1
What Good Is Blockchain?

The phone rang in my office. It was late summer 2017. I answered and was greeted on the other end of the line by the voice of one my colleagues in the Computer Science Department asking, "Is blockchain really a *thing*?" I have been asked this question any number of times since then, though each time it takes a slightly different form. James Mickens, a Harvard computer scientist, produced a video in 2018 entitled *Blockchains Are a Bad Idea* that is a variation on the same theme. In the video, Mickens points out that many of the features of blockchains can be provided by existing technologies. To many, blockchain technology seemingly offers nothing new, since, by and large, it presents an assemblage of preexisting theories, algorithms, and mathematics, as I will discuss in Section 1.3, and a computationally inefficient one at that (see, e.g., Truby, 2018; Li et al., 2019)![1]

Observed from a purely computational or technical (in the sense of information and communications technology) perspective, it is not easy to see what all the fuss is about when it comes to blockchains, nor why there should be such interest in them. As Mickens argues in his video, blockchain systems, such as Bitcoin, have features and capabilities that can be provided by existing systems: tamper resistance can be provided by digital signatures (discussed in

[1] Bitcoin mining consumes an enormous amount of electricity. According to the Bitcoin Energy Consumption Index (see, e.g., Digiconomist, 2021), a single Bitcoin transaction consumes the equivalent of the carbon footprint of 664,375 Visa transactions (299.76 $kgCO_2$) and the same amount of power as the average United States household usage over 21.63 days (631.08 kWh). De Vries (2018, p. 801) states that "The Bitcoin network can be estimated to consume at least 2.55 gigawatts of electricity currently, and potentially 7.67 gigawatts in the future," making it comparable with countries such as Ireland (3.1 gigawatts) and Austria (8.2 gigawatts) in energy consumption based on 2018 data. On this point, see also Das and Dutta (2020). Owing to the amount of energy needed to mine Bitcoin and cool the mining equipment, miners tend to gravitate their operations to places where they can obtain electricity relatively cheaply and where it is easier to keep their equipment cool (Bjarnason, 2019; Baydakova, 2021). Unscrupulous miners have also been known to steal the electricity they need (Nadeau, 2020).

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

Section 1.3) with or without blockchain, as can the ability to prove a claim or to achieve non-repudiability of a transaction (ISO, 2018a, s. 3.48); message ordering can be achieved through the use of hash pointing (discussed in Section 1.3) without resorting to blockchain; and highly available storage needs can be handled by commercial cloud storage (Mickens, 2018). Yet, as an archival scientist – someone who studies the theory of recordkeeping and the long-term preservation of authentic records – blockchain makes sense to me. Even if I doubt some of the claims I hear about it, I see it as a response to a perceived erosion of society's "fact infrastructure" in an age of disinformation and disorders of social trust. It is this perspective on blockchain technology that I will explore in this volume.

1.1 Blockchain Is Meaningless

My colleague's question about blockchain came, not unreasonably, at the peak of what has been described as the blockchain "initial coin offering hype cycle," which was ramping up to its late 2017, early 2018 crescendo. At the time, many were touting blockchain (and their own initial coin offerings, the cryptocurrency community's equivalent to initial public offerings) as a solution to all the world's problems. To illustrate the zeitgeist of the time, technology writer Alex Hern (2016) wrote a (tongue-in-cheek) piece for *The Guardian* in 2016 entitled, "Blockchain: The Answer to Life, the Universe and Everything?" that appropriately began with the sentence, "Have you heard the good news? The blockchain is here – and it's going to save everything."

Don and Alex Tapscott's 2016 book, *Blockchain Revolution: How the Technology Behind Bitcoin is Changing Money, Business, and the World*, set out a vision of how blockchain could be used to transform and change the world for the better by tracking the provenance of digital and real-world assets, banking the unbanked, and unleashing new businesses. Given the lack of real-world evidence at the time, many were (and remain) skeptical, as outlined in, for example, David Gerard's critical 2015 book *Attack of the 50 Foot Blockchain: Bitcoin, Blockchain, Ethereum & Smart Contracts.* At the same time, few really understood what the term blockchain meant. Adrianne Jeffries (2018), writing for *The Verge* in early 2018, described a blockchain Tower of Babel in which everyone was speaking their own incomprehensible blockchain language, concluding that "'Blockchain' is meaningless."

How is it that blockchains are meaningless to some, while others see their potential to transform the world? The old parable of the blind men and the elephant suggests an explanation. In this story, a group of men come across

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

a creature they have never seen before: an elephant. Each man grabs hold of a different part of the elephant and describes it based on their own limited perception and experience. None of the men has the knowledge needed to understand the parts holistically to determine that what they are encountering is an elephant. Our attempts to make sense of blockchains are analogous when we try to provide an explanation of them without taking a holistic view.

It is for this reason that I argue we need to approach understanding blockchains not from a singular disciplinary perspective but holistically. In this volume, I will draw upon Lemieux and Feng's (2021) multidisciplinary "three-layer" model, which conceives of blockchains and distributed ledgers as *socio-informational-technical systems*. The model was "born of the need to develop an appropriate framework for the problem-centered design of blockchains, in which the problems are themselves 'wicked,' multidimensional, and multidisciplinary" (Palmer et al., 2021, pp. 591–592). It is well known that "systems designed from a single point of view have often proved to have 'blind' spots which can render them ineffective, or even dangerous. With this in mind, we aimed to design a framework which encourages holistic problem analysis and affords a common language, underpinned by a reasonably shared ontology and epistemic worldview" (Palmer et al., 2021, p. 592).

The original model was simplistic, recognizing that blockchains had social, informational (or more accurately, as I will discuss in Chapter 6, evidential), and technical dimensions. In 2019, a diverse group of blockchain scholars came together to discuss the original three-layer model, especially the interactions among the three layers. With further theoretical refinements arising from these discussions, the most recent version of the model represents blockchains as complex, dynamic systems with four interrelated sub-systems – the original three layers (the social, the informational, and the technical) and a governance sub-system – which work together to achieve trust among social actors (Lemieux and Feng, 2021).

The technical sub-system is reasonably well understood, even as there remain novel technical challenges to be overcome, being those technical components that implement blockchain and distributed ledger systems. The social sub-system – which encompasses social, political, and economic implications of these tools and platforms – though arguably less well understood, has at least been recognized as an important aspect of blockchain systems. Indeed, common use of the term blockchain "ecosystem," rather than "system," draws attention to the fact that blockchains comprise communities that are often "contentious and non-homogeneous, in which unpredictable agents can disrupt the planned flow of ecosystem participation" and in which, therefore, governance is needed (Palmer et al., 2021, p. 591). The final sub-system, the

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

informational, focuses on the ledger itself. Paradoxically, given that a defining feature of blockchain technology is the production of an "immutable" distributed ledger that features heavily in "archival imaginaries" (Woodall and Ringel, 2020) that posit blockchain and distributed ledger technology as a cure-all for our current epistemic ailments, it is this aspect of the technology that has received the least scholarly and research attention.

Scholars who have addressed the question of the immutability of blockchain and distributed ledgers have noted that "'immutability' of blockchain records is a matter of debate, as high-profile events in the blockchain space have shown that blockchain records are changeable at will by the people who govern the blockchain system, and it currently is unclear which variations of blockchain technology actually create a record that even approaches immutability" (Walch, 2017b, p. 1). This observation highlights an important insight that is only possible from a holistic vantage point on blockchain and distributed ledger technologies – one that takes into consideration the social, informational (or evidential), and technical dimensions of the technology in equal measure. From this vantage point, blockchain immutability is best viewed not as a property of blockchain-based ledgers but as a sustained commitment that a group of individuals holds onto because they believe that the attribute is desirable, even necessary. In the remainder of this chapter, I will explore this idea more deeply.

1.2 The Social Construction of Meaning

Recognizing that it would be difficult to advance scientific discussions about blockchain technology without a stable definition of the term, in 2017, global blockchain experts became involved in an international project to develop a standard blockchain and distributed ledger vocabulary under the auspices of the International Organization for Standardization (ISO) Technical Committee on Blockchain and Distributed Ledger Technologies (TC307). This work, which involved the input of over 300 international experts from 50 countries over the span of almost three years, resulted in what has become the first ISO standard on blockchain and distributed ledger technologies, ISO 22739:2020 *Blockchain and Distributed Ledger Technologies – Vocabulary* (ISO, 2020a; Oclarino, 2020). The working group that developed the vocabulary converged on a set of interlocking definitions that capture a shared understanding of what a blockchain is and, equally importantly, what it is not.

After many months of deliberation, the ISO experts arrived at a definition of blockchain as a "distributed ledger with confirmed blocks organized in an

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

append-only sequential chain using cryptographic links" (ISO, 2020a, s. 3.6), with a distributed ledger being defined as a "ledger that is shared across a set of [distributed ledger technology (DLT)] nodes and synchronized between the DLT nodes using a consensus mechanism" (ISO, 2020a, s. 3.22). Thus, in this volume, when I use the term distributed ledger, it encompasses the concept of a blockchain because blockchains are a type of distributed ledger. The ISO defined a ledger as an "information store that keeps records that are intended to be final, definitive and immutable" (ISO, 2020a, s. 3.43).

The idea that blockchains are a type of distributed ledger was not an uncontroversial position among ISO experts, since some held the view that the unique features of the blockchain's chained block data structure and consensus mechanism made blockchains categorically different from distributed ledgers. Despite the consensus reached by the ISO community about the meaning of blockchain, it remains true, as I have previously observed, that "different epistemic communities have formed their own ideas about what blockchain is, some with very strong political and social views around open source, sharing, and autonomy" (as quoted in Jeffries, 2018). It also remains true that legal definitions of blockchain technology continue to proliferate (see, e.g., Walch, 2017a, 2017b). As a result, it is doubtful that everyone will accept and adopt the ISO definitions. Nevertheless, these definitions can at least provide a stable foundation for discussion of blockchain and distributed ledgers for the purposes of this volume, even if they do not end the debate about the meaning of blockchain and related concepts.

It is significant that the ISO experts did not define blockchains strictly in terms of technical components, such as the networked databases that communicate and interact with one another over a network in order to implement a blockchain. ISO 22739 instead refers to these technical components as instantiating blockchain or distributed ledger technology *systems* (ISO, 2020a, s. 3.33). To attempt to understand blockchain purely in terms of the computational technologies, experts understood, is to miss the mark by focusing on the wrong abstraction layer, to use a concept from computing. In software engineering and computing, abstraction involves thinking about and representing a thing, for example, a system, at different levels of granularity or detail. Abstractions, like models, are representations that help simplify a complex world and focus the mind on important details (Butterfield et al., 2016).

In contrast to focusing on the technical system view in its definition of the term blockchain, ISO TC307 chose to focus on a *higher* level of abstraction. In ISO 22739, by recognizing blockchains as a distributed type of ledger, ISO experts connected blockchain with a long tradition of *recordkeeping*. This, in

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

turn, connects blockchains to the theories, principles, and methods of *archival science*[2], which is the science underpinning recordkeeping. Archival science, as Thomassen (2015, p. 84) explains,

> is an academic and applied discipline that involves the scientific study of process bound information, both as product and as agent of human thoughts, emotions, and activities, in its various contexts. Its field of study encompasses personal documents, records, and archives of communities, government agencies, and other formal organizations, and archival materials in general, whether kept by archival institutions, units, or programs. It covers both the records themselves and their contexts of creation, management, and use, and their sociocultural context. Its central questions are why, how, and under what circumstances human beings create, keep, change, preserve, or destroy records, and what meanings they may individually or jointly attribute to records and to their recordkeeping and archival operations.

Thomassen (2015, p. 85) goes on to explain that archival science focuses on more than just records or archival documents to think about records or archival documents *in context*, that is, "the context of the data within a record and the contexts of creation, management, and use, as well as the socio-political, cultural, and economic contexts underlying these contexts." Although it has existed for centuries as a practical field, archival science as an academic discipline is considered relatively new, even if it has disciplinary forerunners that extend back centuries (Duranti, 1989; Thomassen, 2015). The more practical orientation of most archivists and the relative newness of contemporary archival science might account in large part for the comparative absence of archivists and archival science from discourse on blockchains.

Why should it be so important to recognize blockchains as recordkeeping systems and connect them to archival concepts? For one thing, defining blockchains in this way makes it possible to treat them as a single category. No matter how many different types of blockchains and distributed ledgers there are now in the world, or there might be in the future, they all will have one thing in common – a ledger.

Another reason is that recordkeeping and archival theories, principles, methods, practices, and professionals have been long associated with the preservation of "information created or received and maintained as evidence and as an asset by an organization in pursuit of legal obligations or in the course

[2] The "archive" and archives and recordkeeping research has received a great deal of attention within the academy in the past two decades. This research encompasses a diverse range of disciplinary perspectives on the "archive" and the study of archives and archivists. Such studies can be distinguished from archival science, which has its own discipline and its own unique body of theory and practices. At the same time, the cognate field of archival studies encompasses a "multiverse" of perspectives, including those from archival science and archival studies (on this point, see Duranti and Michetti, 2016; Gilliland et al., 2016).

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

of conducting business," that is, with records (ISO, 2020b, s.3.2.10). Evidence is here not limited to the legal sense of the term but rather is "information that could be used either by itself or in conjunction with other information, to establish *proof about an event or action* [emphasis added]" (ISO, 2020b, s. 3.2.6). In order to offer proof of an event or action, evidence must be shown to be inviolate and complete (ISO, 2020b, s. 3.26). Thus records, in order to offer evidence, must, among other things, possess the characteristics of authenticity (actually be what they purport to be),[3] reliability (complete, accurate, and able to stand for the events or actions they represent),[4] and integrity (complete and unaltered) (ISO, 2016, s. 5.2.2). It follows, then, that if we want to design blockchain and distributed ledger systems capable of creating, capturing, and preserving sources of evidence, then recordkeeping and archival theories, principles, methods, practices, and professionals offer knowledge and experience that can provide valuable guidance.

It is the promise – if not yet the reality – of being capable of producing inviolate and complete evidence – or, as expressed in the definition of a ledger in the international standard on blockchain and distributed ledger vocabulary, of being designed to produce final, definitive, and immutable records (ISO, 2020a, s. 3.43) – that sets blockchains (and other distributed ledgers) apart from other types of information systems, such as the commonly used transaction processing systems, management information systems, or office automation systems.

Indeed, in a datafied world, the capability of producing and preserving immutable evidence, as blockchains are designed to do, is a rare one. As paper records and recordkeeping have gradually fallen away to be replaced by digital records and recordkeeping, greater value has been placed on ensuring that the information created by an organization in the conduct of its business can be reassembled into new information assets that might be mined to advance organizational strategy, more often than not profit-driven, or sold to other organizations for similar purposes. As the now well-worn expression goes, "data is the new oil."[5] New business models have arisen based upon exploiting

[3] ISO 30300: 2020, s. 3.2.2, which reads in full "quality of a record (3.2.10) that can be proven to be what it purports to be, to have been created or sent by the agent (3.1.3) purported to have created or sent it, and to have been created or sent when purported" (ISO, 2020a).

[4] ISO 15489:2016, s. 5.2.2 describes reliable records as ones "whose contents can be trusted as a full and accurate representation of the transactions, activities or facts to which they attest" and "which can be depended upon in the course of subsequent transactions or activities." The standard goes on to note that reliable records are usually created "at the time of the event to which they relate, or by systems routinely used to conduct the transactions" (ISO, 2016). In other texts, this notion is similarly captured in the phrase "made in the usual and ordinary course of business."

[5] Clive Humby is attributed with coining the phrase "data is the new oil," but the phrase came into popular usage following a 2017 article in the *Economist* (Economist, 2017).

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

information as assets. To enable these new business models, what once would have been created as records in fixed form is now created and kept in a malleable and manipulable form. Datafication and the creation and storage of vast troves of information have given rise to the so-called era of Big Data and an entirely new field of endeavor – data science, the art of data manipulation and exploitation. While the ability to manipulate records by transforming them into novel forms of data has led to great innovation and scientific advances, it has also undermined the basis of societal proof about past events and actions and, in so doing, contributed to the emergence of an age of disinformation (a topic that will be discussed more fully in Chapter 4). Blockchain, a unique type of ledger, promises to restore society's evidence base. To understand how and, more importantly, why, it is helpful to reflect upon the genesis of blockchain technology.

1.3 Genesis of Blockchain

The blockchain origin story, like all good origin stories, remains somewhat shrouded in mystery. In October 2008, Satoshi Nakamoto – a pseudonym for a person or persons unknown to the present day[6] – proposed a combined digital asset, bitcoin (I will use "bitcoin" with a lower case "b" whenever I am referring to bitcoin the cryptocurrency and with an upper case "B" whenever I am referring to Bitcoin the network), and peer-to-peer payment system (the Bitcoin blockchain network) in a modest nine-page paper entitled "Bitcoin: A Peer-to-Peer Electronic Cash System" (Nakamoto, 2008a). Against the backdrop of a global financial crisis, the genesis block of the Bitcoin network was mined on January 3, 2009 and the first block thereafter was created on January 8, 2009.[7] Nakamoto (2009b) announced the release of the Bitcoin protocol software as open source the day after the first block was mined.

[6] Many theories exist about the real identity of Satoshi Nakamoto (see, e.g., O'Neal, 2019). Some argue that Nakamoto is the American computer scientist, legal scholar, and inventor of the concept of smart contracts Nick Szabo; others that Nakamoto was the late Hal Finney, a cypherpunk and one of the early contributors to Bitcoin's codebase; and still others posit that Nakamoto is British cryptographer Adam Back, CEO of Blockstream. Yet another possibility is Craig Wright – who has actually claimed to be Satoshi Nakamoto – an Anglo-Australian computer scientist and businessman. Rather interestingly, Wright was granted the United States copyright registrations for the original Bitcoin whitepaper and code, which he still holds (Bitcoin SV, 2019).

[7] The original block hash at Block 0 is 000000000019d6689c085ae165831e934ff763ae46a2a6c172b3f1b60a8ce26f and the hash of Block 1 is 00000000b873e79784647a6c82962c70d228557d24a747ea4d1b8bbe878e1206. As Bitcoin is a shared and transparent ledger, readers can see this for themselves at www.blockchain.com/btc/block/000.

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

In this paper, the pseudonymous Nakamoto described the problem that Bitcoin was designed to solve as being one of *trust*:

> Commerce on the Internet has come to rely almost exclusively on financial institutions serving as trusted third parties to process electronic payments. While the system works well enough for most transactions, it still suffers from the inherent weaknesses of the trust based model. Completely non-reversible transactions are not really possible, since financial institutions cannot avoid mediating disputes. The cost of mediation increases transaction costs, limiting the minimum practical transaction size and cutting off the possibility for small casual transactions, and there is a broader cost in the loss of ability to make non-reversible payments for non-reversible services. With the possibility of reversal, the need for trust spreads. Merchants must be wary of their customers, hassling them for more information than they would otherwise need. A certain percentage of fraud is accepted as unavoidable. These costs and payment uncertainties can be avoided in person by using physical currency, but no mechanism exists to make payments over a communications channel without a trusted party.
>
> *(Nakamoto, 2008a, p. 1)*

Bitcoin solved the problem of trust by introducing a mechanism for making financial transactions computationally impractical to reverse, that is, by solving "the double spending problem," which is discussed further below. It achieved this by "using a peer-to-peer distributed timestamp server to generate computational proof of the chronological order of transactions" – the blockchain – which would be secure "as long as honest nodes collectively control more CPU power than any cooperating group of attacker nodes" (Nakamoto, 2008a, p.1). As such, the blockchain to which Bitcoin gave birth relied on the capabilities of two fundamental technological primitives: cryptography and distributed systems.

Cryptography is a centuries old "discipline that embodies the principles, means, and methods for the transformation of data in order to hide their semantic content, prevent their unauthorized use, or prevent their undetected modification" (ISO, 2020a, s. 3.17).[8] Public-key cryptography, upon which Bitcoin relies, is a type of cryptography in which there is a public key – that, just as the name implies, is made public – and a corresponding private key that must always be kept secret. For encryption, the public key is used for encryption and the private key for decryption; for digital signatures, the keys' roles are reversed (ISO, 2020a, s. 3.62 and 3.65). A digital signature, which relies upon public-key cryptography, is data that when appended to a digital object, such as a document, allows someone to verify its authenticity and integrity. (ISO,

[8] On the history of cryptography, see Dooley (2018).

2020a, s. 3.21). In the Nakamoto paper, bitcoin – the cryptocurrency – was described as a chain of digital signatures (Nakamoto, 2008a, p. 2).

All this might have been quite ordinary, save for the fact that most digital signature schemes rely upon a third-party trust anchor – a certificate authority (CA) – which issues digital certificates to certify who owns a public key. This allows trusting parties to rely upon signatures or on assertions made about the private key that corresponds to the certified public key. In Bitcoin, on the other hand, public and private key pairs are created without reliance upon an external CA (Narayanan et al., 2016).[9] As a result, reliance upon traditional CAs to generate key pairs is not necessary with blockchain technology, which may even be a defense against malevolent CAs.

The use of digital signatures is an important aspect of how the Bitcoin network works. To prevent forgery of digitally signed content, an input message is converted into a hash, which is comprised of a 256-bit string, produced when a hash function algorithm is applied to an input message. The input message can be of any size. For example, in theory you could take the entirety of the Bodleian Library or the Library of Congress, if they were completely digitized, and reduce them to a 256-bit hash. For each bitcoin transaction, the hash consists of a hash value of the concatenation of the following fields: version number, transaction inputs, transaction outputs, and lock time. To improve Bitcoin's robustness, an update of the Bitcoin protocol called SegWit was introduced on August 24, 2017, in block number 481824. It includes another two fields called flag and witnesses (see Chapter 8 for more on the SegWit update). Once the input message is hashed, it is nearly impossible to regenerate it from the hash output. In other words, you cannot reverse the hash function to reproduce the Bodleian Library, the Library of Congress, or a bitcoin transaction from its hash, or at least it is currently considered computationally impossible to do so. Hashing, therefore, is a one-way function, not a two-way function as in the case of encryption (and its counterpart, decryption).

The fact that it is not easy to guess or reproduce the content of the original input message from its hash is helpful for maintaining integrity because the hash helps detect any changes to the input message. A change in the input message will produce a different hash output, signalling that the original message could have been tampered with.

As Nakamoto notes in his paper, however, digitally signing transactions alone would not prevent someone from double spending a coin or transferring

[9] Bitcoin uses the Elliptic Curve Digital Signature Algorithm to do this (see Narayanan et al., 2016, pp. 17–19).

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

it a second time. To solve this problem without relying on a trusted third party, additional features need to be designed into the system. Bitcoin achieved this through establishing a *transparent ledger* visible to all participants on the network: “transactions must be publicly announced” with “a system for participants to agree on a single history of the order in which they were received” (Nakamoto, 2008a, p. 2).

The process of establishing a shared ledger begins with each participant on the Bitcoin distributed network digitally signing transactions that propagate to other participants across the network in a “gossip-like way” (Narayanan et al., 2016; Greenspan, 2017). The Bitcoin network, a network of indeterminate size (Narayanan et al., 2016, p. 69),[10] runs on the Transmission Control Protocol (TCP) network protocol and has a random topology, meaning that each participating node communicates randomly with other participating nodes. There is no hierarchy among the participating nodes; each one can come and go as it pleases from the network without any permission (Narayanan et al., 2016, p. 66). This is why Bitcoin is called a “permissionless” distributed ledger. Each participant verifies that every new transaction it receives complies with Bitcoin’s rules, and checks and validates the transaction, including checking for any conflicts with previous transactions (Narayanan et al., 2016).[11] Once a transaction is verified, it enters the participant’s list of provisional unconfirmed transactions (the “memory pool”) and is forwarded on to other participants on the network. Transactions that fail verification enter the participant’s “orphan pool.”.

At this point, provisional unconfirmed transactions still need to be confirmed. Confirmation begins when the “Merkle root” of a tree of hashes of provisional sets of unconfirmed transactions are grouped together into blocks (see Figure 1.1). Each block is limited to a megabyte, or about 1 million bytes, in size. If each transaction is about 250 bytes, then each block can hold a maximum of about 4,000 transactions (Narayanan et al., 2016, p. 72). However, as transaction sizes can vary significantly, this can only be considered a very rough estimate. Each block also has a header that includes, among other data, a hash of the block transaction data (i.e., the Merkle root, or a hash of all previous transaction hashes of the transactions to be incorporated into a specific block, as shown in Figure 1.1), a timestamp, and a hash link to

[10] Narayanan et al. (2016, p. 69) note that “it is difficult to measure how big the [Bitcoin] network is, since it is dynamic and has no central authority.”

[11] Narayanan et al. (2016, p. 68) explain that “Nodes run the script for each previous output being redeemed and ensure that the scripts return true. Second, they check that the outputs being redeemed haven’t already been spent. Third, they won’t relay an already-seen transaction . . . Fourth, by default, nodes only accept and relay standard scripts based on a small whitelist of scripts.”

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

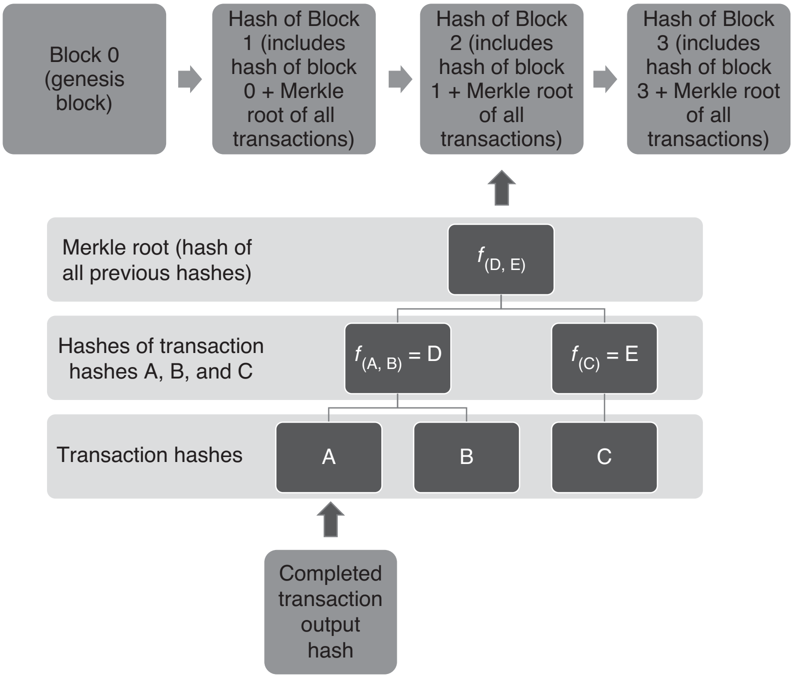
Block 0 (genesis block)
Hash of Block 1 (includes hash of block 0 + Merkle root of all transactions)
Hash of Block 2 (includes hash of block 1 + Merkle root of all transactions)
Hash of Block 3 (includes hash of block 3 + Merkle root of all transactions)
Merkle root (hash of all previous hashes)
$f_{(D, E)}$
Hashes of transaction hashes A, B, and C
$f_{(A, B)} = D$
$f_{(C)} = E$
Transaction hashes
A
B
C
Completed transaction output hash

Figure 1.1 Bitcoin as a chain of hashes

the block that precedes it in order to create an append-only, sequential chain that forms a blockchain (Nakamoto, 2008a, p. 2). This assures that the correct order of transactions is recorded in the ledger and cannot be altered without detection.

To create a transparent and shared ledger on which all participants can agree without resorting to reliance upon a third-party arbiter of the truth requires another novel feature of the Bitcoin network: a mechanism for achieving consensus among network participants on the validity of blocks. To achieve consensus, Bitcoin network participants can choose to work on solving a computational puzzle called Proof-of-Work (PoW) that generates a block hash output that is below a certain difficulty target, typically represented as a hash beginning with the correct number of zeros (19 at the time of writing) when added to a nonce (a random or pseudo-random number) (Nakamoto, 2008a, p. 3; Narayanan et al., 2016, pp. 64–66). Puzzle solvers, called miners, must use central processing unit (CPU) power to find the unknown value. It is computationally very hard to solve the puzzle, so miners are rewarded for their effort with a certain number of new coins (the block reward) in order to

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

incentivize them to expend costly CPU power to maintain the network. Block rewards also help to keep miners honest, on the assumption that it is more profitable for them to expend CPU energy on honestly maintaining the network than subverting the source of their income generation and the basis of their wealth (Nakamoto, 2008a, p. 4).

Mined blocks are then propagated out to other participants across the network. The client software that each participant runs is programmed to validate the blocks it receives first. Once validation is complete, participants update their copies of the ledger, at which time the transactions within the new block are considered confirmed by that participant (Narayanan et al., 2016, p. 68).[12] Once confirmed, any transactions in the participant's memory pool or orphan pool that conflict with those in the new block are then discarded. The confirmation process continues until each participant in the network has completed it. As ideally configured, each network participant is under separate control and independently confirms the validity of blocks. This approach to creating a transparent and shared ledger on which all participants can agree has come to be known as "Nakamoto consensus" (ISO, 2021).

Nakamoto consensus solved a well-known and long-standing problem in distributed computing: the "Byzantine Generals problem." This is a problem that was first theorized by the mathematicians Leslie Lamport, Robert Shostak, and Marshall Pease. In their paper (Lamport et al., 1982), they explain the problem of achieving consensus among peers in an open distributed network when it is unknown which peers might be trusted. To illustrate the problem, they used the metaphor of Byzantine generals on the verge of attacking an enemy city during a siege. The generals are located in different areas surrounding the city and can only communicate via messengers in order to coordinate their attack. However, it is highly probable, or even certain, that there are traitors among the messengers or that the messages of honest messengers have been corrupted. The problem, therefore, lies in the ability to effectively coordinate an attack when it is unknown if messages or messengers have been interfered with. Nakamoto's consensus mechanism solves the problem for Bitcoin by requiring participants to rely upon and add only to the longest chain. Chains containing confirmed updates always grow the fastest and are the longest, since participants producing these updates have solved the PoW puzzle first (Finney, 2008). This helps protect the integrity of the network since,

[12] Narayanan et al. (2016, p. 68) explain, "Validating a block is more complex than validating transactions. In addition to validating the header and making sure that the hash value is in the acceptable range, nodes must validate every transaction included in the block. Finally a node will forward a block only if it builds on the longest branch, based on its perspective of what the blockchain . . . looks like."

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

as explained by Nakamoto, “If a majority of CPU power is controlled by honest nodes, the honest chain will grow the fastest and outpace any competing chains. To modify a past block, an attacker would have to *redo the proof-of-work of the block and all blocks after it and then catch up with and surpass the work of the honest nodes* [emphasis added]” (Nakamoto, 2008a, p. 3). In this way, the Bitcoin network remains trustworthy even if some network participants send incorrect or harmful information.

1.4 Bitcoin Antecedents

Bitcoin’s assemblage of cryptography, distributed networking, transparent ledger, and Nakamoto consensus to incentivize honesty in network participants (i.e., to achieve “trustless trust” [Werbach, 2019], or trust without a central intermediary) was, and remains, novel. It did, however, create a challenge relating to privacy and tracking. To explain, in traditional cash-based systems, ledgers keep track of transactions in third-party recordkeeping systems. For example, when we purchase a cup of coffee with cash, no one asks for our signature, digital or otherwise, though the vendor or their representative (e.g., a cashier) may record the transaction using a cash register and generate a receipt for it. After our purchase, the transaction cannot easily be traced back to us. When we pay with a bank card, on the other hand, we are often required to digitally sign for the transaction by typing in a PIN. In this case, the transactions are easily tracked, are recorded by our bank or third-party payment processor, and can be traced back to us. As a result, we give away a great amount of personal information about our spending habits. Bitcoin, which was designed to cut out these third-party middlemen, aims at affording the digital world the same privacy that comes with using traditional, non-digital cash.

To do this, bitcoin keeps public keys pseudonymous;[13] that is, transacting parties are only identified by their bitcoin address, not by name, and may change their address for each successive transaction (indeed, this is considered good practice), so that it is difficult to link the address to a specific individual. It is in this sense that bitcoin manages to combine the seemingly oppositional

[13] The Nakamoto (2008a, p. 6) paper actually uses the word anonymous; however, even that paper acknowledges that the method is not completely anonymizing, observing, “Some linking is still unavoidable with multi-input transactions, which necessarily reveal that their inputs were owned by the same owner. The risk is that if the owner of a key is revealed, linking could reveal other transactions that belonged to the same owner.” An array of techniques has emerged, aimed at obfuscating the origins of cryptocurrency transactions (cryptocurrency tumblers) and has even given rise to entirely new privacy-preserving blockchains (e.g., Monero). For a further discussion of this, see chapter 6, “Bitcoin and Anonymity,” in Narayanan et al. (2016).

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

properties of transparency and privacy. Note, however, that privacy in the context of bitcoin relates to the identity of individuals originating bitcoin transactions and not to the records of those transactions captured on the blockchain (i.e., ledger records). These remain open for all to see, including any data embedded into those transactions unless the embedded data are purposefully encrypted to preserve confidentiality.

The pseudonymity of bitcoin transactions has the advantage of protecting individuals' financial privacy. At the same time, it can shield from detection those who are using bitcoin for nefarious purposes. This has made the bitcoin cryptocurrency popular with hackers, money launderers, and tax evaders and given bitcoin some of its bad reputation. Proponents of bitcoin argue that traditional "hard" cash can be used for the same purposes, yet there's no denying that the digital nature of bitcoin makes it an attractive form of currency for hackers. New methods of tracking and tracing bitcoin transactions, and of identifying their originators, have emerged. These methods have also been countered by new privacy-preserving techniques adopted for use by alternative cryptocurrencies (e.g., Monero)[14] or cryptocurrency services (e.g., "mixers" or "tumblers" such as the Wasabi wallet).[15]

Although it may seem as though bitcoin sprang "full grown from the head of Zeus," this was not the case. As Clark (2016, p. ix) writes, there was actually a "long road to Bitcoin." One of the first steps along the road, according to Clark, was a company called CyberCash that implemented a protocol called SET, designed to avoid the need for e-commerce customers to send their payment details to merchants or enrol with an intermediary payment processor. Unfortunately, CyberCash did not survive, largely due to user experience problems (Clark, 2016, p. xiii). Another antecedent to bitcoin came from David Chaum's DigiCash company, which was based upon his proposal for e-cash involving cryptographic "blind signatures" aimed at preventing people from spending units of digital currency twice (the "double spending" problem) (Clark, 2016, pp. xiv–xv). Clark argues that the patents Chaum took out on his invention spurred others to invent their own open-source version of e-cash, for example, MagicMoney and Lucre (Clark, 2016, p. xvii).

"Netcash" was the first to propose the idea of minting digital cash through solving a cryptographic puzzle, leveraging cryptographic hash functions similar to the idea of Bitcoin's mining. The idea was first proposed in 1992 by cryptographers Cynthia Dwark and Moni Naor as a potential way to reduce email spam, and a similar idea was discovered independently, according to Clark, by Adam Back in 1997, which Back called Hashcash (Nakamoto

[14] See www.getmonero.org. [15] See https://wasabiwallet.io.

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

[2008a, p. 3] references Back in his Bitcoin paper). In 1991, Stuart Haber and Scott Stornetta proposed, for the first time, the idea for secure timestamping of documents in which clients send documents to a timestamping service and the service signs the document together with a timestamp and a hashpointer to the previous document (Clark, 2016, pp. xx–xxi). B-money and Bitgold (which was proposed by Nick Szabo in 1998, though he did not publicize it until 2005) are both cited as earlier solutions that combined the use of computational puzzles with timestamping to secure transaction records in a ledger (Clark, 2016, pp. xxii–xxiii).

1.5 Blockchain and Immutability

From the original Nakamoto paper and its antecedents emerge the basic characterization of blockchains in intricate and delicate balance: a decentralized ledger, the entries of which are tamper-evident, agreed among all participants, transparent, and of pseudonymous origin. Working together, these characteristics generate arguably the most significant and controversial property of blockchains, *immutability*. Immutability, in the context of blockchain and distributed ledger technology, can be understood as the "property wherein ledger records cannot be modified or removed once added to a distributed ledger" (ISO, 2020a, s. 3.24).

Immutability, however, is better characterized as an emergent property of blockchains and other distributed ledgers. C. D. Broad (1925, as cited in O'Connor, 2020) explains emergentism as,

> the characteristic properties of the whole R(A, B, C) (where R marks their structural arrangement) [that] cannot, even in theory, be deduced from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form R(A, B, C).

That is to say, the structural property of immutability in blockchains cannot be predicted from knowledge of any one of the individual properties of blockchains (e.g., tamper evidentiality, decentralization, transparency, or pseudonymity). Each of these properties must be present, and work in harmony with the others, to produce ledger immutability.

To explain, it would not be possible to achieve tamper evidentiality or resistance without the existence of a transparent ledger, which renders visible any alterations in original input data that might occur. Similarly, if the ledger were to be maintained by a single controlling interest, there would be nothing to prevent that controlling interest from altering ledger records if it were to choose

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

to do so, even if digital signing of transaction records and blocks makes alterations evident. Once control of the ledger is decentralized, however, no single controlling interest can unilaterally alter it because a majority of participants must first agree on the alteration. In the Bitcoin network, this only becomes possible when participants who collectively control more than 50 percent of the CPU effort it takes to confirm blocks agree to change the ledger.

The more decentralized the participants in the network – in the sense of being free from the power and control of one another – the more censorship resistant is the ledger, that is, the more likely it is that an incomplete or altered copy of the ledger held by one of the participants would be rejected. This characteristic also contributes to immutability because it prevents a single controlling interest, or consortium of interests, from exercising control over the ledger in order to rewrite the ledger history. Altering or deleting records written to the ledger requires the agreement of the majority of all the participants on the network to that action, which is very difficult to gain (Daian, 2016), or the destruction of all of the hard disks of all of the participants on the network, which would be nearly impossible if participants are truly decentralized. It is precisely the difficulty of doing this that contributes to a blockchain system's immutability.

Finally, pseudonymity, while protecting privacy, also serves to prevent collusion among network participants since, if participants cannot identify one another, it is more difficult for them to form cartels or fall under one another's control.

Should any of these underlying properties be absent, or altered, ledger immutability may be affected. For instance, so called "51 percent attacks" – when a Bitcoin network participant or group of participants gains more than 50 percent of the hashing power on the network to form a controlling interest – can lead to alterations in the ledger. In this case, those with the controlling interest will have sufficient CPU energy to modify a past block by recomputing the PoW of all blocks to catch up with and surpass the work of the honest nodes (Eyal and Sirer, 2014). In essence, this gives the controlling interest the power to roll back history.

A 51 percent attack is not outside the realm of the possible, given that the formation of mining pools – groups of miners that form consortiums to lower the reward variance for participating miners – create a risk of CPU power concentration on the network. Narayanan et al. (2016, p. 128) noted that, as of 2015, nearly all miners were mining through pools, and in June 2014, one of those mining pools – GHash.IO – actually gained more than 50 percent of the network's CPU power; this led to a backlash against the mining pool. While the community of miners generally agreed to avoid becoming too large, some speculated that the true concentration of mining power may not be visible

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

given that miners can participate in many mining pools simultaneously, which tends to obfuscate their true size (Narayanan, 2016, p. 130). Indeed, a number of mining pools still represent an outsized amount of the network's CPU power (see Figure 1.2).

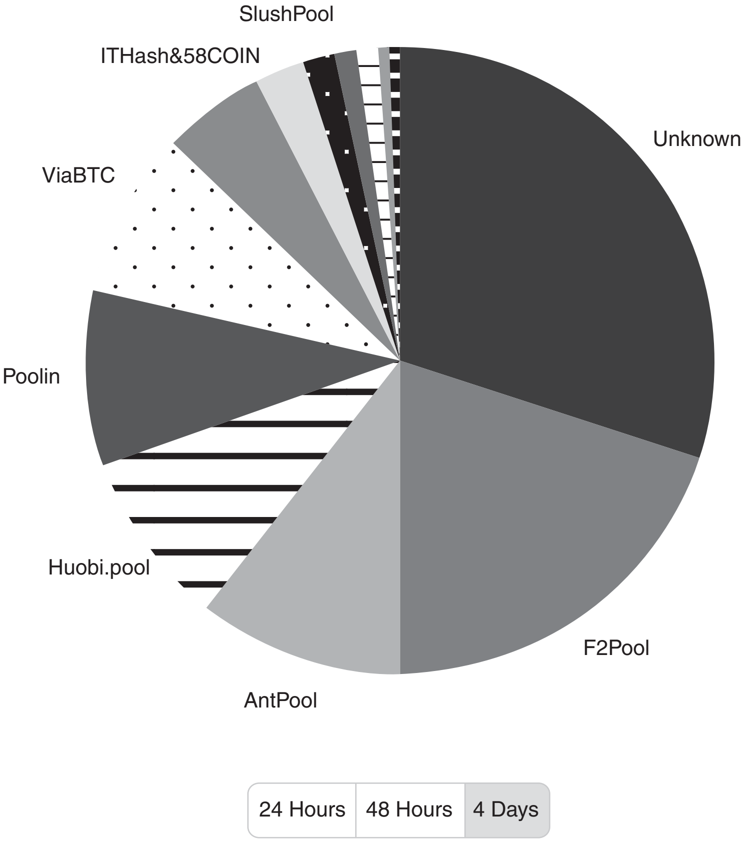

Figure 1.2 An estimation of hashrate distribution among the largest mining pools as of December 27, 2020

(Source: www.blockchain.com/pools, © 2021 Blockchain Luxembourg S.A. All rights reserved)

Note: The source of Figure 1.2 (www.blockchain.com/pools) includes the following caveat: "This graph shows the market share of the most popular bitcoin mining pools. It should only be used as a rough estimate and for various reasons will not be 100 percent accurate. A large portion of blocks are grouped into the "Unknown" category. This does not mean an attack on the network, it simply means it has not been possible to determine the origin."

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

In addition, quantum computing has been identified as an existential threat to the immutability of blockchains, since it allows for the possibility of breaking the cryptography upon which the immutability of the blockchain partially depends (Fernández-Caramès and Fraga-Lamas, 2020). New, quantum-capable computers threaten both the digital signature cryptography and the hash functions used to secure blockchains. This has spurred recent efforts aimed at redesigning blockchains to create what are variously called post-quantum, quantum-proof, quantum-safe, or quantum-resistant cryptosystems (Fernández-Caramès and Fraga-Lamas, 2020; Mashatan and Turetken, 2020; Mashatan and Heinztman, 2021).

In a 2017 paper, legal scholar Angela Walch, commenting on a new law in a state of Arizona (USA) statute[16] that gave recognition to signatures secured through a blockchain as "immutable and auditable and provid[ing] an uncensored truth," drew attention to the problematic nature of the concept of immutability in blockchains (Walch, 2017a). One of the high-profile events to which Walch refers is the DAO exploit of 2016, which involved the Ethereum blockchain.

Programmer Vitalik Buterin originally conceived of Ethereum in 2013. Development was crowdfunded in 2014, and the network went live on July 30, 2015.[17] In its original form, Ethereum bore many similarities to Bitcoin, but a key difference was that it incorporated a virtual machine that could execute Turing-complete scripts and run decentralized programs, called smart contracts. The DAO exploit – DAO standing for decentralized autonomous organization – involved a smart contract, which encoded rules for the operation of the DAO, within which there was written a fatal flaw: it allowed a poorly written function that permitted the repeated withdrawal of DAO funds, rather than the single withdrawal allegedly intended by the contract's author. An attacker was able to exploit this flaw to siphon off Ether (Ethereum's native cryptocurrency) worth an estimated USD 50–70 million (Daian, 2016; Tapscott and Tapscott, 2016; Butijn et al., 2020). All this would have been fine as far as the immutability of blockchain goes, but it was certainly not okay with those whose funds had been misdirected into the attacker's account.

How to handle this obviously unintended situation sparked a fierce debate in the Ethereum community. Some, including Buterin, proposed an update to Ethereum's software that would allow for the recovery of the Ether (Daian, 2016; Butijn et al., 2020). Those who maintained a commitment to immutability – that is, to the idea that "code is law" and that the blockchain should not be

[16] Act of Sept. 21, 2006, ch. 26, ARIZ. REV. STAT. ANN. § 44–7003 (2006) (amended by 2017 Ariz. Sess. Laws 2417). https://legiscan.com/AZ/text/HB2417/id/1528949

[17] See https://ethereum.org/en/history.

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

subject to human interference – stood firmly against this notion. In the end, a majority (about 89 percent) of Ethereum network participants "voted" for the solution supported by Buterin by updating to the new version of Ethereum. Opponents of this approach did not update to the new version, leading to a split in the ledger (a "hard fork," which occurs when blocks generated using a new version of a blockchain protocol are not accepted by those operating an older version [ISO, 2020a, s. 3.38]). Essentially, by majority rule, those who updated to the new version of Ethereum agreed to "interfere" in the operation of the ledger in order to restore DAO funds, while those who did not agree to this interference continued to add to and accept only the original version of Ethereum, which became known as "Ethereum Classic."

As Gideon Greenspan (2017) points out, supporters of Ethereum Classic paid a hefty price for their commitment to immutability: Ethereum Classic is worth a fraction of the price of Ethereum. Although the DAO exploit highlights that the property of blockchain immutability is conditional, it also shows that alteration is not easy. In the DAO example, alteration involved debate among the entire Ethereum community, required a majority of participants on the network to agree to the change, and ultimately caused a division in the Ethereum network. A realistic way to view immutability in the context of blockchains is expressed by Greenspan "there is no such thing as perfect immutability. The real question is: What are the conditions under which a particular blockchain can and cannot be changed? And do those conditions match the problem we're trying to solve?" (Greenspan, 2017).

Threats to immutability may be dismissed as being possible (even increasingly likely) but outside the expected normal operations of blockchain networks. However, even normal blockchain and distributed ledger operations may deliver less-than-immutable ledgers; indeed, it might even be considered desirable for them to do so under certain conditions. Debate surrounding the EU's General Data Protection Regulation (GDPR) and the "right to be forgotten" – which requires the erasure of personally identifiable information from data stores under certain conditions – best illustrates this point, though these regulations are not the only legal requirements for erasure or alteration of information.[18] GDPR establishes data protection as a fundamental right of all EU citizens (Finck, 2018; Hofman et al., 2019). Article 17 of the GDPR makes provision for a "right to be forgotten," or the right to erasure, stating that

[18] The US Fair Credit Reporting Act also requires deletion of personal information, for example. Under this law, consumer reporting agencies must correct or delete inaccurate, incomplete, or unverifiable information, typically within 30 days. The US Federal Trade Commission has estimated that 40 million Americans have inaccuracies in their credit reports under the current system.

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

> [a] data subject shall have the right to obtain the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where [one of the legislative grounds] applies.
>
> *(European Parliament and the Council of European Union, 2016, Article 17, Section 1)*

The right to erasure is not absolute; indeed, it is only required when one of several principles specified in the regulations applies. For instance, a data controller may deny a request for erasure if there is a legal reason to retain the data, or if there is an "archival purpose" for retention of the data, such as to meet public interest, scientific, or historical research purposes. As Hofman et al. (2019) observe, the right to erasure may well apply to blockchains where the ledger contains personal information and the legislated limitations do not apply.

In order to circumvent the apparent incompatibility between the immutability of blockchains and the right to erasure, a number of solutions have been proposed (Hofman et al., 2019), among these being the somewhat infamous and much derided "editable" blockchain. This was an idea proposed by Accenture in 2016. In a *Coindesk* post on the subject, David Treat argued that immutable blockchains would slow adoption of the technology:

> For those who only believe in the permissionless mode of blockchain solutions, there is no need to discuss alternatives. But as industries explore new uses for blockchain beyond cryptocurrency and permissionless systems, there will be situations when that same immutability could make it difficult for the technology to advance.
>
> *(Treat, 2016)*

Treat was referring to the rise of private and permissioned distributed ledgers, or distributed ledgers that are accessible only to a limited number of participants and that require authorization to participate or to perform a specific activity (ISO, 2020a, s. 3.57–3.61). Such distributed ledger variants had arisen in response to government and private organizations' desire to retain control over their technology and platforms, which the large permissionless blockchains and distributed ledgers – such as Bitcoin and Ethereum, in which participants operated independently (even if only theoretically) and without special authorization – do not allow.

The need to comply with regulatory requirements, such as GDPR, was one driver of organizations' preference for permissioned over permissionless blockchains. In the context of permissionless blockchains, Treat noted a number of problems that likely could only be solved by incorporating the capability to edit blockchains. For example, if records can only ever be added to the chain, ledger size and storage costs will continue to grow. Those in favor of blockchain editability also argued that it should be

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

possible to delete illegal content, such as child pornography, or to correct mistakes created by human error or intent, such as the DAO. Finally, the need to comply with regulatory requirements to remove or redact data makes it necessary to be able to edit a blockchain. These issues led Accenture, for whom Treat worked, to conclude that

> On one side of the debate are those who argue that immutability is precisely what makes the blockchain such a significant innovation. On the other side are pragmatists who increasingly see where and how in enterprise environments immutability may prohibit adoption due to human error, mischief and privacy laws.
>
> *(Accenture, 2016, p. 3)*

Accenture's solution was to introduce a new variation of the "chameleon" hash function to enable blockchain editing, allowing the creation of a "virtual padlock" linking two blocks. Deleting or altering a block would involve unlocking the padlock with a private key (called a "trapdoor"), which unchains the blocks and allows insertion of a new or altered block in place of the old one without breaking the integrity of the entire blockchain (Accenture, 2016, p. 7).

Accenture (2016, p. 7) acknowledged that the solution was designed for permissioned systems, wherein there is a designated administrator responsible for managing the system, and rules, procedures, and roles are defined in advance, unlike in Bitcoin. Accenture further noted that redaction should only be available in exceptional circumstances, such as the correction of typos and factual errors, or to bring the data into compliance with the requirements of changed legislation.

Opponents of the approach immediately called out the potential for financial fraud (Kelly, 2016). As Kelly (2016) wrote in a critical *Coindesk* commentary, "The moment you allow someone to change the record you begin to erode trust. While the change may be for the most benign reasons, like human error, it invariably opens the door to the erosion of trust." Commenting on financial market manipulation, Kelly went on to observe that it was "the ability to change the permanent record that enabled Bernie Madoff to commit the largest fraud in financial history" and that "[a] blockchain is a great way to keep a record that you don't ever want changed – this is the heart and soul of a trustless system – it is a feature, not a flaw" (Kelly, 2016).

1.6 Concluding Thoughts

Given the conditionality of blockchain immutability, whether intentional or not, it is inadvisable to view it in essentialist terms as a fixed and stable

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press

blockchain property. Rather, blockchain immutability is as much socially constructed as technically implemented and is best viewed as a sustained commitment that a group of individuals holds onto because they believe that the attribute is desirable and necessary. This commitment has been affirmed by experts from countries around the globe participating in the development of international standards on blockchain and distributed ledger technology, even though they acknowledge that immutability is a design goal of blockchain and distributed ledgers, not something that can be absolutely guaranteed (ISO, 2020a). Why should so many individuals have come to see the property of immutability as desirable or necessary, despite its limitations? Nakamoto's original paper provides the answer: *trust*. It is to the question of trust that we now turn in the next chapter.

https://doi.org/10.1017/9781108877350.002 Published online by Cambridge University Press