Queen Mary University of London, School of Law

Legal Studies Research Paper No. 268/2017

Blockchain Demystified

Jean Bacon, Johan David Michels, Christopher Millard, and Jatinder Singh

Blockchain Demystified[+]

An introduction to blockchain technology and its legal implications

Jean Bacon[*], Johan David Michels[**], Christopher Millard[***], and Jatinder Singh[****]

This paper provides an introduction to blockchain technology and its legal implications. It explains how the underlying technology works and can be deployed in various ways to create applications with different features, including distributed and centralised platforms. It analyses the technology's implications for law enforcement, private law (including contracts, companies, and securities), and EU data protection law. The purpose of this paper is to help legal and other professional advisors understand blockchain technology, so they can tailor appropriate advice, and to alert users of blockchain technology to the current legal uncertainty and associated risks.

Keywords: Blockchain, Distributed Ledger Technology, DLT, cryptocurrency, Bitcoin, Ethereum, smart contract, DAO, ICO, criminal law, contract law, company law, securities, intellectual property, data protection, competition law

[+] This paper has been produced by members of the Microsoft Cloud Computing Research Centre, a collaboration between the Cloud Legal Project, Centre for Commercial Law Studies, Queen Mary University of London and the Department of Computer Science and Technology, University of Cambridge. The authors are grateful to members of the MCCRC team and to attendees at the 4th Annual MCCRC Symposium (September 2017) for helpful comments and to Microsoft for the generous financial support that has made this project possible. Responsibility for views expressed, however, remains with the authors.

[*] Professor Emerita of Distributed Systems, Department of Computer Science and Technology, University of Cambridge.

[**] Researcher, Cloud Legal Project and Microsoft Cloud Computing Research Centre, both at the Centre for Commercial Law Studies, Queen Mary University of London.

[***] Professor of Privacy and Information Law and Project Leader, Cloud Legal Project, Centre for Commercial Law Studies, Queen Mary University of London and Senior Counsel, Bristows LLP. Joint Director of the Microsoft Cloud Computing Research Centre.

[****] EPSRC Research Fellow, Department of Computer Science and Technology, University of Cambridge.

Contents

1. Introduction

In 2017, excitement about blockchain and related technologies soared to new heights. The market caps of the two main cryptocurrencies (Bitcoin and Ethereum) increased by c. 1,200% to a combined total of over US$200bn in the period January to November 2017.[1] Initial Coin Offerings (ICOs) raised an estimated US$1.3bn in the months January to September 2017.[2] Organisations ranging from banks to charities publicly expressed their interest in using blockchain technology. This flurry of activity sparked responses by legislators and regulators, including in Russia, China, the United States, and the EU.[3]

[1] Charles Bovaird, "*Why The Crypto Market Has Appreciated More Than 1,200% This Year*" (2017), https://www.forbes.com/sites/cbovaird/2017/11/17/why-the-crypto-market-has-appreciated-more-than-1200-this-year/#9784906eed3b [accessed 29 November 2017].

[2] Martin Arnold, "*Tech start-ups raise $1.3bn this year from initial coin offerings*" (2017), https://www.ft.com/content/1a164d6c-6b12-11e7-bfeb-33fe0c5b7eaa?mhq5j=e5 [accessed 31 October 2017].

[3] See for example: the Russian government set up a legal framework for cryptocurrencies through presidential orders in October 2017, see https://www.theverge.com/2017/10/31/16387042/russia-putin-bitcoin-regulation-ethereum-blockchain-technology ; China regulators have shut down cryptocurrency exchanges in September 2017, see https://www.forbes.com/sites/leonhardweese/2017/11/29/bitcoin-regulation-in-china-still-unclear-but-chinese-exchanges-thrive-overseas/#43bd2cfa6487; New York State licensed cryptocurrency exchanges in January and November 2017, see

There are many tutorials, reports, and various books on blockchain. However, a lot of the existing material assumes that readers are familiar with the underpinning technologies. Further, some sources fail to distinguish between the core components of blockchain technology and the various ways in which the technology can be applied. Finally, the terminology used to describe blockchain is often unclear or inconsistent.[4] As a result, many discussions of blockchain are marred by misunderstandings and can leave audiences mystified.

This paper aims to demystify blockchain. It first introduces the two core cryptographic functions underpinning blockchain and explains for a non-expert audience how they work. It then reviews key differences between the platform design choices of early cryptocurrencies and the ways in which future applications may utilize blockchain technology.

Finally, the paper explores blockchain's legal implications. It argues that understanding the different ways in which platforms can apply blockchain technology is often key to accurate legal analysis. As a result, there can be no one-size-fits-all legal response to blockchain technology. Instead, this paper aims to help legislators, regulators, and lawyers understand blockchain technology, so they can tailor appropriate legal solutions to each use case. Moreover, the initial exploration of legal implications could be helpful to those considering the use of a blockchain solution in almost any context.

1.1 Core components of blockchain technology

In our view, a blockchain is a type of database, in that it is a structured collection of information. In this paper, we use the term 'blockchain' to refer to a specific type of database that uses certain cryptographic functions to achieve the requirements of data integrity and identity authentication, as set out in the table below. (Since blockchains commonly track transactions, they are often referred to as 'ledgers'.)

Table 1. Two key requirements of blockchain technology

No.	Requirement	Component	Purpose
a)	Data integrity	Hash functions	To create a persistent, tamper-evident record of relevant transactions.
b)	Identity authentication	Public key infrastructure	To authenticate the party or parties associated with each transaction.

In Section 2, below, we explain first how hash functions can be used to create a persistent, tamper-evident record of transactions. Then, we show how public key infrastructure (PKI) can be used to authenticate the identity of the parties associated with transactions recorded on the blockchain.

We use the term distributed ledger technology, or DLT, to refer to a ledger that is stored in a distributed manner across a peer-to-peer network. By this definition, a Distributed Ledger (DL)

https://www.forbes.com/sites/jonathanponciano/2017/11/28/with-regulatory-approval-bitflyer-launches-u-s-exchange-targeting-institutional-traders/#78efff94d0fa; European Securities and Markets Authority issued a warning on ICOs in November 2017, see https://www.esma.europa.eu/press-news/esma-news/esma-highlights-ico-risks-investors-and-firms [all accessed 20 December 2017].

[4] Cf. Angela Walch, "*The Path of the Blockchain Lexicon (and the Law)*" (2017), 36 Review of Banking & Financial Law, pp. 719-735, https://ssrn.com/abstract=2940335 [accessed 20 December 2017].

is also a blockchain if it uses a blockchain data structure to record transactions. However, a blockchain that is stored in a centralised manner is not a DL.

1.2 Platform design

Cryptocurrencies like Bitcoin and Ethereum are the most well-known applications of blockchain technology. They have shaped the public perception of what a blockchain is. However, the Bitcoin and Ethereum platforms were configured to meet specific requirements for creating a secure online currency that anyone could access (known as 'open' or 'permissionless' systems). They were intended to operate in a so-called '**trustless environment'**, *meaning the participants in the network need not trust each other*. These requirements shaped the ways in which these early platforms applied blockchain technology, as set out below in Table 2.

Table 2. Additional requirements of early cryptocurrencies

No.	Requirement	Component	Purpose
c)	Trustless environment	Distributed storage on a peer-to-peer network (P2P)	Early applications assume that no single authority can be trusted. As a result, the blockchain must not be held by any central party, but instead stored in a distributed manner, with each node holding a copy.
		Consensus Protocols	To ensure consistency across the many copies of a distributed blockchain, early applications use consensus protocols to tell nodes which new 'blocks' to add to their local copy of the blockchain.
d)	Open access	Permissionless system	Early applications are open for all to use, store, and update. There are no identity restrictions for participating in the system as a user or provider of system functions.
e)	Public visibility	Publicly available download	Anyone can download the full archive of transactions and view the transaction data.
f)	User pseudonimity	Addresses as identities	Users of early applications could not be identified as real-world entities. Instead, the systems use public keys and 'addresses' to identify users.[5]
g)	No double spending	Payment verification protocols	In order for cryptocurrency coins to have value, the owner must not be able to spend each coin more than once.
h)	On-chain assets	Tokens or coins	Early applications tracked transactions in what are known as 'on-chain assets', i.e. tokens or coins that exist only by virtue of the ledger and do not represent real-world objects. As a result, the ledger provides a complete record of the relevant assets.

[5] A Bitcoin address is a 26-35 digit combination of letters and numbers which is generated based on a hash of a user's public key (what is called a 'Pay to Public Key Hash' or P2PKH system). To simplify payments, a Bitcoin address can also be represented as a QR code. For further details, see https://blockgeeks.com/guides/blockchain-address-101/ [accessed 20 December 2017].

Blockchains can be applied in a variety of ways to create platforms with different properties and features. As blockchain technology is adopted for purposes other than currencies, the above requirements may not be carried forward. Some applications entail 'closed' or 'permissioned' systems, participation in which is limited to a certain group of users. In such cases, there is likely to be a higher level of trust among users, reducing the need for distributed storage and consensus protocols.

In Sections 3 and 4, we review the way in which early platforms have applied blockchain technology in open, permissionless ways to meet requirements (c) through (g) above. We then turn to closed, permissioned platforms and consider how these might differ in terms of their configuration and features. Section 3 looks at the key issue of control over the blockchain: who can store copies of the blockchain and who can propose new 'blocks'. Section 4 sets out the differences between open and closed platforms with regard to visibility of the blockchain record and user identity. Section 5 briefly outlines the role of smart contracts.

In Section 6, we explore blockchain technology's legal implications. We first consider how a platform's design affects the opportunities for law enforcement. We then consider the legal implications of smart contracts, digital autonomous organisations, and initial coin offerings under contract, company, and securities law. Finally, we consider the implications of EU intellectual property, data protection, and competition law.

We conclude that any particular blockchain-based platform may be more or less decentralised and more or less anonymous, based on application requirements and associated technical design decisions. These features in turn have significant legal implications, for instance with regard to the difficulty of reversing past transactions recorded on the blockchain.

This paper does not explore questions around what assets the blockchain relates to, as noted in requirement (h) above. While early applications only tracked on-chain assets (i.e. digital tokens which exist only by virtue of the blockchain), some future applications will use tokens to reflect real-world assets. A separate paper explores the associated legal implications.[6]

2. Core components of blockchain data structures

2.1 Data integrity

Blockchain technology aims to create a persistent, tamper-evident record of relevant transactions. This section shows how hash functions can be used to create a tamper-evident data structure.

2.1.1 Hash values prove the integrity of data

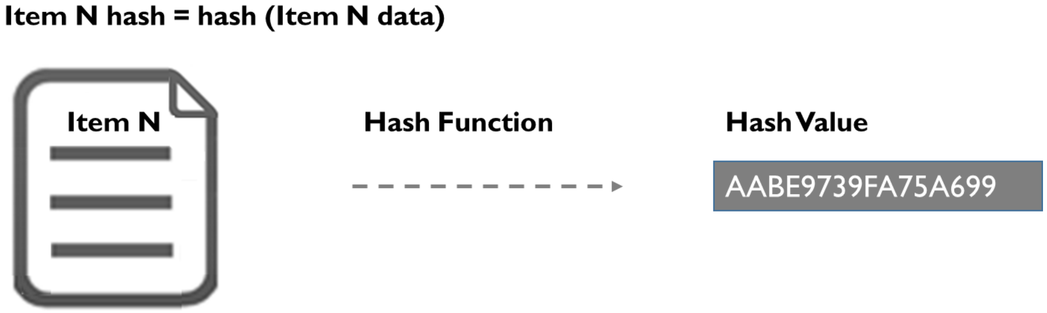
Hashing involves putting the contents of a data item (e.g. a document), through a 'hash function'. This function creates a string of digits of a fixed length which are unique to the input data item. The output is called a 'hash value'. It is practically 'impossible' for two different data items to hash to the same value (i.e. the probability of that occurring is very low[7]).

[6] Chris Reed, Umamahesh Sathyanarayan, Shuhui Ruan, Justine Collins**,** "*Beyond BitCoin – legal Impurities and off-chain assets"* (2017), Queen Mary University of London, School of Law Legal Studies Research Paper No. 260/2017, https://ssrn.com/abstract=3058945 [accessed 20 December 2017].

[7] The probability of a collision in SHA256 is of the order of one in 10^{60}. The transaction rate of e.g. Bitcoin blockchain is currently 10 transactions per second $=3\text{x}10^{8}$ per year, so of the order of 10^{52} years for a collision.

As a result, hashing can be used to prove the **integrity** of the input data. If the original input is changed in any way (by even a single character or space), the function will produce a totally unrelated hash value. To prevent tampering, this requires the hash value - but not the data item itself - to be visible to external observers. Conversely, if the hash value is unchanged, we can be confident that the input data has not been tampered with.

Figure 1: A hash function that outputs a 16-digit (8 byte, 64 bit) hash value



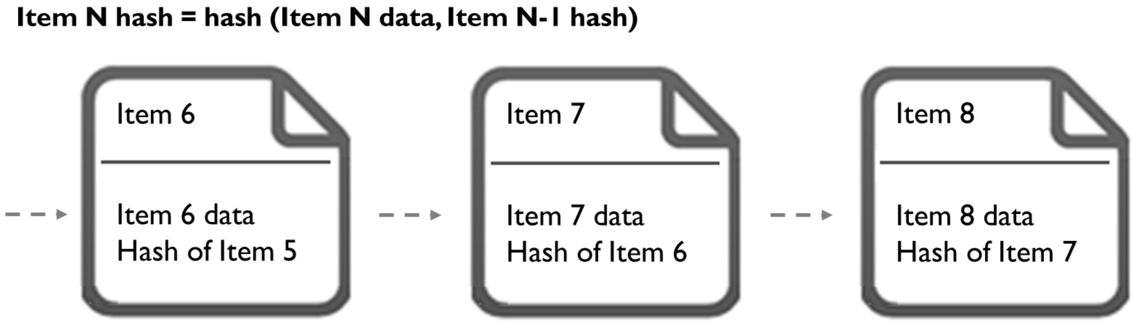
Hashing is 'one-way', in that it is not possible to recreate the original input from the hash value that the hash function outputs. Hashing does not change the input data. Unencrypted data, with the associated hash value, are readable by anyone with access to them. In sum, the combination of a data item and its hash value is tamper-evident, in that it would be evident to any observer if the data item were changed in any way.[8]

2.1.2 Hash pointers create tamper-evident data chains

Hash values can also be used to make a data structure of multiple data items tamper-evident, through 'hash pointers'. Hash pointers prove the integrity of a string of documents, including both their contents and their sequence.

Hash pointers achieve this by linking a series of items together, as illustrated in Figure 2 below for Items 6-8. The data of each item is combined with the hash value of the previous item and put into a hash function. This generates that item's hash value, which is then included in the next item. For example, the hash value of Item 8 is based on both Data 8 and the hash of Item 7. Item 7 contains Data 7 and the hash of Item 6, and so on, back to the start of the chain.

Figure 2. A tamper-evident chain of items using hash pointers



This results in a tamper-evident data chain. The data of Item 7 cannot be changed without changing its hash value. However, any attempt to 're-hash' Item 7 would break the link between the items, since the hash of Item 7 is recorded in Item 8. Provided an external

[8] See Donald E. Knuth, "The Art of Computer Programming" (1998) Volume 3, Addison Wesley Longman, Reading Massachusetts.

observer can view the hash pointers, they can spot any tampering. So if a (fraudulent) change were to be made in the data of Item 7, all subsequent blocks in the chain would have to be re-hashed to rebuild the chain. Further, if there are multiple copies of the blockchain, all copies of the chain would have to be changed in the same way. Having multiple copies of the chain therefore reduces the likelihood of data being changed by an attacker.

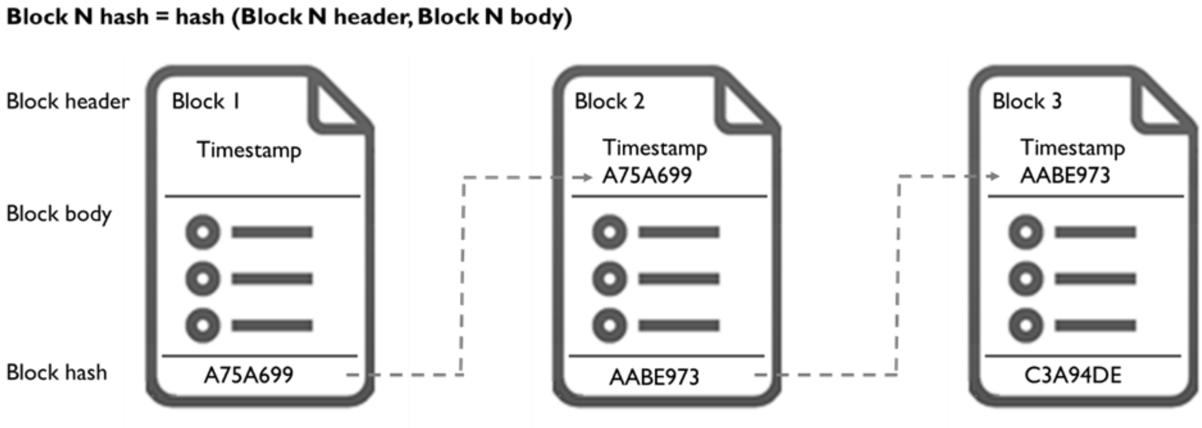
2.1.3 Blockchains group transactions into 'blocks' in a 'chain'

Instead of merely linking single data items, for efficiency reasons blockchains record large numbers of transactions. They achieve this by grouping individual transaction records together, into a **block**, and chaining blocks together using hash pointers.[9]

A block consists of two parts. The 'block body' contains the transactions that the block records. The 'block header' includes the hash of the previous block and some metadata such as a timestamp. Blocks are hashed as a whole, i.e. the header and the body are used as input data for the hash function.

Thus, a block's hash value is created from data that includes the hash of the previous block. Blocks are chained using these block hash pointers, creating a 'blockchain', as shown in Figure 3. (The figure does not represent any particular platform's block header in detail and is intended as an example.)

Figure 3. A simple blockchain representation showing three chained blocks

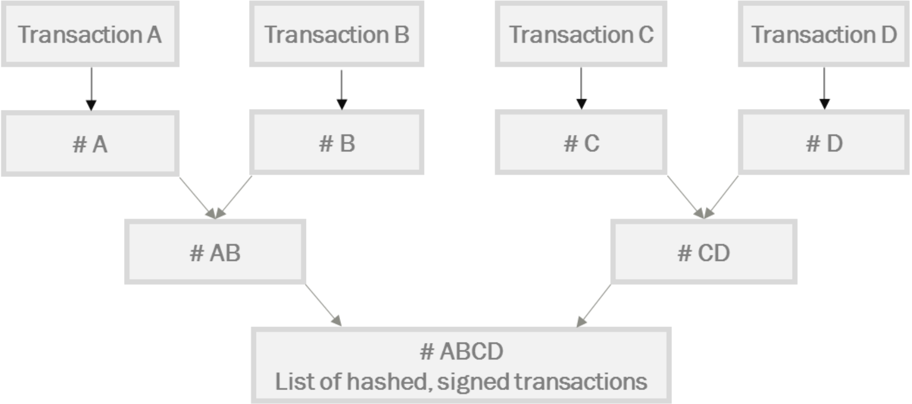


In practice, for scalability and to limit access latency, a more general 'Merkle Tree' structure is used to record the transactions within each block.[10] This makes lookup more efficient, both for proof that a data item exists in a block and for proof that some other item does not exist. Figure 4 shows the idea for a small number of transactions in a block.

[9] Satoshi Nakamoto, "*Bitcoin: A Peer-to-Peer Electronic Cash System*" (2009), https://bitcoin.org/bitcoin.pdf [accessed 19 October 2017], p. 2.

[10] Nakamoto (2009), supra note 9, p. 4. See further Arvind Narayanan, Joseph Bonneau, Edward Felten, Andrew Miller, Steven Goldfeder, "*Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction*", (2016) Princeton University Press, Princeton, pp. 12-15.

Figure 4. A Merkle Tree lists hashed transactions in the body of a block (showing four transactions)



2.2 Identity authentication

To record transactions securely, blockchain technology needs to authenticate the parties to any transaction, before storing them in a tamper-evident database. Otherwise, an attacker could simply pose as another party and propose transactions to be included in blocks. For instance, they could send another user's Bitcoins to an address they control. This section explains how blockchain technology uses public key infrastructure to authenticate users' identities and ward off such attacks.

2.2.1 Public key infrastructure

Public key infrastructure (PKI) comprises a function to generate a **key pair** consisting of a public and a private key, a signing algorithm, and a validation function to check whether a digital signature is correct. The key pair has the following properties: data encrypted with the public key can only be decrypted using the private key and vice versa. If a certain set of data can be decrypted with the public key, this proves the data was encrypted by and therefore came from the holder of the private key.

As the names imply, users must never reveal their private keys, since anyone who knows a private key can masquerade as its owner. Private keys should not be transmitted, even when encrypted. Conversely, the public key is published to represent the individual or entity holding the corresponding private key.

To check whether someone holds a particular private key, users can engage in a challenge-response interaction. The sender takes a message containing a nonce (an arbitrarily chosen number) and encrypts it with the recipient's public key. If the recipient can decrypt this message, add one to the nonce (or apply any agreed function), and then encrypt it with the private key and send the result back, the sender can be confident the recipient held the private key. Thus, PKI can be used to establish a digital identity.

2.2.2 Digital signatures

PKI can be used to create digital signatures, which establish that a transaction emanated from a certain user. To sign a data item, the sender encrypts the data with their private key. If the public key can be used to decrypt the data, this proves that the sender held the private key.

With blockchain, the private key is used to encrypt the transaction record. This effectively establishes that the transaction originated with the associated party. (Either the owner of the corresponding private key must have signed the data, or the key has been compromised by theft or sharing.) Transaction records are signed before being included in blocks.

Each party's private key is their means of access to the blockchain platform. If they lose their private key, the platform can no longer authenticate their identity and will deny them access. Thus, if a user loses their Bitcoin private key, they can no longer access any associated coins.

2.3 Summary of core components

In sum, hash functions can be used to generate hash pointers that link blocks of transactions together in a chain. The hash pointers establish the integrity of the data within each block, as well as the order of the blocks, thereby creating a tamper-evident data structure. For example, Bitcoin generates hash pointers using Secure Hash Algorithm 256 (SHA-256), a well-known hash function that generates a 256-bit (or 32-byte) hash.

A private key can be used to establish an individual's identity through a digital signature. Blockchains combine private keys and hash functions to create a long-term, tamper-evident record of transactions between parties with verified identities. However, the intended long-term storage of blockchain records raises the issue of whether current encryption schemes will continue to be sufficient. Quantum computing may present a challenge to encryption in the long term.

Although hash functions and their uses were first described in the 1950s-70s,[11] and PKI in the 1970s, it wasn't until 2008 that a person (or persons) known as 'Satoshi Nakamoto' combined these two components to describe a blockchain data structure.[12] The next section considers issues of trust and control and explains how Nakamoto used a blockchain database stored across a peer-to-peer network to create Bitcoin, the world's first functional cryptocurrency.

3. Trust and control

For a cryptocurrency to work, users need to value the coins it tracks. This requires users to trust that their coins are secure. As set out above, the use of a blockchain should give users confidence in a tamper-evident ledger of their currency transactions. Further, users should be confident that nobody can spend their coins without access to their private key. In addition, users need to be confident that nobody can spend the same coin twice (known as 'double-spending').

However, using a blockchain does not in itself prevent tampering by whoever controls the ledger. For example, an ill-intentioned record-keeper could potentially change transactions in past blocks and then re-hash all the blocks up to the present block, so that the new hash pointers link the blocks. Thus, in a system where control over the ledger is centralised in the hands of a single party, users need to trust this party not to tamper with the ledger.

This raises a key question for each use of blockchain technology, namely: who will control the blockchain? This can be split into two sub-questions, namely: (i) who stores the current version of the blockchain and (ii) who can add new blocks to it?

Early cryptocurrencies seek to operate in a trustless environment.[13] Instead of centralising control, they allow anyone to store a local copy of the blockchain and propose new blocks for inclusion. To achieve this, they rely on a network of 'nodes' to store copies across a peer-to-peer network and 'miners' to propose new blocks, as explained below.

[11] See Knuth (1998), supra note 8.

[12] Nakamoto (2009), supra note 9, p. 1.

[13] Nakamoto (2009), supra note 9, p. 1.

In contrast, future applications of the technology may operate in environments where there is a degree of trust. As a result, future platforms may feature either a single entity (known as a 'Trusted Third Party' or TTP) or a small group of participants that operate the blockchain.

Below, we first review the design decisions of 'trustless' platforms that operate a distributed ledger. We describe how the decision to operate as fully open, permissionless platforms shaped many of the other features of early applications of blockchain technology, including the need for an energy-intensive consensus protocol. We use Bitcoin as an example to illustrate the workings of a widely distributed platform. We then consider how future closed, permissioned platforms may differ from early cryptocurrencies.

3.1 Trustless environments and widely distributed platforms

3.1.1 Users: permissionless access

Early cryptocurrency platforms involve three (overlapping) groups: users, nodes, and miners. Users participate in the platform by buying and selling coins like Bitcoin or Ether. To participate, they run open source code on their local machine or buy a dedicated hardware wallet.[14] This software broadcasts the users' required transactions onto the network, to be incorporated into blocks by miners. (In Bitcoin, the pool of transactions waiting to be confirmed is called the 'mempool'.[15]) Section 5 covers 'smart contracts' that assist in the definition and broadcast of users' transactions.

Early cryptocurrencies were open or permissionless at the user level. For example, in Bitcoin, anybody can generate a public private key pair and a Bitcoin address via their open source software. Alternatively, they can join a software wallet service online that generates the key pair for them.[16] To start trading, users can buy Bitcoin from online exchanges,[17] or by finding other users to trade Bitcoin in person.[18]

However, using intermediary services such as wallets or exchanges requires a level of trust from participants, since they store copies of the user's private key. While Bitcoin itself has not been hacked, several exchanges have been, resulting in substantial losses.[19]

3.1.2 Nodes: storage and validation

Nodes store a local copy of the blockchain. 'Full' nodes store a copy of the entire blockchain, while 'light' nodes hold only a subset of the blockchain in order to verify transactions.[20]

The early applications of blockchain technology are 'open' or 'permissionless' at the node level. Anybody can become a node by downloading and running the relevant software and

[14] See for instance https://www.bitaddress.org/; https://trezor.io/ [accessed 20 December 2017].

[15] See https://blockchain.info/charts/mempool-size [accessed 20 December 2017].

[16] See for instance https://www.coinbase.com; https://blockchain.info/wallet/#/ [accessed 20 December 2017].

[17] For UK examples, see Bittylicious or Coinfloor.

[18] See for instance https://localbitcoins.com/ [accessed 20 December 2017].

[19] Kevin D. Werbach, "*Trust, But Verify: Why the Blockchain Needs the Law*" (2017), Berkeley Technology Law Journal, Forthcoming, https://ssrn.com/abstract=2844409, p. 27 [accessed 20 December 2017].

[20] Nakamoto (2009), supra note 9, p. 5; Vitalik Buterin, "*A Next Generation Smart Contract & Decentralised Application Platform"* (2013), https://github.com/ethereum/wiki/wiki/White-Paper, p. 10 [accessed 20 October 2017].

storing the blockchain archive on their local machine. In practice, only a subset of users will do so, since this requires significant storage space. As of October 2017, running a full node requires 145GB of free disk space.[21] There are almost 10,000 nodes running on the Bitcoin core network as of December 2017.[22]

Nodes discover and maintain connections with other nodes across a P2P network. When they receive a new block from another node on the network, they check that it is valid. This includes a check to prevent users spending the same coin twice (known as double-spending).[23] They do so by checking the proposed transaction against a list of previous, unspent transaction outputs (known as the UTXO database). A transaction output is the result of a transaction, transferring an amount of coin (which is not yet spent) to a new address. The transaction giving ownership to the payer is removed from the UTXO and the transaction giving the currency to the payee is added to the UTXO. All inputs to a transaction must be in this database for the transaction to be valid. If the block is valid, the node adds it to their local copy of the blockchain and broadcasts it to other nodes on the network.

3.1.3 Miners: distributed control over new blocks

When a user signs a transaction, the transaction record is broadcast to nodes across the network. Users can offer 'transaction fees' in return for priority processing of their transactions. Miners assemble transactions into blocks and broadcast those blocks to nodes across the P2P network, so they can append the new block to their local copies.[24] Miners are rewarded for adding new blocks with newly minted crypto-coins, as well as any transaction fees users have offered.[25]

Bitcoin and Ethereum are open or permissionless at the miner level. Any user can become a miner by running mining software on their local machine. Thus, these early applications were designed to be open or permissionless on three levels, as set out in the table below.

Table 3. An open/permissionless application of blockchain technology

No.	Group	Function	Permission
i.	Users	Propose new transactions	Open: Anyone can join the network and send and receive Bitcoin.
ii.	Nodes	Store copies of the DL	Open: Anyone can download the software and run a Bitcoin node.
iii.	Miners	Propose new blocks	Open: Anyone can mine new blocks and broadcast them to the P2P network

3.1.4 Advantages of distributed storage

Storing a blockchain in a distributed manner (i.e. as a Distributed Ledger or DL) has three main advantages. First, it protects data integrity from tampering by any single centralised party. Second, a DL may be less vulnerable to attack since there is no single master copy of the ledger to target. Finally, a DL is resilient, since there is no single point of failure to target

[21] See https://bitcoin.org/en/full-node [accessed 23 October 2017].

[22] See https://coin.dance/nodes [accessed 23 October 2017].

[23] Nakamoto (2009), supra note 9, p. 3.

[24] Nakamoto (2009), supra note 9, p. 3.

[25] Nakamoto (2009), supra note 9, p. 4.

with a denial of service (DoS) attack. Even if several nodes failed, the network would still continue to function.

However, the major challenge for a DL application is ensuring that all of the nodes hold a consistent and up-to-date copy of the blockchain and that participant/system behaviour is valid and appropriate. In blockchain terms, the nodes must achieve ‘consensus’. ‘Full nodes’ start by downloading the latest version of the ledger. Thus, to achieve consensus, the system needs to ensure that each node adds the same new blocks to their local copy. To this end, all nodes must follow the same rules for deciding when to add a new block. These rules are called a **consensus protocol**. The protocol is embedded in the software each node runs. All distributed consensus protocols are designed to operate using asynchronous communication, where messages can be delayed, lost or corrupted.

3.1.5 Consensus protocols

In order for nodes to accept their blocks, miners need to generate new blocks that accord with the consensus protocol. Traditional computer science distributed consensus protocols are **synchronous,** in the sense that they agree on a total ordering of the accepted blocks by adding agreed blocks one at a time. They operate with a known number of nodes that have to reach consensus (the nodes have to formally join a group). The nodes operate in lockstep, ensuring every node has updated its copy of the DL before moving on to the next block. These protocols work where there is only a limited number of nodes (i.e. a maximum of a few tens of nodes),[26] whereas, as noted above in Section 3.1.2, Bitcoin has almost 10,000 nodes.

A traditional consensus protocol can have an elected leader, as in Lamport’s Paxos protocol.[27] The leader indicates the new blocks to be added to the other nodes in the network. However, a leader-based system requires users to trust the elected leader. When a single party controls which blocks are added to the blockchain, they can abuse this power. For instance, they could withhold certain transactions purposefully or discriminate between classes of users in return for payment. In addition, they could engage in double-spending scams, as set out below (see Section 3.1.5 (iii) and (iv) on 51% attacks).

Traditional consensus protocols do not work for cryptocurrencies that seek to operate in trustless environments and with an unknown and potentially large number of nodes. Instead, they use protocols that are **asynchronous**. Nodes do not wait to synchronise with their peers to agree each block in turn. Instead, they proceed to work on the next block based on the best information available to them.

A simple version of an asynchronous consensus protocol would determine that the block with the earliest timestamp wins. While simple to implement, this scheme relies on the trustworthiness of timestamps, which at present cannot be guaranteed system-wide. In such a system, an attacker could forge the timestamp on a block in order to control which new blocks are added.

Since anyone can join as a miner or node, the platform needs safeguards against malicious actors who may try to take control of the ledger. If starting nodes and mining new blocks were costless, an attacker could flood the system with new nodes and newly mined blocks, in what

[26] Marko Vukolic, “*The quest for scalable blockchain fabric: proof-of-work vs BFT replication*”, (2016) iNetSec, Springer, LNCS 9591, pp. 112-125.

[27] Leslie Lamport, “*The Part-Time Parliament”* (1998) ACM Transactions on Computer Systems (TOCS), 16(2), pp. 133-169.

is known as a ‘Sybil’ attack. To defend against this, early cryptocurrencies make mining new blocks costly, by requiring ‘proof of work’, as explained below.

(i) Proof of work

The consensus protocol of Bitcoin and Ethereum requires miners to demonstrate proof of work (PoW) for each new block. To do so, each miner must find the answer to a computationally difficult puzzle.[28] Solving the puzzle can be seen as a demonstration of good faith, since it requires the miner to invest resources (CPU power and electricity) into updating the ledger. Nodes will only accept blocks which contain the solution to the puzzle.

The puzzle works using hash functions. As set out in Section 2.1.3, the header and body of a block are run through a hash function to generate that block’s hash value. To mine a valid new Bitcoin block, the hash value of that block must achieve a particular pattern, namely it must start with a certain number of zeros.[29] To create a valid block, a miner must add a random number (known as a ‘nonce’) to the header of the block such that the resulting hash value fits the pattern. Miners solve this puzzle by trial-and-error, iterating through different nonces until the hash value has the required number of leading zeros.[30] (The higher the number of zeros required, the harder the puzzle.)

The more computational resources (CPU power) a miner devotes to solving the problem, the more likely they will solve it first. In practice, professional miners use dedicated hardware (known as Application Specific Integrated Circuits or ‘ASICs’) and base themselves near sources of cheap electricity to increase their efficiency in mining Bitcoin.

The puzzle is difficult to solve, but easy to verify. This means that when a successful miner broadcasts a block with the solution to other nodes, they can easily check that the miner has solved it, by rehashing the block containing the miner’s nonce. The nodes then propagate the valid block across the peer-to-peer network.

The difficulty of the puzzle and the amount of computational resource devoted to solving it determine the frequency of new blocks. The Bitcoin protocol dynamically adjusts the difficulty of the puzzle, to ensure that a new block is added every 10 minutes.[31]

Although it enhances security, PoW wastes large amounts of energy. All miners expend energy trying to solve the puzzle, but only one of the miners will successfully create a new block. Researchers have estimated that Bitcoin mining consumes 100–500 MW per day, or 3–16 PJ per year.[32] This is similar to the yearly energy expenditure of c. 200,000-1.2m EU

[28] Nakamoto (2009), supra note 9, p. 3.

[29] Nakamoto (2009), supra note 9, p. 3.

[30] Buterin (2013), supra note 20, p. 7. For a helpful video illustrating this process, see Aners Brownworth, MIT, http://blockchain.mit.edu [accessed 20 December 2017].

[31] Nakamoto (2009), supra note 9, p. 3.

[32] Harald Vranken, “*Sustainability of bitcoin and blockchains*”, (2017) Current Opinion in Environmental Sustainability, 28, 1-9.

households.[33] A single Bitcoin transaction requires an estimated 200kWh of energy, compared to around 0.01kWh per Visa transaction.[34]

To reduce the costs of achieving consensus, Ethereum is considering moving to a consensus protocol that combines proof of stake and sharding (i.e. partitioning a large database into many smaller parts).[35] Other consensus protocols have been proposed as well, as set out in the table below.

Table 4: Possible consensus protocols for DLT

No.	Protocol	Description
a)	Proof of work	Participants must invest resources to solve a computational puzzle, before proposing a valid block.
b)	Proof of stake	Participants must show a 'stake' in the system in order to participate in the protocol.[36] An example would be a weighted raffle based on the number of coins held by each participant.[37]
c)	Proof of space	Participants must provide a specified amount of memory to compute the proof, in order to participate in the protocol.[38]
d)	Practical Byzantine Fault Tolerance (PBFT)	A traditional, synchronous distributed consensus protocol. These protocols have been proved correct and their performance has been evaluated.[39] They have been assumed to operate across a few tens of nodes, for which case they perform well, but they have not been used at large scale.[40]
e)	Earliest timestamp wins	Earliest timestamp wins relies on trusting the timestamp generation hardware in the miners participating in the protocol. A hardware technology called 'roots-of-trust' could be built on.[41]

[33] Based on an average EU household estimate of 3,600 kWh in 2014, World Energy Forum, average electricity consumption per electrified household, available at https://wec-indicators.enerdata.net/household-electricity-use.html [accessed 20 October 2017].

[34] ING, "*Why Bitcoin transactions are more expensive than you think*", (2016), https://think.ing.com/downloads/pdf/opinion/why-bitcoin-transactions-are-more-expensive-than-you-think [accessed 20 October 2017].

[35] Vitalik Buterin, "*Ethereum 2.0 Mauve Paper",* (2016) https://cdn.hackaday.io/files/10879465447136/Mauve%20Paper%20Vitalik.pdf [accessed 20 October 2017], p. 1.

[36] See Iddo Bentov, Ariel Gabizon, Alex Mizrahi "*Cryptocurrencies Without Proof of Work*", (2016), in: Clark J., Meiklejohn S., Ryan P., Wallach D., Brenner M., Rohloff K. (eds) "*Financial Cryptography and Data Security*" (2016), Lecture Notes in Computer Science, vol 9604.

[37] Vranken, (2017), supra note 32, 1-9.

[38] See Ateniese G., Bonacina I., Faonio A., Galesi N, "*Proofs of Space: When Space Is of the Essence*" (2014), in: M. Abdalla, R. De Prisco (eds) "*Security and Cryptography for Networks*" (2014), Lecture Notes in Computer Science, vol 8642.

[39] Miguel Castro and Barbara Liskov, "*Practical Byzantine Fault Tolerance and Proactive Recovery*", (2002) ACM TOCS 20(4): 398-461.

[40] Vukolic (2016), supra note 26, pp. 112-125.

[41] See Jatinder Singh and Johan David Michels, '*Blockchain as a Service*' (2017), available on SSRN.

f)	Hashgraphs[42]	Hashgraphs[43] have been proposed as an alternative to PBFT for lower overhead and increased scalability.

(ii) Longest chain wins

As noted above, in a PoW system, nodes will only accept blocks that contain proof of work. However, a problem arises when two miners solve the puzzle around the same time and both broadcast their competing blocks to the network.

Since the protocol is asynchronous, there is no guarantee of consistency between the copies held by the various nodes at any one point in time. The nodes are operating in a widely distributed system on a best-effort basis. As a result of network latency, blocks broadcast by miners arrive at the various nodes at different times. Since different nodes can receive competing blocks at different times, their copies of the DL may differ.

As a result, competing chains can form between nodes. When there are multiple active chains, this is known as 'forking', with the different chains tracking different branches of new blocks.[44]

Bitcoin's consensus protocol resolves this problem by determining that the **longest chain wins**.[45] Imagine that two miners, Alice and Bob, both solve the puzzle to mine the next block around the same time. Each broadcasts their newly mined blocks to the rest of the network. Nodes that hear of Alice's block first add it to their local copy of the blockchain, while nodes that hear of Bob's block first add his. As a result, there are now two groups of nodes working on two different chains: chain A (ending with Alice's latest block) and chain B (ending with Bob's).

All miners on the network will then start mining the next block. Some will work on chain A (incorporating the hash of Alice's block A into their next block), while others will work on chain B (incorporating the hash of Bob's block B). Which chain will be preserved depends on which miner solves the puzzle first. If a miner operating on chain A solves the puzzle first, they will broadcast their new block. Nodes that were on chain A can add the new block to their chain, since its hash value matches the last block they had stored.

However, nodes that were on chain B must cede to the longer chain A. To do so, they discard the incorrect fork and proceed from the newly-arrived, correct chain.[46] In the process, they will discard block B. A discarded block is called an **orphan block** in Bitcoin (or an 'uncle', in Ethereum). Transactions in the orphan block that aren't already reflected in the longer chain return to the pool of transactions waiting to be processed. (Since cryptocurrencies track on-chain assets, the blockchain reflects the authoritative and complete record of transactions. It

[42] Leeman Baird, "*The Swirlds Hashgraph consensus algorithm: Fair, Fast, Byzantine fault Tolerance*", (2016), Swirlds TR 2016-01.

[43] http://www.pingidentity.com/en/resources/white-papers/distributed-session-management-beyond-the-firewall.html?success=true [accessed 20 December 2017].

[44] This 'accidental' and temporary forking is a routine part of agreeing the blockchain. It should not be confused with the deliberate creation of long-term competing forks aimed at creating an alternate cryptocurrency, as described below in Section 3.3.

[45] 'Longest' chain in this sense is calculated based on which chain has the most proof of work built into it. Nakamoto (2009), supra note 9, p. 3. Ethereum employs a similar rule, known as the GHOST protocol, see Buterin (2013), supra note 20, p. 26.

[46] Nakamoto (2009), supra note 9, p. 3.

is less clear how accidental forks and longest-chain-wins rules would work in a system that sought to track off-chain, real-world assets, since the transactions in the orphan block may have already had real-world effects.[47] A solution may be to wait for a number of blocks to be added to the chain before considering a transaction as final.

(iii) Attacking a PoW blockchain

As set out in Section 2.3 above, DLs feature strong security. For instance, regardless of how much hash power an attacker has, they cannot propose new transactions using another user's Bitcoins, since they do not have access to that user's private key.[48] Nor can they spend the same coin twice in new transactions, since nodes would not verify those payments. Further, because of PoW and longest-chain-wins, they cannot easily change transactions in earlier blocks of the blockchain, since they would have to re-hash those blocks (including solving the proof of work) *and* mine new blocks faster than the rest of the network to create the longest chain.[49]

To attack a DL that uses PoW, the attacker must gather more computational power than the rest of the network combined.[50] This is called a '**51% attack**', since the attacker must control more than 51% of the hashing power in the system.[51] The attacker would be likely to consistently solve PoW first and can thus control the addition of new blocks.

As a result, the attacker will not only consistently reap the mining rewards, but can also reject blocks containing certain transactions in order to obtain benefit. In addition, an attacker controlling 51% of the hashing power could scam other users through a so-called 'double-spending' attack, as described in the table below.

Table 5: Example of a 51% attack

A 51% double-spending attack on the Bitcoin network
An attacker with more than 51% of a network's hash power can scam others using a so-called 'private chain'. To do so, the attacker must mine several new blocks in succession, but keep them private, thereby intentionally forking the chain. Honest nodes continue to mine on the public chain. Since the attacker has more hash power, his private chain will become longer than the public chain.[52] The attacker can now fool other users into thinking he has paid them and reverse the transaction.[53] Suppose the latest block in a blockchain is block N. The attacker mines five extra blocks to block N+5, but keeps them secret. The attacker then enters into a contract with another user to pay them Bitcoin in return for a product or service, for example they agree to trade 0.12BTC for US$500. The attacker submits a transaction transferring BTC0.12 to that user, per their agreement. This transaction is recorded in block N+1 on the public chain. The user sees the transaction

[47] See further Reed et al. (2017), supra note 6.

[48] Nakamoto (2009), supra note 9, p. 6.

[49] Nakamoto (2009), supra note 9, p. 3.

[50] Nakamoto (2009), supra note 9, p. 1.

[51] Buterin (2013), supra note 20, p. 8.

[52] Ittay Eyal and Emin Gűn Sirer, "*Majority is not Enough: Bitcoin Mining is Vulnerable*", (2013), available at https://arxiv.org/abs/1311.0243 [accessed 20 December 2017].

[53] Nakamoto (2009), supra note 9, p. 7.

recorded on the public chain and pays the attacker US$500. However, the attacker does not include it in his private chain. Instead, the attacker includes a transaction transferring the Bitcoin to another address they control in block N+1 on the private chain.

Honest miners continue to mine the public chain, adding 4 new blocks (up to block N+4). The attacker then reveals his private chain of five blocks (up to N+5). Since he has the longest chain, the nodes in the network accept the attacker's private chain as valid. As a result, the Bitcoin transaction from the attacker to the user is removed from the blockchain and the attacker makes off with US$500.[54]

The more blocks since the block in which a transaction is recorded, the safer it is from this type of attack.[55] Consequently, users are recommended to wait a certain number of blocks (such as six, meaning c. one hour), before considering high-value transactions as final.[56]

(iv) Feasibility of a 51% attack

A single attacker would have to spend significant resources in order to obtain 51% hashing power. However, in practice, many Bitcoin miners work together as part of large, centrally operated 'mining pools'. Miners in these pools try to solve the puzzle independently, but agree to share any earnings from mining blocks with others in the pool.[57]

These pools concentrate hashing power and could be used for a 51% attack.[58] In 2015, some of the largest Bitcoin mining pools voluntarily split into smaller pools because the top two pools held a majority of the CPU power.[59] As of October 2017, the largest Bitcoin mining pool (AntPool) had c. 20% of the network's hash power.[60]

That said, even if miners have the ability to engage in a 51% attack, they may not have the incentive to do so.[61] Miners have an economic interest in maintaining high Bitcoin prices. They are paid in Bitcoin for their mining and many have invested in dedicated hardware to support the currency. A successful attack on Bitcoin would reduce its value, thereby jeopardising their ability to generate returns on their investment.

Nonetheless, the system remains vulnerable to a politically motivated 51% attack. This introduces a level of political risk, such as from a terrorist attack or from a government

[54] Buterin (2013), supra note 20, p. 8.

[55] Nakamoto (2009), supra note 9, p. 7.

[56] Meni Rosenfeld, "*Analysis of Hashrate-Based Double Spending*" (2014) https://arxiv.org/abs/1402.2009 [accessed 20 December 2017].

[57] Rainer Böhme, Nicolas Christin, Benjamin Edelman, and Tyler Moore, "*Bitcoin: Economics, Technology, and Governance*" (2015), Journal of Economic Perspectives, 29(2): 213-38.

[58] Primavera De Filippi and Benjamin Loveluck, "*The Invisible Politics of Bitcoin: Governance Crisis of a Decentralized Infrastructure*" (2016). Internet Policy Review, Vol. 5, Issue 4.

[59] Larissa Lee, "*New Kids on the Blockchain: How Bitcoin's Technology Could Reinvent the Stock Market*" (2016), Hastings Business Law Journal, Volume 12, Issue 2, 2016, p. 107.

[60] See https://blockchain.info/pools?timespan=24hrs [accessed 23 October 2017].

[61] Nakamoto (2009), supra note 9, p. 4.

commandeering one or more of the big mining pools (which are presently concentrated in China).[62]

3.2 Centralisation and trusted parties

Above, we've seen how early cryptocurrencies were designed as open, permissionless platforms in trustless environments. As a result, they feature distributed storage and adding of new blocks, with complicated consensus protocols and resource-intensive mining to ensure that no single party controls the addition of new blocks.

However, future blockchain-based platforms need not all operate in trustless environments. Instead, they may avoid costly consensus protocols by re-introducing trusted intermediaries that control the blockchain. Such solutions could be applied in cases where there is a level of trust from users, either in a single centralised Trusted Third Party (TTP) or in a small number of trusted nodes. Indeed, closed, permissioned systems may inspire trust, since parties can limit participation so as to keep malicious actors off the platform.

In many instances, consumers may prove willing to trust reputable companies and citizens may trust government agencies with blockchain management.[63] After all, they currently trust many centralised record-keepers to maintain accurate records, without the use of blockchains.

For example, users currently trust Twitter not to change their past tweets. If Twitter stored tweets on a blockchain, it would make past tweets tamper-evident. However, Twitter didn't need a blockchain to gain users' trust, since users understand that tampering with past tweets would damage the company's reputation. This gives it a strong incentive to refrain from doing so. Users may similarly prove willing to trust reputable operators of closed, permissioned blockchain platforms.

3.2.1 A centralised Trusted Third Party model

Governments are exploring the use of blockchain for large, centrally managed databases, such as land registers.[64] Provided citizens are willing to trust their government, such a blockchain need not be stored at multiple nodes across a P2P network. Instead, it could be stored centrally with a government agency, such as the land registry, acting as a TTP.[65] If the blockchain is publicly visible, this would increase the transparency of the registry's record-keeping, without requiring distributed storage.

Estonia is often considered a pioneer in the use of blockchain technology. Since Estonian independence in 1991, citizens have been issued a proof of identity via a PKI-based identity card.[66] Since 2012, blockchain-like approaches have been used to assist in maintaining the integrity of data and transactions regarding national health, judicial, legislative and other social

[62] Angela Walch, "*The Bitcoin Blockchain as Financial Market Infrastructure: A Consideration of Operational Risk*" (2015), 18 NYU Journal of Legislation and Public Policy 837.

[63] See further Izabella Kaminska, "*Blockchain's governance paradox*", Financial Times, Alphaville, (2017), https://ftalphaville.ft.com/2017/06/14/2190149/blockchains-governance-paradox/ [accessed 27 November 2017].

[64] See for instance Stan Higgins, "*UK Land Registry Plans to Test Blockchain in Digital Push*" (2017), https://www.coindesk.com/uk-land-registry-plans-test-blockchain-digital-push/ [accessed 29 November 2017].

[65] See Vitalik Buterin, "*On Public and Private Blockchains"*, (2015), https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/ [accessed 19 October 2017].

[66] See https://e-estonia.com/solutions/e-identity/id-card/ [accessed 20 December 2017].

services.[67] A small number of dedicated nodes are involved in agreeing on the blocks in the ledger, with a view to checking that the constituent system components are operating correctly.

To bolster trust in a TTP-based system, archival copies of the DL could be stored at independent nodes, such as with an Ombudsman. Any tampering by the record-keeper would then be evident through comparison with the independently-stored copies. (Users would still need to trust the TTP and independent archival nodes not to collude against their interests.)

Table 6: A 'closed' or 'permissioned' centralised blockchain model with a TTP

No.	Group	Function	Permission
i.	Users	Propose new transactions	Open: anyone can join, for instance to register their land on the registry.[68]
ii.	Nodes	Store copies of the DL	Closed: the TTP acts as a single node, holding the master copy.
iii.	Miners	Propose new blocks	Closed: the TTP is the only party who updates the ledger.

Alternatively, the ledger could be widely distributed across a P2P network, but mining of new blocks could be entrusted to a TTP. In this case, the TTP would be unable to alter previous transactions.

3.2.2 A trusted nodes model

Commercial entities such as banks are also exploring ways to use blockchain to settle payments amongst themselves.[69] In such cases, multiple parties could set up a system with a defined number of 'trusted nodes' who each store copies of the blockchain.[70] Since the number of trusted nodes is known, they could follow a traditional, synchronous consensus protocol (see Section 3.1.5, above).

Such closed platforms restrict control over the blockchain by limiting permission to store the ledger and add new blocks to a small number of trusted parties. The platform can limit permission in respect of users, nodes, and miners, depending on the level of trust and desired functionality. The table below shows a possible configuration.

Table 7: A 'closed', 'permissioned' distributed blockchain model with trusted nodes

No.	Group	Function	Permission
i.	Users	Propose new transactions	Closed: only authorised parties can join and participate.
ii.	Nodes	Store copies of the DL	Closed: only trusted nodes store copies of the ledger.

[67] See https://e-estonia.com/wp-content/uploads/faq-a4-v02-blockchain.pdf [accessed 20 December 2017].

[68] In practice, in some jurisdictions, users may need to interact with a land registry through a lawyer or notary.

[69] For instance, R3 leads a consortium of over 100 large financial institutions to research and develop blockchain-inspired applications for the financial system. See https://www.r3.com/about/ [accessed 20 December 2017].

[70] Buterin (2015), supra note 65.

iii.	Miners	Propose new blocks	Closed: trusted nodes act as miners according to a consensus protocol.

The result is a permissioned, 'narrowly distributed' platform with a 'shared' (as opposed to distributed) ledger that should have lower running costs than permissionless widely distributed platforms, since there are fewer nodes. Moreover, the nodes may be able to process transactions more quickly, since transactions can be verified and blocks mined by a small number of trusted nodes, each with high processing power (as opposed to by thousands of distributed nodes).[71] As a result, they may provide a more scalable platform. However, the TTP or trusted nodes will need to invest in traditional security, to protect against hackers gaining access to the ledger or DoS attacks from taking down nodes.

3.3 Governance and control

A final aspect of trust relates to control over platform design: who has the power to determine how the platform operates? This raises issues of blockchain governance: who can change the platform and under what circumstances should past entries in the ledger be changed. In this section, we first consider protocol changes, before turning to the reversibility of past transactions, and the role of service providers.

3.3.1 Developers and protocol changes

From a technical perspective, governance involves the groups discussed above – users, miners and nodes – as well as a fourth group, the *developers*, who produce the software that nodes and miners run to support the blockchain.

In practice, blockchain governance will differ per platform. In the current cryptocurrency context, these groups form a community around a shared interest in maintaining (and increasing) the value of the coin. Each group can use informal governance mechanisms to express its preferences. Examples include discussions of technical improvement proposals,[72] user-wide votes on protocol changes,[73] miner-implemented 'soft forks',[74] and ultimately, hard forks that lead to competing, alternative cryptocurrencies (like those set out in Table 7 below). The result is a system of checks and balances that should generally favour incremental protocol change.

For example, the Bitcoin source code is developed in an open-source manner using GitHub.[75] Anybody can view the code and propose improvements. However, only a small group, known as the 'core developers' have password access to make changes to the version of the software known as 'Bitcoin Core'.[76] Thus, the Bitcoin platform is not fully open/permissionless at this

[71] Buterin (2015), supra note 65.

[72] See https://github.com/bitcoin/bips; For Ethereum Improvement Proposals, see https://github.com/ethereum/EIPs/issues/ [accessed 20 December 2017].

[73] http://carbonvote.com/; http://v1.carbonvote.com/ [accessed 20 December 2017].

[74] See De Filippi and Loveluck (2016), supra note 58.

[75] Github is a web-based service that provides cloud-hosted distributed revision control and source code management for user-uploaded software projects. Anyone can register an account, create a software repository, and/or begin suggesting edits to other public repositories. See www.github.com [accessed 20 December 2017].

[76] Catherine Martin Christopher, "*The Bridging Model: Exploring the Roles of Trust and Enforcement in Banking, Bitcoin, and the Blockchain*" (2016) 17 Nevada L.J. 1, p. 12.

code development level. Ethereum code is similarly developed through an open-source process, though the Ethereum Foundation has ultimate responsibility for code changes.[77] Ethereum's founder, Vitalik Buterin is a member of the foundation and actively participates in governance discussions.

Table 8. The Bitcoin application of blockchain technology

No.	Group	Function	Permission
i.	Users	Propose new transactions	Open: Anyone can join the network and send and receive Bitcoin.
ii.	Nodes	Store copies of the DL	Open: Anyone can download the software and run a Bitcoin node.
iii.	Miners	Propose new blocks	Open: Anyone can (attempt to) mine new blocks and broadcast them to the P2P network
iv.	Developers	Change the Bitcoin 'core' software	Closed: only a small group of core developers can change the 'core' code.

Changes to the code encompass technical issues (like bug fixes), but also platform design decisions that have direct operational impact, such as the size of the blocks that determine the system's throughput. As a result, the platform requires users to place some level of trust in the developers.[78]

Nonetheless, the developers' power is limited by the possibility of 'hard forks'. First, the source code is transparent, as it is available for anyone to access online.[79] Second, anyone can take the code from GitHub and 'fork' it, meaning they can write a new version (typically an extension) of the software and make it available for others to download. There are several versions of the Bitcoin software available as of 2017 that differ in terms of properties or functionality.

This means that anyone can propose new technical features or platform designs, by writing new software with e.g. a different block size or consensus protocol. The deliberate creation of a new version of software that is incompatible with the existing consensus protocol and/or previous data structures is called a 'hard fork'. In the case of a hard fork, miners and nodes must decide which version of the software they want to run. Different miners and nodes may choose to run different versions of the software. If they do, this creates two separate blockchains that track two different cryptocurrencies. Both start from the last block before the fork, but will add different blocks going forward. As a result, they will track two different cryptocurrencies. Examples of such forks are set out in Table 9, below.

In practice, software is forked to provide some different functionality or capability. But of course to have impact, the resulting software must be adopted. This means that each fork will need to attract miners, nodes, and users to their version of the software. The value of each currency,

[77] Nick Tomaino, "*The Governance of Blockchains*", (2017), https://thecontrol.co/the-governance-of-blockchains-5ba17a4f5da6 [accessed 31 October 2017].

[78] Christopher (2016), supra note 76, p. 36; De Filippi and Loveluck (2016), supra note 58.

[79] There appears to be much scrutiny over the Bitcoin codebase, which is perhaps unsurprising given the direct financial incentives involved. However being open-source, while a form of mitigation, does not guarantee that major code issues will be identified before deployment or exploitation. Prominent examples include the Ethereum DAO crisis discussed below and the OpenSSL Heartbleed vulnerability: http://heartbleed.com/ [accessed 20 December 2017].

resulting from a fork, is determined by supply and demand on currency exchanges. This, in turn, depends on each currency's ability to attract miners, intermediaries, and ultimately, users.[80]

Table 9: Bitcoin block size forks

Bitcoin XT, August 2015
In August 2015, two core developers launched a new version of the Bitcoin client software, called 'Bitcoin XT', with an 8MB block size. To gauge miner support, they released the software as a soft fork first.[81] If at least 75% of newly mined blocks signalled support for the new software, it would initiate a hard fork.[82] However, Bitcoin XT failed to gain mass adoption.
Bitcoin Cash, August 2017
In August 2017, mining pool ViaBTC launched 'Bitcoin ABC', also with an 8MB block size. The new fork supported an alt-coin called 'Bitcoin Cash' that competes with Bitcoin. As of November 2017, Bitcoin Cash was worth around US$2,000 (compared to Bitcoin's over $18,000) and there were around 1,400 nodes running on the Bitcoin Cash network (compared to almost 10,000 on the original Bitcoin network).[83] Since both systems feature the same kind of PoW, mining pool operators can switch between the chains when mining new blocks, depending on which is more profitable.[84]
SegWit and SegWit2x, August 2017
To improve the throughput of the Bitcoin Core network, a group of core developers proposed to remove digital signature data and place it in a separate file. The proposal was called 'SegWit' (or 'segregated witness').[85] Since this didn't require a change to the consensus protocol, the proposal was implemented as a soft fork, dependent on the majority of hash power 'signalling' in favour of the update.[86] SegWit activated in August 2017.[87] The developers further proposed to double the block size three months after SegWit's activation. This second proposal was called 'SegWit2x'. However, the proposed block size increase proved controversial and, in November 2017, the developers abandoned SegWit2x.[88]

[80] See Kyle Torpey, "*You Really Should Run a Bitcoin Full Node: Here's Why*", (2017) https://bitcoinmagazine.com/articles/you-really-should-run-full-bitcoin-node-heres-why/ [accessed 31 October 2017]; Aaron van Wirdum, "*On Consensus, or Why Bitcoin's Block-Size Presents a Political Trade-off*", (2016), https://bitcoinmagazine.com/articles/on-consensus-or-why-bitcoin-s-block-size-presents-a-political-trade-off-1452887468/ [accessed 31 October 2017].

[81] See De Filippi and Loveluck, (2016), supra note 58 .

[82] Blocks mined by miners running the Bitcoin XT software would be automatically 'signed' XT.

[83] https://cash.coin.dance/nodes; https://coin.dance/nodes [accessed 20 December 2017].

[84] I.e. depending on the value of each coin, the difficulty of the puzzle, and the level of competition.

[85] See https://github.com/bitcoin/bips/blob/master/bip-0141.mediawiki; Mohit Mamoria, "*Your ultimate guide to the upcoming fork that's splitting the Bitcoin community*" (2017) https://thenextweb.com/contributors/2017/07/24/ultimate-guide-upcoming-fork-thats-splitting-bitcoin-community/ [accessed 31 October 2017].

[86] See https://github.com/bitcoin/bips/blob/master/bip-0091.mediawiki [accessed 20 December 2017].

[87] See Aaron van Wirdum, "*Segregated Witness Activates on Bitcoin: This is What to Expect",* Bitcoin Magazine, https://bitcoinmagazine.com/articles/segregated-witness-activates-bitcoin-what-expect/ [accessed 31 October 2017].

[88] See Alyssa Hertig, "*2x Called Off: Bitcoin Hard Fork Suspended for Lack of Consensus*" (2017), Coin Desk, https://www.coindesk.com/2x-called-off-bitcoin-hard-fork-suspended-lack-consensus/ [accessed 31 October 2017].

In more closed, permissioned instantiations of a DL, the code and its updates might be defined by agreements between the participants involved, including contractual agreements. Governance issues should be more straightforward where fewer parties are involved.

3.3.2 Reversibility of past transactions

A second issue of blockchain governance relates to the reversibility of past transactions recorded on the blockchain. As discussed above, the blockchain aims to create a persistent, tamper-evident record of relevant transactions. However, the ledger is not technically 'immutable'. The nodes of any given platform could undertake a coordinated effort to 'correct' their local versions of the ledger and undo specific past transactions they considered inappropriate. (The nodes would effectively perform a 'hard fork' to a new version of the blockchain without the offending transactions.) Although this goes against the general aim of creating a persistent record, there may be exceptional circumstances in which leaving past transactions unaltered is so costly that it would jeopardise the success of the platform.

How easy this is to achieve differs according to the blockchain platform's design. On a widely distributed platform, establishing a hard fork is costly, both in terms of arranging cooperation between nodes and re-hashing previous blocks where required.[89] Further, there is no guarantee that nodes will agree to a hard fork. The difficulties are illustrated by the example of the Ethereum DAO, set out below in Section 6.2.2.

Conversely, with a centralised platform, it should be easier to arrange a corrective fork. The TTP or small group of trusted nodes can agree to correct the ledger (which would also be less costly if the consensus protocol does not require PoW). As a result, centralised platforms feature better reversibility.

A potentially simpler way to reverse the effects of a past transactions is for the parties to enter a 'correcting transaction', to be recorded in a subsequent block. This second transaction would negate the effect of the initial transaction. However, this requires either both parties' cooperation, or at least access to the counter-party's private key if enforced externally by a third-party. As a result, this is not an option for contested reversals. In such cases, a party looking to change a past transaction must turn to the appropriate governance mechanism (i.e. either proposing a network-wide hard fork to nodes or submitting a request to a TTP, depending on the platform's design).

There may well be ways for automating aspects of reversibility, such as corrective operation that can occur automatically through the use of smart contracts (see Section 5, below).

3.3.3 Service provider considerations

A final trust issue relates to the roles of service providers involved in the blockchain platform.

As noted above, many users rely on intermediaries, such as online (or 'hot') wallets, as an interface to a blockchain system. This requires users to place some degree of trust in such service providers, since the provider determines the nature of the service and manages its operation.

However, service providers may also be directly involved in the operation of a blockchain platform. In such cases, sometimes termed '*Blockchain-as-a-Service'* (BaaS), a third party

[89] Jeremy M. Sklaroff, "*Smart Contracts and the Cost of Inflexibility*" (2017), University of Pennsylvania Law Review, Vol. 166, p. 14.

provides aspects of the platform’s infrastructure.[90] This can range from a completely managed DL, to the hosting of particular nodes, as well as supporting infrastructure such as identity (key) management services. While BaaS is an emerging area, the major cloud providers are active in this space. The current focus is on capturing existing B2B networks through closed, permissioned platforms, for instance to offer improved efficiency and transparency of supply chains. In future, BaaS offerings may target individual consumers in more open, permissionless blockchain-driven applications.[91]

BaaS raises additional trust and governance considerations, given the service provider’s control over the technical infrastructure. Ultimately, any concerns will depend on the precise nature of the service offered, and the degree of power the service provider has over the entire system. A provider managing a single node as part of a large, federated network of nodes raises different considerations to a service provider that hosts all the nodes of a network. There are strong incentives for service providers to ensure the integrity of their platform, as their reputation is crucial to the longevity of their business. Providers may also be able to use advances in technical mechanisms, such as trusted execution environments (‘secure enclaves’) that aim to provide a technical guarantee that specific code was executed.[92]

4. Visibility and identity of participants

This section considers two further questions of blockchain platform design: (i) who can see the record of transactions stored on the blockchain and (ii) to what extent can the blockchain’s users be identified.

4.1 Visibility of the record: public and private platforms

Early cryptocurrencies were permissionless at the level of nodes and miners, meaning anyone could download the entire blockchain archive. Further, all new transactions are broadcast to all nodes for the purpose of mining new blocks. As a result, all blockchain transactions are publicly visible.[93]

For instance, the latest Bitcoin transaction records are available to view online.[94] This record includes the sending Bitcoin address, the receiving Bitcoin address, the amount of Bitcoin, and a timestamp. A Bitcoin address is a 26-35 digit combination of letters and numbers which is generated based on a hash of a user’s public key.[95] Thus, the public can see that a particular address is sending an amount of Bitcoin to another address, but without information linking the addresses to any real-world identities.

[90] See Singh and Michels (2017), supra note 41.

[91] For example, by hosting an individual’s node that acts on a user’s behalf as part of a broader peer-to-peer marketplace; see Singh and Michels, supra note 41.

[92] See Singh and Michels (2017), supra note 41.

[93] Nakamoto (2009), supra note 9, p. 6.

[94] To see the latest Bitcoin transactions submitted for inclusion into blocks, see https://blockexplorer.com/ [accessed 20 December 2017].

[95] For further details on addresses see https://blockgeeks.com/guides/blockchain-address-101/ [accessed 20 December 2017].

This public visibility of all transaction data may prove a barrier to the adoption of widely distributed platforms.[96] For certain use cases, parties may not be willing to share data about their transactions publicly, particularly where access to transaction data may provide a commercial advantage. For instance, if a blockchain were used for trading shares, competitors may be able to discern each other's trading patterns.[97]

Conversely, more centralised systems can limit visibility of the blockchain to certain parties, resulting in a 'private' blockchain. If a permissioned system is stored by a TTP, it can withhold access to the blockchain archive and grant permission to view blocks or entries only in specific cases. However, such a system asks for a higher degree of trust from users, since only the TTP will know whether the blockchain has been tampered with.

Similarly, in permissioned systems using trusted nodes, each trusted node stores and is able to view the entire blockchain. Since the blockchain is stored across multiple nodes, this type of system can limit visibility, while being more tamper-evident than systems with a single TTP.

4.2 Identity of participants

4.2.1 Pseudonym-based platforms

In the physical world, we use birth certificates, passports, and other Government-issued official documents to prove our identity. On online networks, we often start from real-world documents, presenting them as credentials to institutions, for example when enrolling as a student at a university or registering with Airbnb. This is the basis for authentication, proving that 'you are who you say you are', and allowing an electronic identity to be associated with a validated person and their real-world identity. Part of the establishment of an electronic identity is to associate electronic credentials such as a secret password or biometrics with the validated person.

As noted above in Section 2.2, blockchain users use PKI to authenticate their identity. Unlike Government-issued IDs, revealing a public key need not in general reveal the real-world identity of the associated party. As a result, even though each transaction record is public, Bitcoin and Ethereum users enjoy a level of anonymity by using their addresses as pseudonyms.[98]

To bolster their privacy, users could generate a new public key and address for each new transaction.[99] (The public could still see that the Bitcoins had moved to a new address, but they would not know who controlled the new address.) There are also other online services to help users mask their transactions, known as bitcoin 'mixing' or 'laundry'.

In most cases, outside observers will not be able to determine the real-world identity of the parties to a transaction. Nonetheless, the two can be linked through a user's voluntary

[96] There are platforms that aim to use distributed blockchains without a publicly visible record. For example, a project called "Zcash" features a blockchain that only stores encrypted data and uses a technique called 'zero knowledge proofs' for validation, without revealing the data. See https://z.cash/technology/index.html [accessed 20 December 2017].

[97] See ESMA, "*Report: The Distributed Ledger Technology Applied to Securities Markets*", (2017), p. 11 (citing concerns this may lead to front-running or price manipulation).

[98] Nakamoto (2009), supra note 9, p. 6.

[99] Jaume Barcelo, "*User Privacy in the Public Bitcoin Blockchain*", (2014) http://www.dtic.upf.edu/~jbarcelo/papers/20140704_User_Privacy_in_the_Public_Bitcoin_Blockchain/paper.pdf , p. 1 [accessed 20 December 2017].

disclosure. For instance, if a user pays an online merchant in Bitcoin, that merchant will likely ask for the customer's name, email address, and possibly real-world address.[100] Further, it may be possible to use other methods to de-anonymise users, such as by linking public keys to the IP addresses where the transactions are generated.[101] Anyone who knows the real-world identity associated with a Bitcoin address, can review the entire transaction history associated with that address.[102]

4.2.2 Real-world identity-based platforms

Future blockchain platforms may instead require the real-world identity of participants to be established, rather than hidden. This will likely depend on the functionality they seek to offer and on regulatory requirements. For instance, a number of organisations are promoting the use of blockchain to allow individuals to manage their identity online.[103]

In such cases, an individual would probably need to verify their identity using traditional methods to a TTP, who would then vouch for that individual's identity. For instance, a Certification Authority (CA) might act as TTP and offer a service to associate public keys with identities in an identity certificate. A new user would request such a certificate by generating a key pair and sending the public key to the CA together with a request for a certificate to be generated. The CA holds its own private key and can sign and validate the certificate.

For example, the X.509 standard defines public key certificates that are used in many Internet protocols such as TLS/SSL, which forms the basis for https. An X.509 certificate contains a public key and an identity (hostname, organisation or individual) and is either signed by the CA or is self-signed. If the CA has signed the certificate the CA can vouch for its validity, using its own private key. If a certificate is validated by the issuer, and the software examining the certificate trusts the issuer, it can use the public key to communicate securely with the subject. A certificate can only be validated by the entity that signed it.

5. Smart contracts

A further platform design consideration is to what extent the platform will support smart contracts. Introduced by Szabo in 1994, a smart contract is "a computerised transaction protocol that executes the terms of a contract".[104] A smart contract essentially represents computer programs that automatically bring about some action(s), such as carry out transfers of, or execute other actions relating to, digital assets according to a set of pre-specified

[100] See Fergal Reid and Martin Harrigan, "*An Analysis of Anonymity in the Bitcoin System*" (2011), Security and Privacy in Social Networks, 3, p. 15.

[101] Alex Biryukov, Dmitry Khovratovic, Ivan Pustogarov, "*Deanonymisation of clients in Bitcoin P2P network*" (2014), In: "*Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*".

[102] Barcelo (2014), supra note 99, p. 2.

[103] As described in Reed et al. (2017), supra note 6.

[104] Nick Szabo, "*Smart Contracts*", (1994) http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/smart.contracts.html [accessed 8 December 2017];

Nick Szabo, "*The Idea of Smart Contracts*", (1997) http://www.fon.hum.uva.nl/rob/Courses/InformationInSpeech/CDROM/Literature/LOTwinterschool2006/szabo.best.vwh.net/idea.html [accessed 25 October 2017].

rules.[105] Smart contracts can be used to automate agreements between parties according to the set of instructions written into their code.

In a blockchain context, smart contracts relate to on-chain occurrences; in many ways resembling the stored procedures and/or triggers (*event-condition-action* rules), which are common in relational databases.[106] The more a blockchain platform supports smart contracts, the more scope users have to use the platform for different purposes. Developers who want to create a simple blockchain for a predetermined purpose may opt to offer limited support for smart contracts, or only support smart contracts made of predefined components. In Bitcoin, the functionality relating to transactions is brought about through a series of scripts (programs).[107] These scripts can be used to control Bitcoins (i.e. unspent transaction outputs or UTXOs, see Section 3.1.2), thereby supporting basic smart contracts with limited functionality.[108]

In contrast, developers can instead allow their platform to support many types of contract. A key feature of Ethereum is that it supports 'Turing-complete' smart contracts.[109] Turing-completeness means that the language is as fully-featured as a general programming language and is not restricted in what it can compute. As a result, Ethereum can be used as a platform to run a wide array of applications expressed in smart contracts,[110] as opposed to Bitcoin's *Script* which is deliberately limited in scope.[111] Given this, some authors differentiate between Ethereum as an 'application stack' platform and Bitcoin as a simple digital currency.[112] A combination of smart contracts that provide a specific functionality is also called a 'decentralised application' (or *'DApp'*).

In terms of operation, as part of the block mining process, a transaction (i.e. a ledger entry) being processed might trigger a smart contract. This results in the smart contract's code being executed, where the triggering and any resulting transactions (i.e. contract outputs) are encoded in the block (along with any other transactions forming the block). As part of the block validation process, all nodes will also execute and verify the contract with the respect to its triggering and output transactions to ensure that the smart contract was properly executed and that the transactions accord. In other words, as per Buterin: "if a transaction is added into block B the code execution spawned by that transaction will be executed by all nodes, now and in the future, that […] validate block B".[113] In this way, it is said that smart contracts enable

[105] Szabo (1994), supra note 104; Szabo (1997), supra note 104; Buterin (2013), supra note 20, p. 1.

[106] Konstantinos Christidis and Michael Devetsikiotis, "*Blockchains and Smart Contracts for the Internet of Things*," (2016) in *IEEE Access*, vol. 4, pp. 2292-2303.

[107] Aviv Zohar, '*Bitcoin: under the hood*', (2015) Commun. ACM 58, 9, pp. 104-113.

[108] Buterin (2013), supra note 20, pp. 11-12.

[109] Buterin (2013), supra note 20, pp. 1, 12-13.

[110] Buterin (2013), supra note 20, p. 34.

[111] https://en.bitcoin.it/wiki/Script [accessed 20 December 2017].

[112] Paolo Tasca, Thayabaran Thanabalasinghamand, Claudio Tessone, "*Ontology of Blockchain Technologies. Principles of Identification and Classification*" (2017), https://ssrn.com/abstract=2977811 [accessed 20 December 2017].

[113] Buterin (2013), supra note 20.

consensus regarding computation,[114] given that generally all nodes in a network will verify a contract's execution.

Since smart contract code is run on all nodes on the network, it is publicly visible to all participants. This means that generally a smart contract's code can be inspected and re-used by other participants.

6. Legal implications

Above, we have seen that blockchain technology can be applied in various ways to create platforms with different features, including with regard to:

(i) who can propose new transactions to be added to the ledger;
(ii) who stores a copy of the ledger;
(iii) who can add new blocks to the ledger;
(iv) who can view the ledger;
(v) whether users are identifiable; and,
(vi) who controls the platform's underlying software.

In this section we argue that these variables may be significant from a legal perspective. Therefore, it is important for legislators, regulators, and lawyers to understand these technical permutations.

To illustrate this point, we consider the implications for a handful of legal areas. We first consider law enforcement, including criminal law and anti-money laundering. We then look at the ways in which blockchain technology is being used to create structures that resemble existing legal concepts, including contracts, companies, and securities. Finally, we consider issues of EU intellectual property, data protection law, and competition law. For each area, we illustrate how the legal analysis is affected by the way in which blockchain technology has been deployed in a specific use case.

6.1 Law enforcement and regulatory access to data

Their inherent pseudonymity (see Section 4.2.1 above) may make cryptocurrencies attractive to actors that wish to hide financial transactions from law enforcement agencies and regulators. For instance, Bitcoin has been used to facilitate illicit drug and weapons trafficking on 'dark web' markets, as well as for money laundering.[115] Similarly, the attackers behind the May 2017 WannaCry ransomware demanded payment in Bitcoin.[116]

The ease (or difficulty) with which law enforcement agencies and regulators can gain access to data in cases involving blockchain-based tokens will depend on a platform's design. Generally speaking, obtaining access should be more straightforward in cases of centralised platforms that hold information on their users' real-world identities. As with traditional financial institutions, authorities can request information from platform operators. To facilitate regulatory

[114] Ahmed Kosba, Andrew Miller, Elaine Shi, Zikai Wen and Charalampos Papamanthou, "*Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts*," 2016 IEEE Symposium on Security and Privacy (SP), San Jose, CA, 2016, pp. 839-858.

[115] Primavera de Filippi, "*Bitcoin: a regulatory nightmare to a libertarian dream*" (2014), Internet Policy Review, 3(2), pp. 3-4.

[116] See https://www.theguardian.com/technology/2017/aug/03/wannacry-hackers-withdraw-108000-pounds-bitcoin-ransom [accessed 20 December 2017].

oversight, new regulation targeting centralised platforms could require the platform operators to know their customers.[117]

Obtaining access is more complicated in relation to widely distributed platforms such as the early cryptocurrencies. On such platforms, users may be identified merely by disposable addresses and there is no central register matching addresses to real-world identities. Further, unlike centralised platforms, there are no central parties to target with regulations, so imposing know-your-customer (KYC) requirements in respect of widely distributed platforms is difficult. Nonetheless, as set out in Section 3.1.1, many users interact with cryptocurrencies through trusted intermediaries, such as wallet services and exchanges. These intermediaries present targets for regulation. For instance, some jurisdictions require cryptocurrency exchanges to obtain a licence. Licence conditions include a requirement to obtain verified information about customers, similar to anti-money laundering (AML) and KYC rules for financial organisations.[118] In addition, exchanges could be required to report substantial transactions, or patterns of transactions, which indicate suspicious activity. Any regulation should be designed carefully so as not to stifle competition by raising barriers to entry for innovative start-ups.[119]

The challenge presented to regulators by widely distributed blockchains is similar to that presented by peer-to-peer file-sharing, where the absence of a central server may frustrate law enforcement and regulators' ability to prevent copyright violations.[120] In response to such file-sharing, many jurisdictions have targeted intermediaries, for instance by requiring ISPs to restrict their subscribers' access to websites that index copyrighted content (such as ThePirateBay.org).

Imposing KYC requirements on intermediaries may facilitate identification of parties that use such intermediaries. However, it would not provide traceability in all cases. More technologically-adept users can avoid intermediaries, for instance by acquiring coins through mining or buying coins from others in person (so as to avoid exchanges), or by running their own nodes and using hardware wallets (to avoid wallet services).

Nonetheless, in many cases, blockchains are likely to provide a robust evidential trail of transactions that will enable law enforcement officials to track coin transfers from Bitcoin address to address, since each transaction record will include a sending and a receiving address, as well as the amount of Bitcoin. This may allow authorities to trace transactions back to specific individuals, or at least to identify the main services used to launder illicit gains. For instance, in 2017, Alexander Vinnik was arrested and charged with laundering funds from the hack of a Bitcoin exchange (Mt Gox) through two exchanges he owned.[121] While tracking

[117] Kevin D. Werbach and Nicolas Cornell, "*Contracts Ex Machina*" (2017), 67 Duke Law Journal, Forthcoming, https://ssrn.com/abstract=2936294, p. 54 [accessed 20 December 2017].

[118] See for instance New York States Bitlicence framework; http://www.zdnet.com/article/austrac-gets-the-legal-nod-to-monitor-bitcoin-ethereum-exchanges/ [accessed 20 December 2017].

[119] Werbach (2017), supra note 19, p. 29.

[120] Michael Abramowicz, "Peer-to-Peer Law, Built on Bitcoin" (2015), GWU Law School Public Law Research Paper, pp. 10-11, http://ssrn.com/abstract=2573788 [accessed 20 December 2017].

[121] See https://www.theguardian.com/technology/2017/jul/27/russian-criminal-mastermind-4bn-bitcoin-laundering-scheme-arrested-mt-gox-exchange-alexander-vinnik [accessed 20 December 2017].

may be complicated by the use of so-called 'coin mixing' services[122], the intermediaries providing such services may be further targets for regulation.

6.2 Quasi-legal constructs: smart contracts, DAOs, and ICOs

6.2.1 Smart contracts and contract law

(i) Formation: is a smart contract a contract?

The term 'smart contract' is arguably misleading, since it refers to automatically executing computer code (see Section 5). This raises the question whether a smart contract qualifies as a legal contract.

A legal contract is commonly defined as a legally enforceable agreement or promise. It is typically formed through voluntary offer and acceptance, as well as (in common law jurisdictions), consideration: the value offered by each party. Many types of contracts can be established in any form: orally, in writing, or through actions, such as assenting to terms in electronic media by clicking, also known as 'clickwrap agreements'. However, the law in some countries may require that certain types of contracts (such as real estate transactions) be conducted in a specific form, for instance in writing and witnessed by a notary.

One could argue that a smart contract is not a legally enforceable 'promise', but an automated mechanical process.[123] While this may be true at the level of the computer-readable code, it is unlikely to reflect smart contract use in practice. In reality, the creator of a smart contract will ordinarily need to explain his offer to human counter-parties in human-intelligible language. This explanation can form the basis of the agreement between the parties and thereby determine the terms of the contract.

For example, assume party A sets up a crypto-asset exchange contract on Ethereum. The smart contract's instructions are that if a counter-party pays 1 ETH into a specified address, it will in return provide them with a 'CryptoKitty' (a unique cartoon cat, stored on the blockchain).[124] In order to attract human counter-parties to this offer, A will have to explain it to them in a language they can understand, for instance through a website (like https://www.cryptokitties.co/) or other user interface. In doing so, A will communicate the details of her offer. By engaging with the smart contract (in this instance, paying 1 ETH to a specific address), B expresses acceptance, assenting to the terms of A's offer as explained in the user interface.[125] Even though the underlying smart contract code may technically be visible,[126] many users will likely *de facto* rely on A's other communications.

[122] Coin mixing services combine several transactions so as to obfuscate transaction chains by obscuring the senders and recipients of each individual transaction.

[123] Werbach and Cornell (2017), supra note 117, p. 21.

[124] CryptoKitties is a game that runs on the Ethereum blockchain, where users buy and trade drawings of kittens. Users have spent over US$ 1m playing this game to date. See https://techcrunch.com/2017/12/03/people-have-spent-over-1m-buying-virtual-cats-on-the-ethereum-blockchain/ [accessed 20 December 2017].

[125] Werbach and Cornell (2017), supra note 117, p. 46.

[126] For the CryptoKitty smart contract code, see https://etherscan.io/address/0x06012c8cf97bead5deae237070f9587f8e7a266d#code [accessed 20 December 2017].

Parties offering smart contracts may seek, via disclaimers or other provisions in their Terms of Service, to limit their contractual obligations solely to the computer-readable code.[127] For instance, in 2016, the website of '*The DAO'* (a collection of smart contracts set up on the Ethereum blockchain) stated explicitly in its terms of service that the smart contract was the valid legal agreement. Any human-readable documents or explanations on the accompanying website were "merely offered for educational purposes" and could not override the terms of the code.[128]

Whether such disclaimers are legally binding may be tested in court. On the one hand, standard contracts commonly include provisions that exclude all representations outside of the contract terms and provide a hierarchy between various legal documents. Moreover, given that smart contract code is typically visible to all nodes on the blockchain, their function should be legible to a skilled programmer/developer. On the other hand, simply publishing machine-readable code may not provide sufficient notice of contractual terms to non-expert counter-parties.[129] Further, if there is an obligation that some term should be 'fair and reasonable' (typically an exclusion or limitation of liability), that has to be assessed individually in relation to the parties. A term might be generally fair and reasonable, but not so in the particular circumstances.[130]

Given the above, the code may be more likely to be binding in B2B-cases, than in B2C or C2C cases. Further, the code could be binding if the counter-party also uses a smart contract to express their assent (i.e. a machine-to-machine smart contract). The two smart contracts will find agreement only if there is a set of computer-readable terms that are acceptable to both parties. In such cases, there is a strong argument for limiting the contract's obligations to those expressed in the code.

However, even machine-to-machine smart contracts' terms may not be legally binding in all cases. Many jurisdictions limit parties' contractual freedom by determining that certain contractual terms are not enforceable, for instance in order to address power asymmetries between the contracting parties (such as between producers/retailers and consumers; landlords and tenants; or employers and employees) or because the terms are otherwise unconscionable.[131] For example, a smart contract that does not give a consumer a right of withdrawal or refund may fall foul of consumer protection law.

(ii) Performance: can smart contracts be breached?

Since smart contracts self-execute pre-determined code, it could be argued that they cannot be breached. The smart contract will always do exactly what it says in its code. The consensus mechanisms work to ensure that the relevant smart contracts are executed correctly (according to their technical specifications) in the appropriate circumstances. However, as

[127] Werbach and Cornell (2017), supra note 117, p. 30.

[128] DAO Hub Explanation of Terms and Disclaimer, https://web.archive.org/web/20160704190119/https://daohub.org/explainer.html [accessed 25 October 2017].

[129] See Pierluigi Cuccuru, "*Beyond bitcoin: an early overview on smart contracts*", (2017), International Journal of Law and Information Technology, Volume 25, Issue 3, pp. 188-189.

[130] For relevant UK case law, see *Cleaver v Schyde Investments Ltd* (2011) EWCA Civ 929 (Law Society's Standard Conditions of Sale).

[131] See Max Raskin, "*The Law and Legality of Smart Contracts*", (2017), 1 Geo.L.Tech.Rev.305, pp. 325-329 (on unconscionable smart contract terms) and Werbach and Cornell (2017), supra note 117, pp. 49-50 (on the public policy rules embedded in contract law).

noted above, the legal contract between the parties is likely to include obligations beyond the code itself, based on other communications. Not all of these obligations can be captured fully and correctly by the underlying smart contract. As a result, there may be a mismatch between what the parties have agreed and what the smart contract's code executes, resulting in non-performance.[132]

Smart contracts are by their nature limited to those contractual terms that can be specified in computer-readable code,[133] and further limited by any constraints imposed by the blockchain system in which the contract operates. As a result, they are unable to capture the real-world complexity of all but the simplest transactions.[134]

Contractual performance[135] in transactions involving digital assets, such as the exchange of crypto-assets described above, is relatively straightforward to describe and measure. The ledger then provides a reliable record of the transactions and contracts executed. However, the complexity of transactions is magnified if performance involves off-chain, real world assets (such as giving access to a vehicle through a smart lock). In such cases, performance is harder to assess: the quality of the service may be below a reasonable standard (e.g. the vehicle may not start) or otherwise differ from reasonable expectations.[136]

Where performance is disputed, the parties may seek to address this through negotiation, arbitration, or litigation. Thus, while smart contracts might simplify execution, they will not prevent contractual disputes. Nonetheless, the blockchain's consensus mechanisms may give guarantees about the smart contracts that were executed and the surrounding circumstances.

(iii) Modification and enforcement: are mistakes in smart contracts reversible?

The code of smart contracts is subject to human error and will therefore likely contain bugs.[137] Mistakes in the code may prove costly. For instance, if an attacker exploits a poorly-written smart contract, any resulting transactions are written into the tamper-evident blockchain. The offering party can replace the smart contract for future transactions, but cannot edit the existing smart contract or easily reverse its effects. For example, in 2017, a user accidentally triggered a flaw in the code of a smart contract service called Parity. The service provided multi-signature wallets for Ethereum. As a result, an estimated 1m ETH (worth c. US$300m as of November 2017) was permanently frozen in the wallets, with no way for users to access their funds.[138]

Mistakes in legal contracts may be costly too. However, since a legal contract is not self-executing, a party can withhold performance and renegotiate terms.[139] In case of an ongoing

[132] See Eliza Mik, "*Smart Contracts: Terminology, Technical Limitations and Real World Complexity*" (2017), Law, Innovation, & Technology, 9.2, p. 11.

[133] Werbach and Cornell (2017), supra note 117, p. 43.

[134] Mik (2017), supra note 132, p. 21; Levy, K., 'Book-Smart, Not Street-Smart: Blockchain-Based Smart Contracts and the Social Workings of Law' *Engaging Science, Technology, and Society* 3 (2017), 1-15.

[135] I.e. the performance of legal contract, not performance in the computer science sense.

[136] Raskin (2017), supra note 131, p. 326.

[137] Werbach and Cornell (2017), supra note 117, p. 42.

[138] The Guardian, "*'$300m in cryptocurrency' accidentally lost forever due to bug*", (2017), https://www.theguardian.com/technology/2017/nov/08/cryptocurrency-300m-dollars-stolen-bug-ether?CMP=share_btn_tw [accessed 15 November 2017].

[139] Sklaroff, supra note 89, p. 23.

dispute, the parties can turn to litigation. In that case, a judge may be able to correct obvious mistakes or incompleteness in contract language through interpretation, by assessing the intent of the parties.[140] Thus, parties can seek to correct mistakes in legal contracts after signing.

Resolving mistakes in smart contracts is more complicated. Since a smart contract is self-executing, its automated performance is written into the blockchain. As outlined in 3.3.2, reversing past blockchain transactions would require either a 'correcting transaction', or encouraging the other participants in a blockchain network to initiate a 'hard-fork'. As explained above, centralised platforms feature better reversibility than distributed platforms. For example, the only way to release the 'frozen' Parity funds is reported to be through an Ethereum hard fork.[141] While the Parity team has proposed such a 'hard fork',[142] as of December 2017, it was unclear how the rest of the Ethereum participants would respond.

It may be possible to incorporate logic for unwinding a transfer into the smart contract at the outset. For instance, enforcement of the contract could, in theory, be structured to allow for arbitration by a third party adjudicator with their own private key.[143] This would however re-introduce a requirement of trust in a third party and add a further layer of complexity to the code. It is not clear at present who these adjudicators would be or what procedural and substantive rules they would apply in resolving disputes.[144]

As noted above, with a centralised platform, it should be easier to arrange a corrective fork. The TTP or small group of trusted nodes can agree to correct the ledger. As a result, centralised platforms feature better reversibility. Companies are also exploring the possibility of 'editable' blockchains, changing the way hash pointers link blocks so that a small number of authorised parties can change past blocks.[145]

(iv) Confidentiality and trade secrets

As noted above in Section 5, the code of a smart contract is executed by all nodes on the network and is publicly visible. However, many contracts contain commercially sensitive information. Thus, smart contracts are inappropriate for contracts that contain information that would otherwise be subject to a non-disclosure agreement or confidentiality clause. In a worst case scenario, revealing information through a smart contract could lead to inadvertent loss of trade secret protection or to a breach of confidentiality.

6.2.2 Digital Autonomous Organisations

Smart contracts can be used to manage the assets and determine the structure, purposes, and functioning of an organisation. Such organizations are known as Decentralised

[140] Werbach and Cornell (2017), supra note 117, p. 47.

[141] The Guardian (2017), supra note 138.

[142] See https://paritytech.io/blog/on-classes-of-stuck-ether-and-potential-solutions-2.html [accessed 20 December 2017].

[143] Aaron Wright and Primavera De Filippi, "*Decentralized Blockchain Technology and the Rise of Lex Cryptographia*" (2015), p. 50, https://ssrn.com/abstract=2580664 [accessed 20 December 2017]; Werbach and Cornell (2017), supra note 117, p. 17.

[144] Sklaroff (2017), supra note 89, pp. 39-40.

[145] See also Accenture, "*Editing the Uneditable Blockchain: Why Distributed Ledger Technology Must Adapt to an Imperfect World*" (2017), https://newsroom.accenture.com/content/1101/files/Cross-FSBC.pdf [accessed 20 December 2017].

Autonomous Organizations (DAOs). For instance, a DAO could have a number of members who could together, say by two-thirds majority, decide how to spend funds. This replicates some of the elements of a legal company, such as dividend-receiving shareholders and tradable shares, but using only smart contracts for enforcement.[146] As a result, a DAO could allow strangers to contribute capital towards a common enterprise pseudonymously, without needing to trust a management team to exercise control over the company and its associated capital.

However, most jurisdictions feature formal requirements for creating a company, such as registration with a central registry like the UK's Companies House. As a result, a DAO might not qualify as a recognised 'legal person' or offer shareholders limited liability. Participants in a DAO may face significant legal risk, potentially being held personally responsible for the DAO's liabilities.[147]

Further, DAOs suffer from the same risk of human error as smart contracts, as illustrated by the hack of '*The DAO'* set up on the Ethereum platform. *The DAO* was set up in 2016 as a funding vehicle for Ethereum-based projects, such as using 'smart locks' to let people share their physical assets (e.g. cars, boats, apartments).[148] It attracted over US$160m in Ether funding from around 11,000 members. However, a hacker exploited a vulnerability in *The DAO*'s smart contracts and siphoned off almost a third of its funds.[149]

Since Ethereum is a widely distributed platform, the only way to 'correct' the effects of this hack (without the hacker's cooperation) was to convince nodes to replace their local copy of the blockchain with a chain in which the funds were still held by *The DAO*. In the event, this was achieved and a majority of Ethereum nodes moved to the new chain, thereby reversing the attack. (A minority of users refused to do so and stayed on the original branch, which was renamed 'Ethereum Classic'.)[150] However this was a particularly high-profile, one-off incident, affecting a large number of participants. Such contentious hard forks may be unlikely for issues affecting fewer parties or deemed by the community as insignificant compared with the cost of implementing a fork.

As noted above, such reversibility is easier on centralised platforms. A single TTP or group of trusted nodes can revert transactions and modify balances if necessary, for instance to enforce court decisions.[151] Ultimately, reversibility and trust are linked.[152] As a result, any platform that seeks to function in a completely trustless environment will face difficulties in providing reversibility, with potentially significant legal risks.

[146] Buterin (2013), supra note 20, p. 23.

[147] See Dirk Zetzsche, Ross P. Buckley, and Douglas W. Arner, "*The Distributed Liability of Distributed Ledgers: Legal Risks of Blockchain*" (2017), Law Working Paper Series number 2017-007, p. 36.

[148] See "*The DAO in-depth: Interview with Stephan Tual, Slock.it CTO*" (2016), available at http://forklog.net/the-dao-in-depth-interview-with-stephan-tual-slock-it-cto/ [accessed 20 December 2017].

[149] Voshmgir Shermin, "*Disrupting Governance with Blockchains and Smart Contracts*", (2017), Strategic Change, Volume 26(5), p. 506.

[150] Shermin (2017), supra note 149.

[151] Buterin (2015), supra note 65.

[152] Nakamoto (2009), supra note 9, p. 1.

6.2.3 Initial Coin Offerings

An initial coin offering (or ICO) is a way for a start-up with a new blockchain-based service to raise money by selling an initial set of tokens.[153] The company then uses the money raised to launch the service. Purchasers of the tokens can then 'spend' their tokens to access the service or sell them on cryptocurrency exchanges. Many ICO-tokens are built through smart contracts on top of the Ethereum blockchain. ICOs raised an estimated US$1.3bn in the period January to September 2017 (up from US$222m in all of 2016).[154]

Purchasing tokens in an ICO is a high-risk activity. Many companies offering ICOs provide investors with little more than a whitepaper and a website. If the service proves popular, the tokens' value can increase exponentially. Conversely, if the start-up fails to launch a valuable service, the tokens may become worthless, leaving investors with losses. The recent management battle at Tezos, which raised US$232m through an ICO in 2017, illustrates the risk of such early stage investment.[155]

From a regulatory perspective, the details of the offering are important, particularly the rights associated with the token. Some ICO token sales may focus on offering investors the opportunity to profit from a company's success for instance by emphasizing the opportunity to sell tokens on secondary markets.[156] Such ICOs resemble more traditional sales of securities, like Initial Public Offerings (IPOs). Unlike with an IPO, however, ICO investors typically do not gain equity in the company and may have very limited ability to influence the company's direction. In other instances, ICO token sales may focus on the use of the service (so-called 'utility tokens'). In such cases, the offer of tokens might appear closer to a pre-launch 'sales' arrangement, similar to other crowdfunding arrangements such as Kickstarter.[157] The legal status of utility tokens is a topic of ongoing debate and may differ per jurisdiction.

Nonetheless, the characterisation of tokens is important, since many jurisdictions seek to protect investors by regulating securities. A key means of providing protection is to mandate disclosure, so that investors have access to the information they need in order to assess the risk associated with an investment. This can help to address the informational asymmetry between potential investors and the investment's promoters.[158]

In line with this, regulators have begun to take action in relation to ICOs. For instance, in July and December 2017, the US Securities and Exchange Commission determined that ICO

[153] Jin Enyi and Ngoc Dang Yen Le, "*Regulating initial coin offerings ("crypto-crowdfunding")*" (2017), 8 JIBFL 495.

[154] Martin Arnold, "Tech start-ups raise $1.3bn this year from initial coin offerings" (2017) https://www.ft.com/content/1a164d6c-6b12-11e7-bfeb-33fe0c5b7eaa?mhq5j=e5 [accessed 31 October 2017].

[155] Maria Terekhova, "*What Tezos Crisis Could Mean for the ICO Space*", (2017), http://uk.businessinsider.com/what-tezos-crisis-could-mean-for-the-ico-space-2017-10?r=US&IR=T [accessed 31 October 2017].

[156] See SEC Chairman Jay Clayton, 'Statement on Cryptocurrencies and Initial Coin Offerings', 11 December 2017, https://www.sec.gov/news/public-statement/statement-clayton-2017-12-11, ('SEC Chairman Statement') p. 3 [accessed 20 December 2017].

[157] SEC Chairman Statement, supra note 156, p. 3 [citing the example of a book club participation token]. See further Enyi and Ngoc Dang Yen Le, supra note 153; https://www.coindesk.com/asic-on-blockchain-australias-securities-watchdog-unlikely-to-regulate-icos/ [accessed 20 December 2017].

[158] Werbach (2017), supra note 19, p. 29.

tokens can qualify as securities, depending on the facts of the case.[159] In the specific case of *The DAO* (set out above), the tokens were deemed securities and their unregistered sale violated securities regulations.[160] Similarly, in relation to Munchee Inc., the SEC ordered the company to cease its token sale since it violated securities law, noting (inter alia) that token purchasers had a reasonable expectation of profits from their investment in the Munchee enterprise.[161] The Monetary Authority of Singapore and the UK Financial Conduct Authority have also announced that ICOs would, under certain circumstances, be regulated as securities.[162]

Regulators will need to consider how to regulate ICOs in order to protect investors without chilling innovation.[163] Without protection, investors risk being misinformed and suffering losses, which in turn is likely to undermine confidence in the market as a whole.[164]

6.3 Intellectual Property: patents, copyright, and database rights

Various forms of intellectual property may apply to different aspects of blockchain technology. First, blockchain-related inventions that provide a novel, non-obvious technical solution that is capable of industrial application may be patentable.[165] While the core components of blockchain technology are public knowledge, companies are filing patent applications for improvements relating to security and encryption techniques.[166] A patent will grant the right holder exclusive rights to the commercial exploitation of the protected invention.[167]

Second, the structure of a specific blockchain database could be protected by copyright. Under the EU's 1996 Directive on the legal protection of databases (the 'Databases Directive'), EU member states are required to provide copyright protection for databases which, 'by reason of the selection or arrangement' of contents 'constitute the author's own intellectual creation'.[168] To attract copyright protection, the author must have expressed his creative ability in an

[159] United States Securities and Exchange Commission, "*Report of Investigation Pursuant to Section 21(a) of the Securities Exchange Act of 1934: The DAO*", Release No. 81207 (2017) ('DAO Report'), p. 10; United States Securities and Exchange Commission, "Order Instituting Cease-and-Desist Proceedings Pursuant to Section 8A of the Securities Act of 1933, Making Findings, and Imposing a Cease-and-Desist Order" File no. 3-18304 in the matter of Munchee Inc., 11 December 2017 ('Munchee Order').

[160] DAO report, supra note 159, p. 10.

[161] Munchee Order, supra note 159, pp. 5, 8-10.

[162] http://www.mas.gov.sg/News-and-Publications/Media-Releases/2017/MAS-clarifies-regulatory-position-on-the-offer-of-digital-tokens-in-Singapore.aspx; https://www.fca.org.uk/news/statements/initial-coin-offerings [both accessed 20 December 2017].

[163] For a detailed discussion, see Peter van Valkenburgh, "*Framework for Securities Regulation of Cryptocurrencies*" (2017), https://coincenter.org/entry/framework-for-securities-regulation-of-cryptocurrencies [accessed 21 November 2017].

[164] Werbach (2017), supra note 19, p. 30.

[165] Tim Press, "*Patent Protection for Computer-related Inventions*" (2012) in: "Computer Law", Chris Reed, Seventh Edition, OUP Oxford.

[166] See The Economist, "*A rush to patent the blockchain is a sign of the technology's promise*", 12 January 2017, available at https://www.economist.com/news/business/21714395-financial-firms-and-assorted-startups-are-rushing-patent-technology-underlies [accessed 20 December 2017].

[167] Press (2012), supra note 165.

[168] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the Legal Protection of Databases, Art. 3(1).

original manner by making free and creative choices in setting up the database. By contrast, the database will lack the required originality, if the setting up was dictated by technical considerations.[169] Since a blockchain's structure will ordinarily be determined by technical design decisions based on a platform's desired functionality, it seems likely that most blockchain databases will not be protected by copyright under EU law.

Third and finally, a blockchain database could be protected as a database (i.e. as a collection of data arranged in a systematic or methodical way).[170] The Databases Directive established a *sui generis* protection right for databases.[171] The *sui generis* right protects databases for which there has been (qualitatively or quantitatively) a substantial investment in the obtaining, verification, or presentation of their contents.[172] This right is not affected by making the database publicly accessible.[173]

It could be argued that the TTP or trusted nodes of a new, centralised blockchain platform make a substantial investment in setting up the blockchain database. For instance, they deploy human, financial, and technical resources in writing the blockchain software and setting up and running nodes.[174] As a result, the resulting blockchain database may be protected under the sui generis right.[175] Determining the right holders would be more difficult for distributed platforms, where developers write the initial software, but the database itself is compiled through the efforts of (large numbers of) nodes and miners. Provided they have sufficient links to an EU Member State[176], the right holders of a centralised platform could have the right to prevent extraction and re-utilization of (all or a substantial parts of) the contents of that database.[177] This may be relevant to instances of 'forking', where a third-party takes an existing blockchain database as the basis for starting a new blockchain platform (see Section 3.3 above).

6.4 Data protection law

The application of EU data protection law to blockchain-based platforms raises difficult questions. In this section, we first consider data protection law's applicability, before reviewing who would qualify as controllers or processors and considering compliance and liability.

[169] ECJ, Football Dataco et al. v. Yahoo! UK et al., Judgement, Case C-604/10, (2010), paras 37-40.

[170] Directive 96/9/EC, Art. 1(2), defines a database as 'a collection of independent works, data or other materials which are arranged in a systematic or methodical way and are individually accessible by electronic or other means'.

[171] Directive 96/9/EC, Arts. 7-11.

[172] Directive 96/9/EC, Art. 7(1).

[173] ECJ, The British Horseracing Board et al. v. William Hill, Preliminary Ruling, Case C-203/02, para. 67.

[174] See ECJ, Fixtures Marketing v. Oy Veikkaus, Preliminary Ruling, Case C-46/02, (2004), paras 37-40.

[175] See Chris Reed, "*Database Protection*", in: "*Computer Law*" (2012), Chris Reed, Seventh Edition, OUP Oxford.

[176] Directive 96/9/EC, Art. 11 requires beneficiaries to be nationals or habitual residents of an EU Member State, or to be companies and firms formed in accordance with a Member State's laws with their registered office, central administration, or principal place business within the Community.

[177] Directive 96/9/EC, Art. 7(1), (5).

6.4.1 Will data protection laws apply?

EU data protection laws apply to the processing of personal data that falls within the regime's territorial scope. The General Data Protection Regulation (GDPR)[178] will enter into force in May 2018 and have a broad territorial reach. The GDPR applies to controllers and processors 'established' in the EU.[179] Existing case law suggests that the test for establishment will be applied expansively.[180] The GDPR also applies to controllers and processors not established in the EU, where the processing relates to the offering of services to data subjects who are in the EU or monitoring of their behaviour that takes place within the EU.[181]

As a result, the activities of many blockchain operators will fall within the regime's territorial scope. For example, since anybody can use an open/permissionless platform, operators of such platforms may be deemed to offer services to data subjects in the EU. For instance, it could be argued that the nodes and miners who collectively support the Bitcoin network offer a payment service to EU data subjects. In contrast, to avoid the GDPR's applicability, non-EU-established operators of closed/permissioned platforms could attempt to prevent data subjects located in the EU from using their platforms. (Further, if the blockchain platform provides a 'publicly available electronic communications service' to end-users in the EU, then the specific regime of the e-Privacy Directive will apply to the processing of personal data, instead of the GDPR.[182])

'Processing' is broadly defined. It refers to any operation or set of operations performed on personal data.[183] As a result, blockchain users, nodes, and miners may engage in processing of personal data when sending, verifying, and storing transaction data.

The definition of 'personal data' is also very expansive. It covers any information that relates to an identifiable person, i.e. a person who can be identified "directly or indirectly".[184] To

[178] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

[179] GDPR, Art. 3(1).

[180] See ECJ, Google Spain v. AEPD and Mario Costeja González, Judgment, (2014), paras 54-59 (the concept of establishment was interpreted so as to provide a broad territorial scope in order to "prevent individuals from being deprived of the protection guaranteed by the directive and that protection from being circumvented").

[181] GDPR, Art. 3(2). This is a change from the test in the Data Protection Directive which focussed on use of 'equipment' in the EU rather than offering services to or monitoring the behaviour of data subjects in the EU. Both approaches, however, are potentially very broad in scope. For more on the 'long arm' reach of EU data protection laws, see W. Kuan Hon, Julia Hőrnle, Christopher Millard, "*Which Law(s) Apply to Personal Data in Clouds?*", in: Cloud Computing Law (2013), Oxford University Press, Oxford.

[182] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector ('e-Privacy Directive'); Council of the EU ePrivacy Regulation Proposal, 2017/0003 (COD), 8 September 2017. Under the proposed e-Privacy Regulation (which is intended to replace the existing e-Privacy Directive), the definition of such services will cover 'interpersonal communications services', including 'Over the Top' (OTT)-services such as VoIP and instant messaging. Thus, if blockchain technology were used as the basis for a new OTT service (such as a web-based email service), this would be likely to fall under the proposed e-Privacy Regulation. See for instance http://johnmcafeeswiftmail.com/ or http://www.cryptamail.com/ [both accessed 20 December 2017].

[183] GDPR, Art. 4(2).

[184] GDPR, Art. 4(1). See further W. Kuan Hon, Christopher Millard, Ian Walden, "*What is Regulated as Personal Data in Clouds?*", in: Cloud Computing Law (2013), Oxford University Press, Oxford.

determine whether a person can be indirectly identified, account should be taken of all the means likely reasonably to be used by the controller or by any other party to identify the person.[185]

A blockchain database is likely to contain at least two types of data. First, it will store metadata related to transactions, namely the sender's and recipient's addresses and a timestamp. Second, it will store data on the object of a transaction. For instance, in the case of Bitcoin, the object would be an amount of BTC. We will now consider each of these in turn.

(i) Metadata as personal data

With respect to metadata, if the platform's users are natural persons, the sender's and recipient's addresses will almost always qualify as personal data. This is most obvious where these addresses directly reveal a person's identity. For specific applications, for example a land registry blockchain, titles to property may be transferred from one named individual to another. While a platform operator may opt to 'pseudonymise' the data (by replacing a person's name with a pseudonym), the GDPR makes it clear that such pseudonymised data will still qualify as personal data.[186]

Even Bitcoin's disposable public keys and addresses may qualify as personal data. The ECJ has determined that information can constitute personal data even where only a third party has the additional data necessary to identify the person.[187] In such cases, whether information is personal data may depend on if the possibility of combining the two sources is "a means likely reasonably to be used" to identify the data subject.[188]

Although there is no central register matching addresses to individuals, a Bitcoin address might still be linked to a real-world identity by combining it with other data. Intermediaries, such as wallet services or exchanges, may register users' real-world identities, for instance to comply with regulatory requirements. Further, the counter-parties a user transacts with, such as merchants, may register users' real-world identities for their own commercial purposes. Combining intermediary records with the public blockchain would reveal the real-world identity behind a Bitcoin address.

In addition, as noted in Section 4.2.1 above, it may be possible to determine Bitcoin user identities by other methods, such as linking public keys to IP addresses. Finally, metadata may reveal a pattern of transactions with publicly known addresses (such as merchants) that could be used to single out an individual user, such as through a technique known as 'transaction graph analysis'.[189] For example, if a certain restaurant accepts Bitcoin as payment and its address is publicly known, then payments to that address would suggest that the sender visited that restaurant at a certain time.[190] Given this, Bitcoin addresses and public keys might in certain circumstances qualify as personal data.

[185] GDPR Recital 26.

[186] GDPR Recital 26.

[187] ECJ, Breyer v. Germany, Judgment, Case C-582/14, (2016), paras 31, 39.

[188] ECJ, Breyer v. Germany, Judgment, Case C-582/14, (2016), para. 45.

[189] See Adam Ludwin, *'How Anonymous is Bitcoin'*, (2015) available at https://coincenter.org/entry/how-anonymous-is-bitcoin [accessed 23 November 2017].

[190] Ludwin (2015), supra note 189.

Conversely, if a platform's users are all legal persons (such as businesses) the platform could be designed such that the metadata does not contain information related to natural persons. For example, a group of banks could set up a closed/permissioned platform to settle end-of-day inter-bank payments for their own accounts (i.e. reflecting the sum total of individual transactions at the inter-bank level). In this case, the addresses might refer only to the sending and receiving banks in question and need not relate to any identifiable person.

(ii) Object of transactions data as personal data

With respect to the object of transactions, in many use cases this information will not relate to an identifiable person. For instance, in Bitcoin, the amount of BTC transferred does not necessarily relate to an identifiable person,[191] nor would payment data for an overall end-of-day settlement between banks.

Nonetheless, in other use cases, the object of the transaction could be linked to real-world identities. For example, a group of retail banks could set up a blockchain platform to share KYC information on their customers with each other. In that case, the object of transactions would be information about natural persons and could be written into the blockchain.

However, it may be possible to design a blockchain platform such that any personal data is not stored 'on the chain', but is stored in encrypted form in a separate, 'off-chain' database. The blockchain transaction data would then only contain the information needed to access and decrypt the personal data in the separate database.[192] In this manner, it may be possible to use off-chain storage to prevent personal data being stored on the blockchain.

6.4.2 Who will be subject to legal obligations as controller(s) and processor(s)?

Data controllers and processors are responsible for ensuring compliance with data protection law. The controller is the natural or legal person who determines the purposes and means of processing personal data.[193] The ECJ has held that 'controller' should be interpreted broadly, so as to ensure the effective and complete protection of data subjects.[194] A processor is any natural or legal person who processes personal data on behalf of the controller.[195]

As noted above, the personal data stored on a blockchain may consist of metadata and data on the object of each transaction. The 'purposes and means' of processing this personal data can be analysed from two perspectives. On the one hand, at the macro-level, looking at the blockchain infrastructure as a whole, the purpose of processing personal data is to provide the associated service. For instance, with regard to Bitcoin, the purpose of processing transaction data is to provide a peer-to-peer system of electronic cash. At this level, the 'means' of processing will generally consist of (i) the software that nodes and miners run to find new blocks and store and update the blockchain database, and (ii) the hardware that nodes and miners use for this purpose.

[191] In some cases, the amount paid could be combined with information about the recipient address to specify the product or service paid for, which may help 'single out' the sender.

[192] For a technical description of such a system, see Guy Zyskind, Oz Nathan, Alex Pentland, "*Decentralizing Privacy: Using Blockchain to Protect Personal Data*", (2015) IEEE CS Security and Privacy Workshops.

[193] GDPR, Art. 4(7).

[194] ECJ, Google Spain v. AEPD and Mario Costeja González, Judgment, (2014), paras 34-38.

[195] GDPR, Art. 4(8).

From this macro-level perspective, nodes and miners arguably decide to engage in processing for their own ‘purposes’, namely to facilitate the platform. They would also determine the ‘means’ of processing, by deciding which software and hardware to use.

On the other hand, at the micro-level, looking at individual transactions, the ‘purpose’ of the processing is to record a specific transaction onto a blockchain. At this level, the ‘means’ would refer to the choice of blockchain platform.

From this micro-level perspective, users enter personal data into the system when submitting their transactions. Thus, for each specific item of personal data, the individual user arguably determines the ‘purposes’ of processing, namely: to record a specific transaction onto the blockchain. The user also determines the ‘means’, namely: to use that blockchain platform to execute their transaction. From this perspective, nodes and miners simply facilitate access to a blockchain database, while the users determine which data are stored there.

In many cases, the analysis of who ‘determines the purposes and means of processing’ will depend on whether you adopt the macro- or micro-level perspective. Since data protection law is concerned with the processing of specific items of personal data, we consider the micro-level perspective more appropriate. To illustrate this point, we consider three different use cases below.

(i) Centralised platform: land registry

With a centralised platform, the platform operator will be likely to determine the ‘means’ of processing at the macro-level. For instance, with a land registry, the government agency setting up the platform could either develop the underlying software in-house, or buy in software from a third-party developer. The agency could then run a single node and miner as a TTP on its own hardware. Similarly, at the macro-level, the agency would decide to process personal data on a blockchain platform for the purpose of providing a registry of titles to land. Seen from the macro-level, one could argue that the land registry should be considered a controller.

However, the micro-level perspective instead focuses on individual transactions. The personal data processed on a land registry blockchain would be metadata: in this case, the sender and recipient’s identifiers. Users enter this personal data onto the platform for their own purposes, namely to register and/or transfer titles to land. Users also arguably determine the means of processing by choosing the blockchain-based land registry as the medium to execute their transfers.[196] Once users have submitted transactions, the government agency may merely process associated data on the users’ behalf. Seen from this micro-level perspective, the user should arguably be considered the data controller, with the agency acting only as a processor on their instructions.

In practice, the government agency need not perform the processing itself, but may engage a sub-processor. For instance, it might use a Blockchain-as-a-Service (BaaS) offering, whereby a third party provides the underlying supporting infrastructure. The agency could pay a third party to run the miner and node on the third-party’s own hardware.[197] In such cases, the BaaS-provider could qualify as a sub-processor when processing personal data for the land registry.

It is important to note, however, that there may be multiple data controllers in relation to a particular set (or subset) of personal data. For example, even if we categorise the government

[196] This assumes that users have a choice between various ‘means’ of executing their transfer.

[197] See Singh and Michels (2017), supra note 41.

agency running the blockchain land registry platform as a 'mere processor' for the purpose of individual transactions, it may nevertheless be a controller for other purposes such as assessment and collection of taxes relating to land transactions.

(ii) Narrowly distributed platform: inter-bank customer data sharing

As a second example, assume a group of parties decide to set up a narrowly distributed blockchain platform with a small number of trusted nodes. For instance, a group of retail banks may set up a blockchain to share information on their customers for KYC purposes. The platform is closed (only the founding banks, or others they authorise, can use it) and private (only the participating banks can view the blockchain database).

At the macro-level, the parties who set up the platform would determine the means of processing by designing the platform. Thus, in specifying the software, the banks would determine which data to store on the blockchain and how it is processed through the consensus protocol. They could also dedicate hardware to running the nodes. The banks also arguably determine the purpose of processing, namely to share KYC information. Seen from this perspective, the banks could be argued to be controllers when setting up the platform and when acting as nodes and miners.

However, as above, the micro-level perspective would instead focus on individual transactions. In this case, the personal data would be the object of the transactions (namely: the customer records). Banks enter this information onto the platform when submitting transactions. In line with the above analysis, as users, each bank would arguably act as a controller with regard to the customer data it submits to the platform. Further, when processing the data as nodes and miners, the banks might be acting only as processors with regard to the customer data that other participating banks have submitted.

The above applies so long as the group of banks acting as nodes/miners only process the transaction data for the purposes determined by the sending user (namely to execute the transaction). If they engage in further processing of the data for their own purposes, they would likely become controllers of that data.[198] For example, a bank could analyse the customer data stored on its copy of the blockchain to glean commercial insights.

(iii) Widely distributed platform: cryptocurrency

For widely distributed, open/permissionless platforms, such as Bitcoin, determining controllers and processors is difficult. The definition used by data protection law is arguably ill-suited to distributed platforms, which purposefully lack a central administrator who could bear responsibility for compliance. Instead, control is deliberately distributed.

At the macro-level, the platform's purpose is to facilitate a peer-to-peer system of electronic cash. The means consist of the Bitcoin core software and the hardware provided by nodes and miners. It is generally accepted that these purposes and means were originally envisaged by a person (or persons) known as 'Satoshi Nakamoto'. Today, the Bitcoin core code developers control the core software (see Section 3.3 above). This arguably gives them a high degree of factual control over the 'why' and 'how' of processing Bitcoin transaction data.[199] However, the developers do not process any personal data themselves (unless they also

[198] See Article 29 Working Party Opinion 1/2010 on Cloud Computing 01037/12/EN WP 196 (2012), p. 8.

[199] See Article 29 Working Party Opinion 1/2010 on the concepts of "controller" and "processor", 00264/10/EN WP169, (2010), pp. 11, 14.

happen to run nodes or mine new blocks). They merely make the software available for others to use. As a result, they are unlikely to qualify as either controllers or processors under data protection law.

Nodes and miners process personal data in the form of Bitcoin addresses, for instance when they store and broadcast transaction data. They decide to process such data for the purposes of facilitating the cryptocurrency and, in the case of miners, to reap a reward for mining new blocks. They provide means in the form of hardware and do so on their own behalf (rather than on instructions from any other party). Given this, the miners and nodes arguably could be controllers.

However, they have limited factual influence over the 'why' and 'how' of processing. Nodes and miners cannot easily change the 'core' software and its consensus protocol. (At most, they can propose changes or move to a different fork.) Instead, they download the software and run it passively on their computers. Given this distribution of control, it is not straightforward to determine controllers at the macro-level.

As above, the micro-level perspective focuses instead on individual transactions. The personal data in this case are the sender and recipient's Bitcoin addresses and (potentially) the related transaction data (e.g. timestamp and amount of BTC). Thus, with respect to each item of personal data, the sending user decides to submit it to the Bitcoin platform for their own purpose, namely to transfer a certain quantity of value to the recipient. In addition, they arguably determine the means of processing, by deciding to use Bitcoin for their transaction. Seen from this micro-level perspective, the users should be considered data controllers. Nodes and miners may simply process this data on behalf of each user.

One could argue that the users have even less factual influence over the means of processing, since they cannot change the Bitcoin software that nodes and miners run. However, this imbalance in power over the means of processing applies to other cases involving individual users and large service providers. For instance, with cloud services, customer can often only use a commoditised cloud service without modification and may have no choice but to accept the standard contractual terms offered by a large cloud service provider if they wish to use a particular service. Nonetheless, as controllers, the customers remain responsible for their decision to use a certain service. They should choose a cloud provider that guarantees compliance with the relevant requirements of data protection law.[200] Similarly, cryptocurrency users should arguably choose a platform that is compliant with data protection law.

Nonetheless, some users may benefit from an exemption for personal activities. Data protection law does not apply to natural persons in the course of a purely personal or household activity.[201] Thus, if a group of friends use a cryptocurrency like Bitcoin to make payments to each other, they may be exempt from data protection law in relation to such processing. However, if a user makes payments outside of a personal or household activity, such as for commercial, political, or charitable goals, the exemption may not apply.[202] Further, the exemption would be unlikely to apply to legal persons who make Bitcoin payments. In such circumstances, users will typically subject to the full responsibilities of a data controller, as set out in Section 6.4.3 below.

[200] Article 29 Working Party Opinion 1/2010 on Cloud Computing 01037/12/EN WP 196 (2012), p. 8.

[201] GDPR, Art. 2(2)(c).

[202] See Article 29 Working Party Opinion 1/2010 on Cloud Computing 01037/12/EN WP 196 (2012), p. 6.

Finally, anybody who accesses the data stored on a public blockchain and processes it for their own purposes becomes a data controller. Thus, if a node analyses the payments data in its local copy of the blockchain to glean commercial insights, it would become a controller in respect of that data.

(iii) Conclusions regarding data controllers and processors

In each of the above examples, the analysis of controllers and processors depends on whether a blockchain use case is analysed from the macro- or micro-level perspective. Given that data protection law is concerned with the processing of specific items of personal data, the micro-level perspective is arguably a more appropriate starting place. Following this line of reasoning, users would be considered data controllers in respect of the personal data they submit to the blockchain platform, since they determine both purposes (to execute the transaction) and means (in choosing the platform). They delegate decisions on the technical and organisational details of the processing to the collective of developers, nodes, and miners.

Whether nodes and miners should also be deemed controllers would depend on the facts of each case. If nodes and miners merely process transaction data on behalf of users, they could arguably qualify as processors, rather than controllers. In some cases, they may simply facilitate the processing of transactions on behalf of users, by passively running the relevant software. In this respect, they could be compared to providers of cloud computing services. Cloud providers offer Internet-based, flexible, location-independent access to computing resources, including processing capability and storage. In many cases, the cloud customer acts as the data controller, with the cloud provider merely processing the data on their behalf.[203] Similarly, a blockchain-based platform provides access to a distributed application for processing and storing transaction records.[204] Just like cloud providers, nodes and miners who provide users with access to hardware and applications are likely to be processors with regard to the personal data submitted by users.[205]

If, however, nodes and miners take a more active role with regard to the transaction data, they may also be deemed to be controllers. In that case, nodes and miners could be compared to SWIFT, a financial messaging service that facilitates international money transfers for financial institutions. In doing so, SWIFT processes personal data such as the names of the payer and payee. SWIFT initially presented itself as a mere processor, relaying messages on behalf of the financial institutions. However, the Article 29 Working Party determined that SWIFT should be considered a controller, since it acted with a significant level of autonomy in respect of the personal data it processed, including by developing, marketing, and changing the services it offered, deciding to establish a data centre in the US and to disclose data to the US Treasury.[206] Thus, the more autonomy and 'effective margin of manoeuvre' the nodes and miners have in respect of the personal data they process, the more likely they are to be considered controllers.

[203] Article 29 Working Party Opinion 1/2010 on Cloud Computing 01037/12/EN WP 196 (2012), pp. 7-8.

[204] See Singh and Michels (2017), supra note 41.

[205] For a further analysis of the role of cloud computing providers under the current Data Protection Directive, see W. Kuan Hon, Christopher Millard and Ian Walden, "*Who is Responsible for Personal Data in Clouds?*" in: Cloud Computing Law (2013), Oxford University Press, Oxford.

[206] Article 29 Working Party Opinion 10/2006 on the processing of personal data by the Society for Worldwide Interbank Financial Telecommunication (SWIFT) 01935/06/EN WP128 (2006), pp. 10-11.

6.4.3 Can blockchain controllers and processors comply with data protection law?

The uncertainty around the status and roles of controllers and processors complicates many aspects of data protection compliance and liability, including in particular the following:

(i) Lawful grounds for processing

First, controllers need a lawful ground for processing personal data.[207] Processors do not need to establish independently that the controller has valid grounds for processing.[208]

We argue above that users of blockchain platforms are likely to be deemed controllers. Consequently, each user will need to be able to demonstrate one or more lawful grounds for processing the personal data they submit to the blockchain. For instance, in the above banking example, each bank will need to determine whether its existing grounds for processing customer data extend to processing by means of a private, permissioned blockchain. Further, users of distributed, public platforms such as Bitcoin need to have a legal basis for processing the recipient's address (unless they are covered by the household exemption). It could be argued that, in signing up for a Bitcoin address, a recipient has implicitly consented to the processing of that address for transaction purposes.

Conversely, we argue that miners and nodes may be deemed 'mere processors', provided they don't also process personal data for their own purposes (for which they would then need a justification). As a result, they would not normally need to verify whether the controller has valid grounds for processing.

(i) Controller and processor obligations

Second, joint controllers must determine their respective responsibilities for compliance by means of an arrangement between them.[209] Further, controllers must put in place a contract with processors to determine how personal data will be processed, including the subject-matter, duration, nature, and purpose of the processing.[210]

Thus, for each platform, the (joint) controller(s) and processor(s) must establish contractually their data protection responsibilities amongst themselves. Achieving this should generally be easier with centralised platforms, since fewer parties are involved. For instance, in the above banking example, the banks would need to establish their respective responsibilities by means of a contract. Similarly, the land registry would need to establish its legal responsibilities as a processor via contracts with the users of its platform. It would probably seek to do so by requiring users to agree to terms of service before accessing the platform.

For widely distributed, open/permissionless platforms, it is unclear how controllers and processors might comply with these obligations. In theory, large numbers of users, nodes, and miners would need to enter into detailed contracts in order to establish their responsibilities. In practice, the most feasible way to achieve this may be through standard-form terms and

[207] Possible grounds include, *inter alia,* (i) data subject consent; or that processing is necessary for compliance with (ii) a contract to which the data subject is party, (iii) the controller's legal obligation, or (iv) for legitimate interests the controller is pursuing, provided that the controller's interests are not overridden by the interests, fundamental rights and freedoms of the data subject. See GDPR Art. 6.

[208] That said, if, in their opinion, the instructions they receive from controllers infringe the GDPR, processors may be under an obligation to inform the controller accordingly. See GDPR, Art. 28 (3).

[209] GDPR, Art. 26.

[210] GDPR, Art. 28.

conditions to be agreed whenever a user, node, or miner first uses a platform. These terms would set out the parties' legal obligations. Software developers could require nodes and miners to agree to contractual terms when downloading or updating the relevant software. For example, the Ethereum website contains standard-form terms and conditions governing the use of its software platform (the Bitcoin Core website does not).[211] That said, Ethereum's terms do not cover data protection compliance. Since users need not interact with the software directly, user-facing intermediaries (such as wallets and exchanges) would need to require users to agree to a platform's terms and conditions during user sign-up.

In addition, the GDPR imposes conditions on transfers of personal data from the EU to 'third countries'.[212] Thus, controllers must ensure they have an appropriate legal basis for any international data transfer to third countries. Returning to our earlier examples, in order to run a node in a third country, the land registry and banks would need to ensure an adequate level of protection or appropriate safeguards (for instance by locating the node in a country subject to a Commission adequacy decision or by putting in place adequate safeguards such as via approved standard contract clauses).

However, widely distributed, permissionless platforms are by design unconstrained by international borders: typically anybody, anywhere, can download the entire transaction archive and start processing new transactions as a node or miner. As a result, use of these platforms is likely to entail data transfers to third countries. Since any party in any third country can download the archive, adequacy decisions and appropriate safeguards (including binding corporate rules) are unlikely to provide sufficient coverage. Since implicit consent does not suffice for international transfer, the platform's terms of use would need to provide for explicit user consent.[213]

(ii) Data subject rights

Third, data subjects have rights in respect of their personal data. These include a right to rectification of inaccurate personal data and to data erasure (also known as the 'right to be forgotten').[214] At first glance, these rights appear to run counter to blockchain technology's 'immutability'. However, the ability to comply with such requests differs, depending on the design of the blockchain platform.

Centralised platforms feature better 'reversibility' and can limit visibility of a record to certain parties (see Sections 3.3.2 and 4.1 above). As a result, they would be in a better position to comply with data subjects' requests to rectify or erase data in past blocks. For example, if a user requested the banks to rectify a specific piece of information in their customer record on the blockchain, each bank could comply by altering the relevant transaction record and re-

[211] Cf. https://bitcoincore.org/en/about/; https://www.ethereum.org/agreement [accessed 20 December 2017].

[212] GDPR, Chapter V, Transfers of personal data to third countries or international organisations. For an analysis of the broadly similar rules under the EU Data Protection Directive, see W Kuan Hon and Christopher Millard, "*How Do Restrictions on International Data Transfers Work in Clouds?*" in: Cloud Computing Law (2013), Oxford University Press, Oxford.

[213] GDPR, Art 49(1)(a). Consent is one of various 'derogations' from the data transfer restrictions. The standard for an effective 'consent' is high and an individual must have been informed of the possible risks of the transfer.

[214] GDPR, Art. 16, 17.

hashing the subsequent blocks in their copy of the ledger. Operators of centralised platforms should similarly be able to comply with requests for erasure.

For widely distributed platforms, it is unclear how individual participants at the user, node or miner level would comply with such requests. (A node can only alter its own local copy of the ledger.) Thus, even if all users, nodes, and miners were considered controllers, this would not necessarily provide effective protection for data subjects.[215] (In theory, all nodes could agree by contract to 'fork' to a new version of the blockchain periodically, to reflect requests for rectification or erasure. However, in practice, this level of coordination may be difficult to achieve among widely distributed nodes.)

Beyond altering the chain, there may also be other technical approaches to assist data protection compliance. For instance, in many blockchain applications it is not the data that is stored on-chain, but rather links to data residing externally. Therefore, implementing mechanisms that allow deletion of data (despite a link persisting on a block) may be enough to satisfy a request for erasure; so too might deleting all instances of a private key for encrypted data (be it stored on- or off-chain). Technical approaches targeting data protection issues are an active area of research and they are likely to receive an additional impetus due to legal obligations to demonstrate '*data protection by design'*.[216]

(iii) Liability

This preliminary analysis illustrates the significant uncertainty as to how EU data protection law might apply to blockchain applications and in particular to widely distributed blockchain platforms. Given the severity of available penalties under the GDPR, there is a risk that this legal uncertainty will have a chilling effect on innovation, at least in the EU and potentially more broadly. For example, if all nodes and miners of a platform were to be deemed joint controllers, they would have joint and several liability, with potential penalties under the GDPR of up to EUR 20m or 4% of global turnover/revenues (whichever is higher).

As a result, it might be helpful if the Article 29 Working Party (or its successor body under GDPR, the European Data Protection Board) were to issue guidance regarding the application of data protection law to various common blockchain models. At a national level, as part of its Information Rights Strategic Plan 2017-2021, the UK Information Commissioner's Office has launched a programme to fund research into the privacy implications of various new technologies, including blockchain.[217]

6.5 Competition law

While using a centralised platform offers better reversibility (see above), it could raise potential legal issues under EU competition law. For instance, depending on the visibility of the record, competitors' ability to see each other's activities could be used to facilitate tacit collusion. Another competition issue could arise if platform operators limit who can participate. For instance, such issues may arise if a group of competitors (such as banks) agree to use a closed, permissioned platform amongst themselves (for instance, to settle payments). It is

[215] Matthias Berberich and Malgorzata Steiner, "*Blockchain Technology and the GDPR - How to Reconcile Privacy and Distributed Ledgers*" (2016), 2 Eur. Data Prot. L. Rev., pp. 424, 426.

[216] GDPR, Art 25.

[217] See https://ico.org.uk/about-the-ico/what-we-do/grants-programme/ [accessed 20 December 2017].

possible that, by lowering transaction costs, such a platform may provide participants with a significant competitive advantage.

These early participants may seek to refuse entry to new members or make it unduly difficult or costly for them to join.[218] If access to the platform becomes essential to providing a service, this may make it difficult or impossible for effective new competitors to emerge, with negative consequences for the cost and quality of services. As a result, participants may be engaging in a horizontal agreement between competitors with the effect of preventing, restricting, or distorting competition within the internal market.[219] If the platform were to be designated an 'essential facility', the participants may be forced to extend access to other competitors.[220] However, this issue is largely speculative at present, since we are at the early stages of commercial blockchain adoption.

7. Conclusion

7.1 Core components of blockchain technology

Blockchain technology utilises two core components to create a persistent, tamper-evident record of transactions between parties whose identity has been authenticated. First, hash pointers link blocks of transactions together, such that tampering with transaction data in past blocks breaks the links between the blocks. Second, Public Key Infrastructure is used to establish parties' identities, with private keys used to encrypt data and provide digital signatures.

7.2 Distributed vs. centralised applications

Blockchain technology can be deployed in various ways to create platforms with different features, including with regard to:

(i) who can propose new transactions to be added to the ledger;
(ii) who stores a copy of the ledger;
(iii) who can add new blocks to the ledger;
(iv) who can view the ledger;
(v) whether users are identifiable; and,
(vi) who controls the platform's underlying software.

Early cryptocurrencies were set up with the following features: a widely distributed platform (anybody can store a copy of the ledger as a node, and contribute to the process of adding new blocks as a miner); open/permissionless use (anybody can join as a user); and, a public database (anybody can view the transaction records stored on the blockchain or store a copy of the ledger).

In order to ensure the consistency of the many distributed copies and to ward off attackers, these systems rely on resource-intensive consensus protocols. They offer strong data integrity, since it is much harder to tamper with large numbers of distributed copies, as well

[218] See ESMA, "*Report: The Distributed Ledger Technology Applied to Securities Markets*", (2017), p. 11 (arguing this may raise "fair competition issues").

[219] See TFEU Art. 101(1).

[220] See ECJ, Bronner v. Mediaprint, Judgment, Case C-7/97, (1998), para. 41; ECJ IMS Health v. NDC Health, Judgment, Case 418/01 (2004), paras 28, 52.

as high resilience. However, the need to support thousands of small nodes and run proof of work limits transaction throughput and increases costs.

These systems also provide high transparency, since anyone can view all transaction data. To afford users some level of privacy, these systems are pseudonymous: users are identified only by their public key and an address which cannot easily be linked back to a real-world identity.

By contrast, future applications of blockchain are likely to feature a more centralised platform, closed/permissioned use, and a private database. In such arrangements, a TTP or group of trusted nodes store copies of the ledger, contribute new blocks, and determine which users can access the platform. This requires users to have a level of trust in the centralised blockchain administrator(s) to maintain an accurate ledger and keep the ledger secure and the system running. Since there are only a limited number of trusted nodes, such systems do not need resource-intensive consensus protocols. Further, by working with a small number of high-capacity nodes, the systems may be better able to scale and process large numbers of transactions.

In sum, from a technical perspective, whether to use a widely distributed or more centralised approach will depend on the degree of trust needed and the application's requirements for data integrity, resilience, scalability, and confidentiality.[221] Finally, it is not clear that using a centralised blockchain offers significant advantages over existing database solutions for all applications.[222] While a blockchain data structure offers strong data integrity, traditional databases can perform many of the same functions.[223]

7.3 Legal implications

Given the diversity of possible blockchain platform designs, no 'one-size-fits all' legal analysis is possible. Instead, each application of blockchain technology will need to be considered on its facts. From a legal perspective, centralised platforms are generally likely to entail lower risks in the areas we have reviewed. The TTP or group of trusted nodes can be targeted with regulation, and may be able to coordinate compliance, limit visibility of records, and reverse past transactions if necessary. Achieving these goals is harder on distributed platforms that deliberately lack a central administrator with control over the ledger.

For example, in relation to regulatory compliance, legislators and regulators can target operators of centralised platforms with know-your-customer and other obligations. This may facilitate prosecutions and other enforcement actions in cases where blockchain-based applications are used for money laundering or other unlawful purposes. With widely distributed platforms, regulation can target intermediaries (such as wallet services and exchanges) and law enforcement can track suspicious transactions.

Users have built quasi-legal constructs on blockchain platforms, such as smart contracts, DAOs, and ICOs. These constructs may give rise to significant legal risks. Communications

[221] See further Gideon Greenspan, "*Blockchains vs centralized databases: Four key differences between blockchains and regular databases"* (2016) https://www.multichain.com/blog/2016/03/blockchains-vs-centralized-databases/ [accessed 17 October 2017].

[222] See further Singh and Michels (2017), supra note 41.

[223] See Arvind Narayanan, "'*Private blockchain' is just a confusing name for a shared database*", http://freedom-to-tinker.com/2015/09/18/private-blockchain-is-just-a-confusing-name-for-a-shared-database/ [accessed 19 October 2017]. For a contrary view, see Greenspan (2016), supra note 221.

relating to smart contracts may also qualify as enforceable contracts, giving rise to legal obligations that go beyond the underlying computer code. Conversely, DAOs will not qualify as legal persons in all jurisdictions and, as a result, participants will not necessarily benefit from limited liability. Finally, some ICOs may be deemed securities, in which case their promoters will need to comply with securities laws, including obligations to provide investors with appropriate information.

Reliance on code written by fallible humans in combination with a persistent, tamper-evident data structure presents risks. For example, participants in a DAO that makes offers through a smart contract may enter into contractual agreements with counterparties resulting in financial obligations. If a bug in the smart code leads to non-performance, or incorrect performance, each participant could conceivably be liable for damages. With a centralised blockchain platform, the TTP or group of trusted nodes may be able to correct past mistakes by amending transactions in the ledger copies they control. However, on a widely distributed platform, correcting past mistakes would require cooperation from potentially thousands of nodes in a 'hard fork'.

Finally, in relation to EU data protection law, the users, nodes, and miners of blockchain platforms may be either data controllers, processors, or potentially both, if the platforms are used to process personal data. If so, they will need to comply with data protection obligations and may be exposed to substantial penalties for breaches of data protection laws. With widely distributed platforms, even if all parties involved were deemed joint controllers, it is not clear how they would comply with their obligations (such as to establish responsibilities by contract and respond appropriately to the exercise of data subjects' rights).

The legal uncertainties and risks associated with the use of distributed platforms may limit their adoption. In some cases, technical solutions may be available, or be developed, to provide greater certainty and reduce such risks. In other cases, legislators, regulators, and legal advisers are likely to face significant challenges in designing legal solutions that protect important public interests without unduly stifling innovation.

Appendix 1 – Key technical terms

Term	Meaning
Asymmetric encryption scheme	In PKI technology, data encrypted with the public key of a key-pair can only be decrypted using the private key and vice versa.
Asynchronous consensus protocol	Asynchronous consensus protocols do not require system-wide agreement on the next block to be included in the blockchain before each node begins work on subsequent blocks.
Bitcoin Cash fork	In August 2017 Bitcoin Cash was forked competitively from Bitcoin, aiming to achieve a larger block size and higher transaction rate.
Bitcoin archival nodes	A subset of full nodes that store the entire ledger and accept requests for and return old blocks.
Bitcoin full nodes	Full nodes maintain a complete copy of the blockchain database and can verify any transaction without any external lookup.
Bitcoin miners	Miners collect new transaction records into a block, and create a block header (metadata) including the previous block hash and a timestamp, and commence proof-of work on the block as part of creating this block's hash. Miners then broadcast blocks with completed proof-of-work to other nodes.
Bitcoin Simplified Payment Verification	SPV nodes validate blocks, carrying out checks including that the same coin has not been spent twice (double spending).
Blocks	For scalable management of a large number of transactions, individual transaction records are grouped, in timestamp order, into blocks. The blocks are chained using block hashes, creating a 'blockchain'.
Blockchain	A specific type of database that uses certain cryptographic functions (namely, hash pointers and PKI) to achieve the requirements of data integrity and identity authentication
Certification Authority	An example of a Trusted Third-party that offers a service to associate public keys with identities in an identity certificate.
Challenge response interaction	In PKI technology, a challenge response interaction can be used to determine whether someone holds the private key associated with a public key.
Consensus Protocol	A protocol executed by defined participants in a DLT system to reach agreement on the content of blocks to be added to the publicly agreed ledger.
Distributed Ledger	A database of transactions that is shared across a peer-to-peer network of nodes and uses a consensus protocol to agree database contents.
Forks	The Bitcoin consensus protocol is asynchronous, so miners take the first valid block they receive and proceed to work on creating the next block. Short, incorrect chains (forks) can therefore arise before the correct blocks are broadcast throughout the whole system. This is not to be confused with the forking of competing systems, see Bitcoin Cash fork above.
Hashing	Hashing involves putting the contents of some data item, e.g. an electronic document, through a hash function (such as SHA 256). This function creates a string of digits of a fixed length (the 'check digits') which are unique to and associated with the document.
Identity certificate	Created, signed and issued by Certification Authorities.
Nonce	An arbitrary number included as part of some algorithm or protocol, to be manipulated and not to represent any application semantic.

Term	Meaning
Proof of Work	Proof-of-work is part of some consensus protocols, including Bitcoin's, to ensure that miners must invest resources (CPU power and electricity) in order to participate in the protocol. PoW is hard to compute but easy to check.
Pseudonymity (digital identity)	Given a key-pair generation algorithm and peer-to-peer distribution of public keys, the private key can function as a pseudonym for a real-world entity, without revealing that entity's identity. All that is known about a participant is that it holds a private key, and is assumed to be the unique holder of that key.
Public Key Infrastructure	A PKI supports the distribution and identification of public encryption keys, enabling users and computers to securely exchange data over networks and verify the identity of communicating parties.
Signature or signed	Transaction records are signed to indicate the party or parties with whom the transaction is associated. To sign some data (often a message to be sent), the private key of the owner is used to encrypt the data.
Smart Contract	Software to create a transaction record as a ledger entry designed to capture the obligations of the transacting parties.
Synchronous consensus protocol	In Computer Science, traditional distributed consensus protocols are synchronous. The number of participants that have to reach consensus is known and they agree on a total ordering of the accepted blocks (in this case), adding the agreed blocks to the DL one at a time.
Tamper-evident	Tamper-evident means that a data structure cannot be changed without this being evident.
Trusted Third Parties (TTPs)	A Trusted Third Party is any part that acts as an intermediary to fulfil a certain function in a system, requiring participants to trust this party to carry out its function