge and control could be viewed as a threshold for creating the need for trust. By contrast, if someone were fully aware of or able to control another's action, the discretion and uncertainty involved in the relationship would cease to exist, and so would the need for trust.

This entangled relationship between trust and knowledge is particularly distinct when comparing laypeople's perceptions of controversial technologies with those of scientists. It is evident that public perceptions of controversial technologies tend to follow the "cognitive miser model," in which value predispositions (such as trust) and other heuristic entities play key roles rather than scientific knowledge (Fiske and Taylor 2013; Kahneman 2011). The situation is the same in cyberspace, where consumers have increasingly relied on heuristics for trust instead of being more rational (Pesch and Ishmaev 2019). On the one hand, this is because most people are incapable of rationally assessing the relevant information due to limits of time, resources, and technical expertise (Nickel 2013). On the other hand, trust provides an easier approach—a cognitive shortcut—for people to make decisions on whether to take the risk of doing something, which, as argued by Luhmann (1979, 8), functions as an effective form of reducing the complexity of living in society. By contrast, for scientists and experts who possess an in-depth understanding of the related technology, attitudes toward the technology are typically formed on the basis of domain-specific and general knowledge rather than trust (Ho et al. 2019). As a result, it is likely that resources endorsed by scientist communities can engender only limited epistemic confidence among laypeople, leading laypeople's judgment to false negatives; or that resources encouraging laypeople's trust cannot reach the same effect in scientist communities, leading laypeople's judgment to entail false positives. Both situations ask for efficient communication between the two groups. On the one side, public communication of scientific knowledge is needed. On the other, predictive resources used by "cognitive misers" should be critically examined and carefully incorporated into the design and communication process.

In this regard, whether there is a need for technology trust depends not only on the practical interest of using a particular technology but also on the extent to which the trustor knows about the technology. When the trustor has limited knowledge of and control over the trustee's action, either explicitly or implicitly, trust is usually needed for technology adoption, and the acquisition of knowledge can enhance the epistemic reason for trust. In the case of the Bitcoin blockchain, studies have shown that blockchain knowledge plays an epistemic role in enhancing users' trust (Sas and Khairuddin 2017; Ostern 2018).

While the word "user" is adopted almost everywhere to mean someone who uses the Bitcoin network for different purposes, according to the preceding discussion on the relationship between trust and knowledge, it causes confusion about whether a certain user has a need for trust. The existence of different classes of users with different levels of knowledge fundamentally explains why blockchain technology is sometimes considered trust-free or trustless and sometimes viewed as a trusted technology. On the one hand, the word "trustlessness" is commonly found in scientific research on blockchain technology, as authors often make unrecognized

and unspoken assumptions that the audience can fully understand how the system flows and how to control its functioning by developing and maintaining the codebase. In this case, trust is not a necessity, as little uncertainty and discretion from the system's side are considered. On the other hand, media reports and academic research from other communities often describe a blockchain as a trusted or trustable system potentially in place of third parties, since a complete understanding of every detail is hardly possible for people with no technical background. In this case, the performance of the system is considered not fully understandable, whether the mystery comes from the complex algorithms, the unknown developers and other network participants, or the novel applications. Therefore, even in the case of blockchain systems where trust is expected to be omitted by the developers, trust still plays an important role in shaping the opinions of people who have limited technical expertise.

Thus, an explicit clarification regarding who requires trust for blockchain adoption is needed. Following the analysis above, it is arguable that *normal users* who are actively or passively associated with the network but have limited, rudimentary, and fragmentary blockchain expertise can be defined as the main trustor group of the original blockchain. For example, active users can be those nodes who contribute their computation power to the execution of the system for economic purposes (that is, miners) but do not possess thorough blockchain knowledge; passive users can be those who merely use the system as an instrument for transactions and investments. By restricting trustors of blockchain-based systems to these specific groups, such a definition partially addresses the conceptual confusion over the understanding of the relationship between trust and blockchain. Equally important, it highlights the inherent risks of trusting complicated systems, which concern not only one's vulnerability with respect to the systems' discretion but also the epistemic-impairment position of assessing the actual trustworthiness of the systems (Ishmaev 2018; Nickel 2013). The vulnerable position of the trustor suggests serious moral concern over trust manipulation and mistrust associated with blockchain-based interactions, especially considering blockchains' irreversibility nature that leaves almost no room for redeeming the loss. I return to the discussion about the risk involved in trusting blockchains in section 5. Before that, let us look at what elements of blockchain-based systems are potential targets of users' trust.

3 | A FRAMEWORK FOR UNDERSTANDING THE STRUCTURE OF BLOCKCHAIN TRUST

Based on the preceding discussion, we can say that from a user-centered perspective trust is still needed in blockchain-based interactions. To understand users' blockchain trust systematically, this section explores the different elements that potentially invite trust in the original blockchain. These trust-inviting elements are then integrated into the blockchain engineering framework (BEF) outlined by Notheisen, Hawlitschek, and Weinhardt (2017). The resulting framework, called user-centered blockchain trust framework (BTF), explicates how the potential targets of trust are associated with the pivotal elements underlying blockchain-based platforms in multiple layers. In doing so, it provides a holistic view that captures the structure of users' reliance relations on blockchain applications.

As mentioned, existing research has argued that the original blockchain does not eradicate trust. Instead, it enables a shift of trust from third-party authorities (such as banks and governments) to the system's algorithms, the network's stakeholders, and the underlying economic mechanisms enabled by the blockchain's performances.

Consider, first, the trust shifted to the algorithms. An algorithm is a set of rules that give a sequence of operations for solving a specific type of problem. With features of "finiteness, definiteness, input, output, and effectiveness," algorithms generally provide some predictability that allows users to predict the system's outcome (Knuth 1997). As a result, increasing

epistemic authority is placed in algorithms to assess and predict the trustworthiness of diverse information sources. The idea of embedding epistemic authority in nonhuman agents such as algorithms has been argued for as a new form of trust that requires us to remain particularly vigilant about the algorithms' transparency (Simon 2010). It seems that blockchain technology has fostered such transparency to a great extent. In the Bitcoin network, while no node is given a privileged position to control the database, participants validate transactions and maintain the database collectively by using consensus algorithms and rules, which ultimately result in a singly valid, tamperproof, and publicly accessible database that can identify any double-spend attempt without the need for any third party (Nakamoto 2008). Thus, interactions processed by the network are based on algorithmic trust that allows users to predict the system's future behavior and act accordingly rather than trust between human agents (Swan and De Filippi 2017). In this regard, the underlying algorithms (for example, proof of work) are elements that directly invite users to trust the system. Trust in Bitcoin's algorithms, as Lustig and Nardi (2015) state, can be viewed as the trust placed in the legitimate power of the open-source codebase to verify information and direct human action, which is considered more predictable than opaque actions of large institutions.

Consider, second, the trust shifted to the network contributors. Although what users directly rely upon is the correct functioning of the blockchain system, the performance of the system is enabled by a chain of network contributors, mainly including developers and miners. At the protocol layer, when users adopt Bitcoin they put faith in the developer community and regard it as collective trustees who are responsible for maintaining the codebase (Mallard, Méadel, and Musiani 2014). As a result of lacking knowledge and time, normal users have to depend on coders who are able and willing to take responsibility for writing and verifying the blockchain code. Considering Bitcoin's open-source nature, while this coder community can be as large as whoever contributes to patching proposals, peer review, and testing, the trustworthiness of "core developers" (known as maintainers) is of importance, since they exert the decision-making power over judging the appropriateness of all pull requests. Trust developed at this layer can be personal or impersonal, depending on whether the trustor places trust in a particular, known developer or the developer community as a professional group. At the application layer, when users adopt Bitcoin, trust is distributed to a network of miners that contribute to validating and securing transactions collectively (Kasireddy 2018; Werbach 2018).

Consider, third, the trust shifted to underlying economic mechanisms enabled by the blockchain's performances. Examples of such mechanisms include new e-commerce models fueled by cryptocurrencies, online exchanges, and digital wallet services (Morisse 2015). Beyond the appearance potential, implicit in these mechanisms is a wide variety of sociotechnical factors that surpass the mathematical properties of algorithms and the direct efforts of the network contributors. As a distributed database technology initially designed to be a decentralized solution that aims to divorce centralized authorities, the Bitcoin system allows users to choose other transaction ways in addition to those provided by colossal banks, giant companies, and nation-states. It thus manifests immediate implications with respect to individual freedom and financial sovereignty, showing the preference for a cryptographic, decentralized, and transparent solution over a bureaucratic, centralized, and opaque system. In this regard, it can be argued that one important kind of motivation blockchain proponents have for using the system is that they share similar attitudes with the set of economic, political, and moral consequences underpinned by what the system affords. For example, as De Filippi and Loveluck (2016) argue, cryptography tools championed by cypherpunk groups and libertarians are utilized as a means of resistance to traditional authorities and human rights abuses. In trusting the economic mechanisms of the system, people expect that the associated sociotechnical values—such as decentralization, autonomy, and transparency—could be brought about by the system's performances and hence contribute to addressing the apprehension about information aggregation and power centralization caused by hierarchical structures.

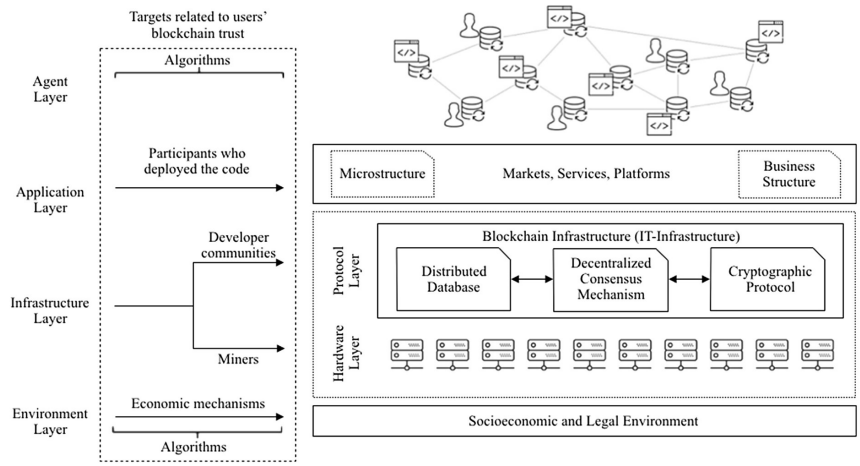

FIGURE 1 User-centered BTF (based on the BEF proposed by Notheisen, Hawlitschek, and Weinhardt 2017)

Before discussing the conception of blockchain trust developed on the basis of these elements, let me introduce the BEF as a tool to systematically structure how these trust-inviting elements are associated with the pivotal elements underlying blockchain-based platforms. Figure 1 illustrates the extended framework (that is, the BTF) resulting from a combination of the BEF and the content newly integrated by this paper (that is, the content in the dotted square).

The BEF creates a common approach for structuring the layers and pivotal elements of blockchain-based platforms in general. It consists of four layers. (1) *The environment layer* constructs economic, social, and legal foundations of the actions in other layers, which could correspond to the trust shifted to the underlying economic mechanisms brought about by the blockchains' performances. (2) *The infrastructure layer* comprises the protocol layer that defines the basic elements of blockchain infrastructure and the hardwire layer that connects a heterogeneous crowd of devices running the virtual machine. On the one side, this layer introduces trust into the developers' collective ability to write and verify the code. On the other, in the case where mining is necessarily involved, users need to make a trust judgment about whether miners will use their computation power to maintain but not manipulate the integrity of the network. (3) By integrating a full-fledged programming language known as the smart contract functionality, microeconomic designs such as autonomous market mechanisms and services can be built into *the application layer*. As the code in this layer is controlled by participants who deploy the code (Glaser 2017), the realized blockchain-based services also introduce trust to the application designers. (4) Governed by the applications' rules and characteristics, human and artificial agents can interact in *the agent layer* and process services available within the self-sufficient, closed ecosystem. By contrast, if the system is not self-sufficient and needs to be bound to external services and interfaces to realize certain functions, it pushes the trust issue back to those third-party authorities, such as exchanges and online markets. In addition, algorithmic trust pervades throughout the design and execution of the ecosystem, providing epistemic authority for users to reasonably expect the system's output and act accordingly.

The proposed user-centered BTF shows that trust shifted from traditional third-party authorities to blockchain-based platforms is multilayer structured. Intimately linked with different elements constituting blockchain systems, the targets that potentially encourage users' trust include the blockchain algorithms, the relevant stakeholders (that is, protocol developers, miners, and application designers), and the economic mechanisms potentially brought about by the system's performances. The resulting framework contributes to first systematically

structuring the elements of blockchain-based systems that potentially invite users' trust, providing a way for developers and users to reflect on the actual trustworthiness of these elements when interacting with a blockchain system. For less demanding accounts that use trust as reliance, trust can probably reside in any or a combination of the trust-inviting elements identified above. In this way, trusting a blockchain-based system simply means that people expect that the system will perform reliably for achieving specific goals, with no additional requirements. If, however, we delve into the reliance relationship grounded in the third category of trust-inviting elements, a more demanding account of trust seems to be applicable to the systems, which enables us to take a step back from rational-choice accounts and begin to think about the sociotechnical values involved in blockchain trust.

4 | A THEORETICAL EXPLORATION OF BLOCKCHAIN TRUST

As mentioned, more demanding accounts of trust hold that trust is a distinctive concept that contains some morally, normatively, or affectively loaded elements, going beyond mere expectations about the trustee's reliability. For these accounts, such elements essentially explain why one would like to take a leap of faith and form interactions under uncertainty. For these accounts, a blockchain is a viable target of trust if and only if one's attitudes toward the trust-inviting elements of blockchain-based systems can bear a family resemblance to the distinctive feature of trust. These accounts show the unique explanatory power of trust, which allows one to take the risk of doing interactions in the lack of knowledge and control. In an effort to explore this question, this section conceptualizes blockchain trust in line with the normative accounts of trust, with the aim of providing a theoretical foundation on which the sociotechnical factors built into blockchains can be based.

To understand whether a rich conception of trust can be applied to blockchain systems, it is imperative to first look at the nature of the trust notion and trust in technologies as an extension of trust. In philosophical studies on the concept of trust, human-to-human trust is regarded as the original and dominant paradigm of trust relations; the possibility of trust arising between humans and technologies is often overlooked or considered implausible. This situation can be primarily attributed to the widely shared assumption that the trustee must be committed to the trustor's dependency or vulnerability (McLeod 2020). As Jones (1996) argues, only those who have a will can be the object of trust. This will-based assumption resonates with the mainstream ethical theory regarding ethics as an activity that is unique to moral agents; and in philosophy, such moral agency is often thought to be restricted to human subjects due to the possession of some characteristics—such as free will, intentionality, and emotions—that indicate the capacity to be responsible for decisions and practices (Verbeek 2011; van de Poel 2020).

The most influential theoretical foundation of this assumption is Baier's (1986) classic exposition about moral trust. For this account, to trust is to rely on another's good will and competence to "pursue, promote, preserve, and protect" certain of one's goods and vulnerabilities (Alfano 2016). Here the relevant characteristics of the trustee show that one cares about, or at least will not use the discretionary power to harm, the things that are valued by the trustor. Grown out of this statement, much of the literature challenges the viability of applying the trust notion to technologies by arguing either that technologies can only be paired with strategic reliance or that this attempt is merely an extension of interpersonal trust (Nickel 2013). Regarding the former aspect, Pettit (2004), for example, argues that machines and technological systems cannot be targets of trust, since they lack consciousness and agency manifesting any will concerning the trustor. Regarding the latter aspect, Pitt (2010) and Cook (2010), for example, argue that trusting a specific technology is eventually an issue of trusting a certain

person to do certain things. These two challenges, however, are fraught with the problem that a will-based account is not omnipotent. Although accounts built on the trustee's specific motives might be rich in justifying certain forms of trust (such as trust in friends), O'Neill (2002) has pointed out that the trustee's good will is neither sufficient nor necessary for understanding a wider variety of trust forms that are more diffuse and complex—such as trust in professionals and strangers.

Differently, normative accounts of trust argue that the distinctive feature of trust lies in one's normative expectations of another's responsible behavior. Such expectations are grounded in our moral standards presumably shared with others, such as the shared expectations that promises should be honored and duties should be performed (Hollis 1998; Jones 2004). Trust here is more like a stance that a trustor holds toward trustees, expecting that the trustees will do what they should while leaving their motivations open to different contexts—be it the desire for good will, good repute, a good character, the fear of sanctions, the pressure of social constraints (Walker 2006, 81). As Simpson (2012) contends, this sort of account is more suitable for characterizing more distant relations in a communal sense, such as trust in people with obligations (such as firefighters and doctors). For these accounts, it is the normatively loaded expectations that distinguish the trustor's trust from reliance and the trustee's trustworthiness from reliability.

There are two things that are clear about trust. First, the way of depicting trust is strongly dependent on the particular trust relation in question, since different trustees and situations to which the trust concept applies vigorously shape the actual meaning of trust (Simon 2013). Thus, rather than giving a single established definition that is amenable to all counterexamples, it seems more plausible to focus on the variable value of trust in context (Simpson 2012; van den Hoven 2008). Relatedly, the second thing about trust concerns the multifaceted nature of trust. As Baier pointed out in 1994 already, if trust is a cognitive, affective, or conative phenomenon, it is all three phenomena. This means that although different situations might emphasize different facets of trust, the nature of trust is not black and white (Baier 1994). Acknowledging that trust is a notion that is both contextual and multifaceted, researchers have explored various frameworks in support of the arguments that trust can be invited by and placed in not only physical persons but also roles and professionals (Becker 1996; Pellegrino 1991), institutions (Townley and Garfield 2013), and technologies (Nickel 2013; Coeckelbergh 2012; Nguyen 2023). Although built on different considerations, these accounts shed light on exploring the explanatory power of trust in the context of abstractly characterized entities.

A particularly interesting and robust trend is the normative accounts mentioned, providing a relatively consistent way to understand trust when applied to multifarious trustees. For example, Walker (2006, 83) argues that, besides people with obligations, people also hold normative expectations toward institutions and businesses, which are more like a default stance we habitually stand in with respect to the good state of their services. When applied to technological systems, such expectations make trustors believe that they are entitled to what the systems are supposed to do (Nickel 2013). Nickel (2021) further argues that here the systems are the direct targets of trust, whereas the trust developed with engineers and designers behind the systems is considered indirect, impersonal, and abstract. In this sense, similar to the case with roles and professionals, it appears that institutions and technological systems are also embodiments of certain moral and normative standards that can invite people's shared beliefs about their performances. From this perspective, it seems natural to say that trust can be placed in these abstractly characterized entities even though we are not aware of those strangers who occupy the roles of doctor, banker, or software engineer. What we generally rely upon is just the standard performances that we normatively expect of those in that profession, institution, or technological system rather than any unique tie or personal concern a particular human agent may give to us. This broader conception seems to enable trust, though in a limited sense, to be created without any complete agential state of the trustee being assumed.

In this regard, it might be said that the essential distinction between the normative expectation we hold toward a specific person and the one we hold toward an abstractly characterized entity lies in the different sources that ground such expectations. While the normativity of the specific person comes mainly from the moral understandings we presumably shared with others (Walker 2006, 66), that of the abstractly characterized entity is suggested by a wide range of sociotechnical factors embedded in and embodied by these entities' performances—such as different values and norms, laws and regulations, and codes of conduct. This claim resonates with the broader philosophical view that technologies are value laden and can inform normatively significant decisions and practices (Verbeek 2011). This indicates that technological systems can be designed in a way that shapes and encourages users to expect certain performances of the system. For instance, if a blockchain system is designed to promote individual freedom and financial sovereignty, it is reasonable for one to normatively expect that the system's patterns of action will display explicit cues that reflect these features. Relying on such a system, in this sense, carries significant ethical implications concerning the moral acceptability and social desirability of the system.

The preceding analysis of trust in technologies provides a feasible way to reach a relatively rich conception of blockchain trust, highlighting the importance of the sociotechnical values embedded in and potentially embodied by blockchains' performances. Accordingly, the attitude of relying on these values is used not just in a *predictive* sense, in that one believes that such values will be brought about by the systems' performances, but also in an *evaluative* sense, in that one thinks that these values are desirable things that should be folded into the systems. In a word, combined with the trust shifted to the algorithms and network contributors, if a person X trusts a blockchain system Y in a normatively loaded way, X relies not only on Y's functionality but also on certain moral, economic, political, or social values built into Y's performances that are considered appropriate to X.

Betrayal, in this case, is not about how our trust has been frustrated by others' commitments but about how we feel alienated from the *appropriateness* that we grant to the normative consequences potentially brought about by the system's performances. Just as promises are not always honored, trust-inviting cues are not always reliable, and they do not lead to the fact that the related entity is indeed trustworthy. Recalling the trustor's epistemic vulnerability discussed earlier, when lacking sufficient knowledge, time, and resources to understand a complex system, users might hastily and carelessly bear an unquestioning attitude toward the system and thus trust more than the system's trustworthiness warrants, ultimately resulting in misplacement of trust (Nguyen 2020). Thus, we need to assess such appropriateness in order to approach more warranted trust decisions. By presenting the challenges faced by achieving two fundamental normative values of the original blockchain's underlying economic mechanism, the focus of the rest of this paper shifts from the blockchain's trust-inviting elements to the creation of trust-deserving or trustworthy applications.

5 | EXAMINING THE NORMATIVE VALUES RELATED TO BLOCKCHAIN TRUST

As a first step in assessing the appropriateness of what people normatively expect from the blockchain's performances, two questions seem to be germane: (1) whether the embedded values are realized in the systems' applied context and (2) whether they are realized without conflicting with other important values. In this section, these two questions are discussed via an examination of two of the most promising values put forward by the original blockchain: namely, decentralization and transparency. On the one hand, these values are inherent in this technology's basic infrastructure and could bring about significant sociotechnical implications, such as making transactions without relying on third-party authorities and their

bureaucratic processes. On the other hand, the system is certainly appealing to those who are eager for a free-floating digital economy mechanism controlled by no human-run institutions and no person in particular. It thus excites the proponents' interests and provides excellent reasons for them to provide their initial trust. From the proponents' perspective, it might be said that the above values behind the blockchain's economic mechanisms are the potential sources of the blockchain's normative desirability that can encourage rich and justified blockchain trust decisions. In what follows, this blockchain trust is examined, and the analysis shows that there is a tension between the pressing values that are intended to be achieved by developers and the predicament situations caused by current blockchain implementations.

5.1 | Decentralized network versus power centralization

While a top-down, centralized authority and hierarchical structures provide useful means of facilitating valid social interactions in modern societies, they are often fraught with a crisis of trust due to data aggregation, undue censorship, surveillance, and consequent moral apprehension such as the erosion of individual freedom, privacy, and autonomy (Chaum 1985; Al-Saqaf and Seidler 2017). By contrast, data validity of the original blockchain is fueled by a series of mathematical rules and executed on a large network of computing devices, peers, and developer communities. By replacing the role and functions of third-party authorities with technical settings, the decentralized system provides another choice that not only improves individual freedom but also mitigates the moral issues engendered by traditional authorities. The system thus carries an explicit normative message that the decentralized network is considered more desirable than the traditional mechanism. Such a message can invite the trust of people who favor the moral desirability and other effects potentially brought about by the system's decentralization promise. Based on the rich conception of blockchain trust discussed above, to see whether trust relations grounded in this striking value are well grounded we need to take a look at the real-world performances of blockchain applications for realizing this promise.

As Reijers et al. (2021) point out, the governance of blockchain-based systems can be divided into two categories: on-chain governance, where interactions are solely determined by the rule of code, and off-chain governance, where the reference community might be affected by self-regulation and exogenous rules such as laws and regulations. Yet, it has been argued that an inherent degree of centralization exists in both the enforcement of rules and the collective governance of blockchains (Azouvi, Maller, and Meiklejohn 2018). First, at the application layer, it is uncertain to what extent the network's decentralization promise can be realized in the fact that several mining conglomerates control a considerable amount of computing power.[4] The monopoly of mining makes it possible for conglomerates to collude with each other and exploit the system. A decentralized network, in this sense, does not guarantee decentralized power (Brekke 2019).

Second, the governance structure of the protocol layer is also quite centralized. Research has shown that the same developer has created around 7 percent of all Bitcoin documents, and half of all the comments in its GitHub repository were written by only eight contributors (Azouvi, Maller, and Meiklejohn 2018). While the codebase of the project is maintained by only a few developers, vital decisions within the community are reached through the exchange of opinions among members on mailing lists without any transparent decision-making process being known by the multitude of users (Gervais et al. 2014). Such a situation causes concern over the appropriateness that users grant to the blockchain's decentralized setup, since it uses a few developers to replace the complex social roles previously filled by a wide range of people and institutions. This seems

[4]For Bitcoin's hashrate distribution, see https://www.blockchain.com/en/pools.

incompatible with the significant role of developers and the consequence their actions may cause. Compared to the well-established legislation imposed on traditional third parties (such as corporations and banks), explicit rules and procedures that can be imposed on developers to guarantee the security of the network are scarce. Although there is always a possibility of a fork, just as how the Ethereum blockchain recovered from the attack on the decentralized autonomous organization (DAO), more explicit strategies for self-governance are required, especially strategies that can inform communities about how to deal with crises in a systematic way.[5]

The increasing power centralization of the two layers together with some centralized services surrounding the Bitcoin system (such as web wallets and exchange platforms) indicate that a decentralized infrastructure does not necessarily lead to decentralization of power. Also, it makes the question of who controls the system of crucial ethical and political importance, since these people impact the sociotechnical implications of the whole ecosystem (Reijers and Coeckelbergh 2018). These issues about power centralization in relation to how Bitcoin is implemented today make blockchain's decentralization promise questionable. Nevertheless, one should note that threats of monopoly on both layers are not equal to actual monopoly, and in fact this peer-to-peer network has survived countless cyberattacks. To say the least, there is always the possibility for the community to alter the code by voting. But we should also be aware of the risks involved in the blockchain's decentralization issue as well as the associated problems related to forks. This means more solutions are needed to make the blockchain a safer place more deserving of trust, where normal users' rights and assets can indeed be protected via the immutable and autonomous code.

5.2 | Data transparency versus privacy concerns

Transparency is arguably another important value that often invites users to trust blockchains in a normatively loaded way. The normative implications of transparency can be understood from the crucial role played by information transparency in promoting people's trust in traditional institutions. As characteristics of information, transparency dimensions, including information disclosure, clarity, and accuracy, are positively related to an institution's trustworthiness to encourage trust (Schnackenberg, Andrew, and Tomlinson 2016). But given the privileged position centralized authorities played in data processing (such as the recording, collection, storage, and using of data), users are almost always in a passive and vulnerable position caused by information asymmetry and its knock-on effects. In this regard, an open-source, public blockchain appears to be an ideal medium to facilitate transparency by permitting users fair access right to the database and source code. Rather than relying on the good will and sense of responsibility of centralized authorities and relevant individuals, transparency in blockchain-based systems is guaranteed by network protocols that directly mitigate the situation of information asymmetry.

One flip side of such blockchains' transparent and immutable nature is the challenge posed to private data protection (De Filippi 2016). As all Bitcoin transactions are publicly available and traceable but not fully anonymous, they can reveal the identity of coin owners when linked to other information or datasets. Given the risk of reidentification and privacy loss, it can be argued that the trust judgment on the public accessibility of a blockchain should be evaluated together with the system's capacity for coping with privacy-related issues. Despite the benefits enabled by blockchains' peculiarities, it is thus important for users to understand the privacy issue involved, particularly considering the fact that in the context of blockchain systems no one is legally responsible for users' loss of privacy.

At the same time, blockchain applications are trying to solve this dilemma in different ways. Take the case of the Enigma project, which is seen as one of the most promising solutions

[5]For more information about the DAO attack, see https://cryptobullsclub.com/the-dao-attack/.

for preserving privacy in the blockchain context. The way that Enigma addresses the privacy concern is to use a cryptography tool called secure Multi-Party Computation that allows data to be split, encrypted, and computed by nodes at a second layer off the ledger (Zyskind, Nathan, and Pentland 2015). This means that nodes of the network could verify smart contract computations without seeing any decrypted data. Although some metadata are still required to be stored in the ledger to keep track of data ownership and the distribution of data, this solution is much more privacy friendly than the original blockchain. Furthermore, with the establishment of blockchain-based data markets, projects like Enigma purport not merely to protect privacy but also to unlock new value by allowing data owners to share, trade, and get rewards from their private data. This arouses the interest of many blockchain proponents who would like to capture the market value of their data. Unlike fiat money, data are non-rivalrous and non-fungible, meaning that a given piece of data could be used for multiple purposes concurrently and cannot be replaced by another piece of data (MIT Technology Review Custom and Oracle 2016). For many crypto fans, these unique characteristics of data are thus intriguing trust-inviting features. Nevertheless, the privacy risk of this market-centric solution should not be ignored, which may exacerbate the situation of data secondary usage and other informational-based harm (van den Hoven 2008).

To conclude, based on the argument that the normative values inserted into blockchains' infrastructure are building blocks of users' blockchain trust, this section has examined two core values that possibly invite justified trust decisions. These values sit well with the proponents' interests and expectations and are thus valid trust-inviting cues for users' initial trust. Yet, the promise of decentralization is shown to be restricted by how Bitcoin is implemented today in its sociotechnical context, and the blockchain's transparency feature might cause risks to users' privacy. Both aspects imply that, on the one hand, trust decisions invited by these promised values should be carefully reflected before bestowing appropriateness. On the other hand, the blockchain community should ensure greater decentralization and robust privacy to make the digital ecosystem more trustworthy.

6 | CONCLUSION

Whom and what can and should we trust? This is a fundamental philosophical question, as it forms the background of nearly all social cooperation—from dyadic interactions in situations well modeled as prisoners' dilemmas and stag hunts to large-scale, longitudinal interactions between anonymous groups. Nevertheless, arriving at a well-grounded trust decision is a non-trivial task, especially when it comes to complex and novel systems such as blockchains. On the trustor's side, the epistemic vulnerability of users impedes their collecting and extracting accurate information from a vast number of resources online, creating barriers to capturing a relatively complete picture of the situation. On the trustee's side, while blockchain infrastructure has the potential to revolutionize the way we interact, the entire blockchain industry is still in its infancy. A number of internal and external uncertainties regarding the moral concerns, legal constraints, and technical limitations of blockchain implementations add unforeseen dynamics to the trust decisions we are able to make at the moment (Swan 2015).

To explicate the role and risk of trust related to blockchain-based interactions, this paper has critically engaged with the concept referred to as blockchain trust. It provides a philosophical analysis of the notion of trust in the context of blockchain technology, encompassing four aspects: (1) a clarification of the trustor group of blockchain technology; (2) a systematic analysis of the elements potentially inviting users' blockchain trust; (3) an investigation into how the distinctive feature of the notion of trust can be understood in blockchain context; and (4) a reflection on the appropriateness one may accord the core values built into blockchains' potential. The upshot of the paper is that more effort should be made to improve the

blockchain's decentralization and privacy-preserving capacity in order to make the system more trustworthy and protect users' vulnerability, which can also provide a positive feedback loop between trust-inviting features and trust-deserving results. This study only starts the inquiry into justified blockchain trust. Future research could build on the conceptual analysis provided and systematically explore self-governance solutions to approach more warranted trust in the context of blockchain technology.

REFERENCES

Alfano, Mark. 2016. "The Topology of Communities of Trust." *Russian Sociological Review* 15, no. 4: 30–56.

Al-Saqaf, Walid, and Nicolas Seidler. 2017. "Blockchain Technology for Social Impact: Opportunities and Challenges Ahead." *Journal of Cyber Policy* 2, no. 3: 338–54.

Azouvi, Sarah, Mary Maller, and Sarah Meiklejohn. 2018. "Egalitarian Society or Benevolent Dictatorship: The State of Cryptocurrency Governance." In *Financial Cryptography and Data Security: FC 2018*, edited by Aviv Zohar, Ittay Eyal, Vanessa Teague, Jeremy Clark, Andrea Bracciali, Federico Pintore, and Massimiliano Sala, 127–43. Berlin: Springer.

Baier, Annette C. 1986. "Trust and Antitrust." *Ethics* 96, no. 2: 231–60.

Baier, Annette C. 1992. "Trusting People." *Philosophical Perspectives* 6: 137–53.

Baier, Annette C. 1994. "Trust and Its Vulnerabilities." In *Moral Prejudices: Essays on Ethics*, edited by Annette C. Baier, 130–51. Cambridge, Mass.: Harvard University Press.

Becker, Lawrence C. 1996. "Trust as Noncognitive Security About Motives." *Ethics* 107, no. 1: 43–61.

Brekke, Jaya K. 2019. "Disassembling the Trust Machine: Three Cuts on the Political Matter of Blockchain." Doctoral thesis, Durham University. http://etheses.dur.ac.uk/13174/. Accessed 19 June 2020.

Chaum, David. 1985. "Security Without Identification: Transaction Systems to Make Big Brother Obsolete." *Communications of the ACM* 28, no. 10: 1030–44.

Coeckelbergh, Mark. 2012. "Can We Trust Robots?" *Ethics and Information Technology* 14, no. 1: 53–60.

Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge, Mass.: Belknap Press of Harvard University Press.

Cook, S. D. Noam. 2010. "Making the Technological Trustworthy." *Knowledge, Technology and Policy* 23, nos. 3–4: 455–59.

De Filippi, Primavera. 2016. "The Interplay Between Decentralization and Privacy: The Case of Blockchain Technologies." *Journal of Peer Production*, no. 7.

De Filippi, Primavera, and Benjamin Loveluck. 2016. "The Invisible Politics of Bitcoin: Governance Crisis of a Decentralized Infrastructure." *Internet Policy Review* 5, no. 3.

Fiske, Susan T., and Shelley E. Taylor. 2013. *Social Cognition: From Brains to Culture*. London: Sage.

Gambetta, Diego. 1988. "Can We Trust Trust?" In *Trust: Making and Breaking Cooperative Relations*, edited by Diego Gambetta, 213–37. Oxford: Basil Blackwell.

Gervais, Arthur, Ghassan O. Karame, Vedran Capkun, and Srdjan Capkun. 2014. "Is Bitcoin a Decentralized Currency?" *IEEE Security and Privacy* 12, no. 3: 54–60.

Glaser, Florian. 2017. "Pervasive Decentralisation of Digital Infrastructures: A Framework for Blockchain Enabled System and Use Case Analysis." *In Proceedings of the 50th Hawaii International Conference on System Sciences*, edited by T. Bui and R. Sprague Jr., 1543–52. University of Hawaii.

Hardwig, John. 1991. "The Role of Trust in Knowledge." *Journal of Philosophy* 88, no. 12: 693–708.

Ho, Shirley S., Alisius D. Leong, Jiemin Looi, Liang Chen, Natalie Pang, and Edson Tandoc Jr. 2019. "Science Literacy or Value Predisposition? A Meta-Analysis of Factors Predicting Public Perceptions of Benefits, Risks, and Acceptance of Nuclear Energy." *Environmental Communication* 13, no. 4: 457–71.

Hollis, Martin. 1998. *Trust Within Reason*. Cambridge: Cambridge University Press.

Holton, Richard. 1994. "Deciding to Trust, Coming to Believe." *Australasian Journal of Philosophy* 72, no. 1: 63–76.

Ishmaev, Georgy. 2018. "Rethinking Trust in the Internet of Things." In *Data Protection and Privacy: The Internet of Bodies*, edited by P. De Hert, S. Gutwirth, R. van Brakel, and R. Leenes, 203–30. Oxford: Hart.

Jacobs, Mattis. 2020. "How Implicit Assumptions on the Nature of Trust Shape the Understanding of the Blockchain Technology." *Philosophy and Technology* 34: 573–87.

Jones, Karen. 1996. "Trust as an Affective Attitude." *Ethics* 107, no. 1: 4–25.

Jones, Karen. 2004. "Trust and Terror." In *Moral Psychology: Feminist Ethics and Social Theory*, edited by P. DesAutels and M. U. Walker, 3–18. Lanham, Md.: Rowman and Littlefield.

Jones, Karen. 2012. "Trustworthiness." *Ethics* 123, no. 1: 61–85.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Toronto: Doubleday Canada.

Kasireddy, Preethi. 2018. "What Do We Mean by 'Blockchains Are Trustless'?" https://www.preethikasireddy.com/post/what-do-we-mean-by-blockchains-are-trustless#:~:text=When%20we%20say%20blockchains%20are%20%E2%80%9Ctrustless%2C%E2%80%9D%20what%20we,a%20consensus%20on%20what%20the%20canonical%20truth%20is. Accessed 20 February 2019.

Knuth, Donald Ervin. 1997. *The Art of Computer Programming*, vol. 1. Boston: Addison-Wesley.

Luhmann, Niklas. 1979. *Trust and Power*. Chichester: John Wiley and Sons.

Lustig, Caitlin, and Bonnie Nardi. 2015. "Algorithmic Authority: The Case of Bitcoin." *In 48th Hawaii International Conference on System Sciences*, edited by T. X. Bui and R. H. Sprague Jr., 743–52. University of Hawaii.

Mallard, Alexandre, Cécile Méadel, and Francesca Musiani. 2014. "The Paradoxes of Distributed Trust: Peer-To-Peer Architecture and User Confidence in Bitcoin." *Journal of Peer Production* 4: 1–10.

Mcleod, Carolyn. 2020. "*Trust*." https://plato.stanford.edu/entries/trust/. Accessed 6 June 2021.

MIT Technology Review Custom, and Oracle. 2016. "The Rise of Data Capital." *MIT Technology Review*. http://files.technologyreview.com/whitepapers/MIT_Oracle+Report-The_Rise_of_Data_Capital.pdf?_ga=2.227337053.1748626409.1654567372-2146615841.1643334756. Accessed 8 June 2020.

Möllering, Guido. 2006. *Trust: Reason, Routine, Reflexivity*. Amsterdam: Elsevier.

Morisse, Marcel. 2015. "Cryptocurrencies and Bitcoin: Charting the Research Landscape." In *AMCIS 2015 Proceedings*. AIS Electronic Library, Association for Information Systems.

Nakamoto, Satoshi. 2008. "Bitcoin: A Peer-To-Peer Electronic Cash System." *Bitcoin*. https://bitcoin.org/bitcoin.pdf. Accessed 9 June 2019.

Nguyen, C. Thi. 2023. "Trust as an Unquestioning Attitude." In *Oxford Studies in Epistemology 7*, edited by T. S. Gendler, J. Hawthorne, and J. Chung. Oxford: Oxford University Press.

Nickel, Philip J. 2013. "Trust in Technological Systems." In *Norms in Technology*, edited by M. J. De Vries, S. O. Hansson, and A. W. M. Meijers, 223–37. Dordrecht: Springer.

Nickel, Philip J. 2021. "Trust in Engineering." In *The Routledge Companion to Philosophy of Engineering*, edited by D. P. Michelfelder and N. Doorn, 494–505. London: Routledge.

Notheisen, Benedikt, Florian Hawlitschek, and Christof Weinhardt. 2017. "Breaking Down the Blockchain Hype: Towards A Blockchain Market Engineering Approach." In *Proceedings of the 25th European Conference on Information Systems*, edited by I. Ramos, V. Tuunainen, and H. Krcmar, 1062–80. Guimarães, Portugal.

O'Neill, Onora. 2002. *Autonomy and Trust in Bioethics*. Cambridge: Cambridge University Press.

Ostern, Nadine. 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.

Pellegrino, Edmund. 1991. "Trust and Distrust in Professional Ethics." In *Ethics, Trust, and the Professions*, edited by E. D. Pellegrino, R. M. Veatch, and J. P. Langan, 69–85. Washington, D.C.: Georgetown University Press.

Pesch, Udo, and Georgy Ishmaev. 2019. "Fictions and Frictions: Promises, Transaction Costs and the Innovation of Network Technologies." *Social Studies of Science* 49, no. 2: 264–77.

Pettit, Philip. 2004. "Trust, Reliance and the Internet." *Analyse und Kritik* 26, no. 1: 108–21.

Pitt, Joseph C. 2010. "It's Not About Technology." *Knowledge, Technology and Policy* 23, nos. 3–4: 445–54.

Reijers, Wessel, and Mark Coeckelbergh. 2018. "The Blockchain as a Narrative Technology: Investigating the Social Ontology and Normative Configurations of Cryptocurrencies." *Philosophy and Technology* 31, no. 1: 103–30.

Reijers, Wessel, Iris Wuisman, Morshed Mannan, Primavera De Filippi, Christopher Wray, Vienna Rae-Looi, Angela Cubillos Vélez, and Liav Orgad. 2021. "Now the Code Runs Itself: On-Chain and Off-Chain Governance of Blockchain Technologies." *Topoi* 40, no. 4: 821–31.

Sas, Corina, and Irni Eliana Khairuddin. 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*, 6499–6510. New York: Association for Computing Machinery.

Schnackenberg, Andrew K., and Edward C. Tomlinson. 2016. "Organizational Transparency: A New Perspective on Managing Trust in Organization-Stakeholder Relationships." *Journal of Management* 42, no. 7: 1784–1810.

Simon, Judith. 2010. "The Entanglement of Trust and Knowledge on the Web." *Ethics and Information Technology* 12, no. 4: 343–55.

Simon, Judith. 2013. "Trust." In *Oxford Bibliographies in Philosophy*, edited by D. Pritchard. New York: Oxford University Press.

Simpson, Thomas W. 2012. "What Is Trust?" *Pacific Philosophical Quarterly* 93, no. 4: 550–69.

Simser, Jeffrey. 2015. "Bitcoin and Modern Alchemy: In Code We Trust." *Journal of Financial Crime* 22, no. 2: 156–69.

Swan, Melanie. 2015. *Blockchain: Blueprint for a New Economy*. Sebastopol, Calif.: O'Reilly Media.

Swan, Melanie, and Primavera De Filippi. 2017. "Toward a Philosophy of Blockchain: A Symposium: Introduction." *Metaphilosophy* 48, no. 5: 603–19.

Townley, Cynthia, and Jay L. Garfield. 2013. "Public Trust." In *Trust: Analytic and Applied Perspectives*, edited by P. Mäkelä and C. Townley, 95–108. Amsterdam: Rodopi.
van den Hoven, Jeroen. 2008. "Information Technology, Privacy, and the Protection of Personal Data." In *Information Technology and Moral Philosophy*, edited by J. van den Hoven and J. Weckert, 301–22. Cambridge: Cambridge University Press.
van den Hoven, Jeroen, Johan Pouwelse, Dirk Helbing, and Stefan Klauser. 2019. "The Blockchain Age: Awareness, Empowerment and Coordination." In *Towards Digital Enlightenment*, edited by D. Helbing, 163–166. Cham: Springer.
van de Poel, Ibo. 2020. "Embedding Values in Artificial Intelligence (AI) Systems." *Minds and Machines* 30, no. 3: 385–409.
van Lier, Ben. 2017. "Can Cyber-Physical Systems Reliably Collaborate Within a Blockchain?" *Metaphilosophy* 48, no. 5: 698–711.
Velasco, Pablo R. 2017. "Computing Ledgers and the Political Ontology of the Blockchain." *Metaphilosophy* 48, no. 5: 712–26.
Verbeek, Peter Paul. 2011. *Moralizing Technology: Understanding and Designing the Morality of Things*. Chicago: University of Chicago Press.
Walker, Margaret Urban. 2006. *Moral Repair: Reconstructing Moral Relations After Wrongdoing*. Cambridge: Cambridge University Press.
Weckert, John. 2005. "Trust in Cyberspace." In *The Impact of the Internet on Our Moral Lives*, edited by R. J. Cavalier, 95–120. Albany: SUNY Press.
Werbach, Kevin. 2018. "Trust, But Verify: Why the Blockchain Needs the Law." *Berkeley Technology Law Journal* 33: 487ff.
Zyskind, Guy, Oz Nathan, and Alex Pentland. 2015. "Enigma: Decentralized Computation Platform with Guaranteed Privacy." arXiv preprint: 1506.03471. https://arxiv.org/abs/1506.03471. Accessed 6 August 2017.

How to cite this article: Teng, Yan. 2022. "What does it mean to trust blockchain technology?" *Metaphilosophy* 00: 1–16. https://doi.org/10.1111/meta.12596.