BLOCKCHAIN AND CYBER SECURITY

a thematic report prepared by

THE EUROPEAN UNION BLOCKCHAIN OBSERVATORY & FORUM

EUBlockchain
Observatory and Forum

An initiative of the

European Commission

About this report

The European Union Blockchain Observatory & Forum has set as one of its objectives the analysis of and reporting on a wide range of important blockchain themes, driven by the priorities of the European Commission and based on input from its Working Groups and other stakeholders. As part of this it will publish a series of thematic reports on selected blockchain-related topics. The objective of these thematic reports is to provide a concise, easily readable overview and exploration of each theme suitable for the general public. The input of a number of different stakeholders and sources is considered for each report. For this paper, these include:

- Members of the Observatory & Forum's Working Groups as well as the Obeservatory's Convergence Sub-Working Group (please see next page).
- "Blockchain and cybersecurity: a taxonomic approach", by Stefano De Angelis, Gilberto Zanfino, Leonardo Aniello, Federico Lombardi, Vladimiro Sassone of the University of Southampton, an academic partner of the EU Blockchain Observatory & Forum
- Input from participants at the "Cyber Security" workshop held in Brussels on 29 October, 2019.
- Input from the Secretariat of the EU Blockchain Observatory & Forum (which includes members of the DG CONNECT of the European Commission and members of ConsenSys).

CREDITS

This report has been produced by ConsenSys AG on behalf of the European Union Blockchain Observatory & Forum.

Written by: Tom Lyons, Ludovic Courcelas
Thematic Report Series Editor: Tom Lyons
Report design: Benjamin Calméjane

v1.0 - Published on 22 May, 2020.

DISCLAIMER

The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the European Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.

ACKNOWLEDGEMENTS

The EU Observatory & Forum would like to expressly acknowledge the following for their direct contributions and feedback to this paper as members of the Cyber Security Sub-Working Group:

- Stefan Beyer
- Janis Graubins
- Konstantinos Votis

NOTE

While we have done our best to incorporate the comments and suggestions of our contributors where appropriate and feasible, all mistakes and omissions are the sole responsibility of the authors of this paper.

Contents

Executive summary

In this paper we examine the issue of cyber security as it pertains to blockchain through a number of different perspectives. Considering that many blockchain use cases involve transactions and custody of value, whether in the form of digital assets or high-value data, this is certainly one of the most important themes in the blockchain space.

We start with the question: are blockchain protocols secure? The short answer is: yes. As we show, users can have high confidence in both the distributed ledger in which blockchain data is saved, and the various consensus mechanisms used to validate transactions and agree on their order. This does not mean that there are no vulnerabilities. Data on a blockchain ledger, secured against tampering by a Merkle tree of hashes (based on well-known cryptographic principles), could be vulnerable if the cryptography currently in use is broken. Yet this is a problem facing all encrypted data and communications today, not just those on a blockchain. The consensus mechanisms currently in use all have different vulnerabilities as well, depending on how the mechanism is designed and the environment it operates in. Yet, if applied in the appropriate settings, people can be confident that these mechanisms, and the blockchain protocols they serve, are safe and dependable.

The next question we ask is: are digital assets on blockchains secure? Here unfortunately the short answer is: not really. Firstly smart contracts, which power digital assets, are susceptible to a Pandora's box of vulnerabilities, and this has already led to a number of serious hacks involving the theft and loss of millions of dollars. These vulnerabilities tend to be related to the complexity of the underlying code and the business logic, as well as to the fact that this is relatively new technology with as yet no widely accepted set of standards and security best practices. Luckily new tools and techniques to audit smart contracts and publicise vulnerabilities and best practice are being developed, and we can expect smart contract security to improve. Secondly, in most blockchain implementations, the blockchain itself makes up a rather small portion of the overall platform. Digital assets are therefore susceptible to a large array of what we might call traditional cyber security vulnerabilities, for example vulnerabilities in database software, websites or APIs; or vulnerabilities related to human error. While the steps needed to mitigate these vulnerabilities are often well known, they are unfortunately as often overlooked. This too is not a problem specific to blockchain.

EXECUTIVE SUMMARY

The third question we tackle is: are blockchains private? We trust our banks to keep our transaction data private because that is their job, and that is the law. But can we trust a public blockchain run by a large but loose community of anonymous nodes? Here the answer is mixed. Contrary to popular belief, even though they contain no personally identifying information, transactions on public blockchains like Bitcoin often can be traced back to real identities. This is possible because the ledger is public and therefore open to forensic analysis. It is much easier to protect data on private blockchains, yet here too there are vulnerabilities, as data privacy depends more on best practice and the honesty of network participants. There is, however, good news on the horizon for data privacy on blockchains. New data obfuscation and privacy-preserving technologies, like ring signatures, homomorphic encryption and zero-knowledge proofs, are maturing and will provide tools to greatly enhance data security.

Last but not least, we look at the question of whether or not we can use blockchain to enhance cyber security generally. We find that, despite some of the privacy concerns with public ledgers, blockchains can potentially be used to enhance the security of data. For example, they can be used to defend against unauthorised access to data, and so enhance data confidentiality; they can be used to prevent data tampering and provide audit trails of transactions that can be used to investigate fraud, and so help support the integrity of data; and they can be used to help secure information on the provenance and validity of data, and so support data authenticity. We also look at a number of specific cyber security use cases for blockchain, including at the network level or in such areas as supply chains, medical records, verifiable software updates and anti-counterfeiting.

We close with some recommendations. Firstly, policy makers should encourage the systematic disclosure and documenting of protocol and smart contract vulnerabilities to help more quickly spread the word about issues and so support best practice. Second, we think that policy makers should strongly recommend – and, potentially, require – that all smart contracts be professionally audited. In lieu of direct regulation, policy makers could consider issuing quality certificates for smart contracts indicating if they have been audited and how. We believe that education and best practice will be a key element in increasing blockchain security as well. Policy makers should look to support efforts in this direction.

Security issues in blockchain

Blockchain is designed to allow a group of strangers, potentially even adversaries, to come to consensus on information and then save trusted, shared records of that information, all without recourse to a third party authority.

There are a wide variety of use cases for blockchains, from economic transactions involving cryptocurrencies or other kinds of digital assets, to sharing trusted data in supply chain situations, to – via smart contracts – automating business workflows or machine-to-machine data sharing and transacting.

In many if not most cases, blockchain protocols are used as part of platforms to transact and store value, often in large amounts. As of this writing, the Bitcoin protocol has a market capitalisation of USD 175 billion.[1] This naturally raises questions regarding the extent to which blockchain platforms are secure. Below we examine – on a high level – these questions, first in terms of blockchain technology, and then in terms of digital assets.

ARE BLOCKCHAINS SECURE?

In technical terms, a blockchain can be described as a multi-party infrastructure based on a distributed network of actors working on a common task who may or may not trust each other and need to communicate and transact over a shared infrastructure.[2] To accomplish this, a blockchain protocol has to have a) a way for all the nodes on a peer-to-peer network, which may find itself in a hostile, unpredictable environment like the public Internet, to come together to validate data (typically transactions); b) a means of permanently recording this agreed-upon data in an agreed-upon order in a shared ledger in such a way that it cannot be tampered with and that it can be verified at any time by anyone on the network; and c) a way to prove ownership/control of a piece of data (often, a digital asset) and to allow that ownership to be transferred.

The question of blockchain security thus can be reduced to two main points: is the distributed ledger on which data is stored secure (meaning correct and immutable), and are the mechanisms used to come to agreement secure (meaning they produce the desired outcome, and cannot be manipulated)?

Blockchain ledgers are highly secure

The distributed ledger on a blockchain serves the latter two purposes: providing an immutable record of data; and providing a means to prove and transfer ownership of data. In both cases, we can say that blockchains are highly secure.

Immutability. Blockchain ledgers achieve immutability through the use of cryptographic hashes: unique numbers of a fixed length that can act as a kind of seal and unique, unforgeable ID for a set of digital data (this is often referred to as a digital fingerprint). In a blockchain, when a miner or validator has assembled a valid block, it hashes that block to seal and identify it. Assuming the miner or validator has received the right to append the next block, it adds the new block to the end of

[1] According to coinmarketcap.com the market capitalisation of Bitcoin was some 160 billion US dollars on 12 May 2020. Considering the volatility of bitcoin, this number may change dramatically over time.

[2] See Cyber Security Workshop Report, EU Blockchain Observatory & Forum, 3 April 2020.

its copy of the blockchain and then hashes the whole ledger. This seals and creates a unique ID for the updated version of the whole ledger that includes the new block. The process is then repeated for each new block, so that in effect as the chain grows, it has a unique seal and ID at every new link. This results in two important security properties. Keeping in mind that distributed ledgers are propagated around the network – every node keeps its own copy, so that there are many copies in circulation – it provides a way to check if a given copy is the agreed-upon one. Any interested party can simply re-hash the ledger (a simple procedure) and see if the resulting hash matches the agreed-upon value. If not, then something is different, and this copy of the ledger can be rejected. This guards against tampering. If some mischievous node comes along with a copy of the ledger in which it has attempted to add bogus transactions to enrich itself, or to double-spend funds, everyone on the network can easily detect the fraud.

Proving ownership. Blockchain uses a combination of public-key cryptography and digital signatures to prove ownership of data and allow that ownership to be transferred. This is highly important on blockchain where data often represents valuable digital assets. Public key cryptography can be used as a means to provide or derive addresses of the senders and receivers of transactions. A public key or the address derived from it is often likened to a public drop box. Someone can put something in the box – in this case information – but only the possessor of the private key can get the information out. Public key cryptography and digital signature techniques are well known and proven technologies, and so can be considered highly secure. That said, history has shown that cryptographic techniques are usually broken over time. If quantum computing becomes powerful enough, this process could be sped up dramatically with regards to current cryptography, exposing existing blockchain ledgers.[3] Therefore there are some risks involved, although these are not imminent and are not related solely to blockchain. Despite these risks, users today can be highly confident that data on a blockchain is secure, that blockchain ledgers are indeed append-only (information can only be added, not removed), and that they are for all intents and purposes immutable. We would, however, note that just because a blockchain ledger is immutable does not mean that the data it contains is correct or private. We look at these two issues in more detail below.

Blockchain consensus mechanisms are secure

A transaction history recorded on an immutable distributed ledger is useless if the transactions themselves are invalid or the record-keeping is inconsistent. So another key security question in blockchain is whether the methods used to verify transactions and agree on their order are trustworthy and fit-for-purpose. As opposed to centralised databases, in blockchain validation and record-keeping is a shared activity that requires coordination among independent nodes on the network – a process referred to as “consensus”.

Different blockchains use different types of consensus mechanisms. While blockchain consensus makes for a fascinating and highly technical field of study,[4] for our purposes it is sufficient to have a schematic overview of the processes involved. Generally, consensus in a

[3] The quantum threat applies to all forms of digital cryptography.

[4] For an overview, see the Observatory’s Academic Paper on the subject. Blockchain and cybersecurity: a taxonomic approach, Stefano De Angelis, Gilberto Zanfino, Leonardo Aniello, Federico Lombardi, Vladimiro Sassone, University of Southampton, October 2019.

SECURITY ISSUES IN BLOCKCHAIN

distributed blockchain network involves a) the election through some predefined process of a leader node (in blockchain parlance, a “block producer”) that proposes the next block to be added to the blockchain ledger, b) a process for the other nodes to accept or reject this addition, and c) a means to define and be sure that a block is permanently added (referred to as “finality”).

While there are many different consensus mechanisms (many of which pre-date blockchain), the most relevant in blockchain are Proof-of-Work (PoW), Proof-of-Stake (PoS), Byzantine Agreement Consensus protocols and Proof-of-Authority (PoA). Each of these has a number of sub-variants. For our purposes we can classify them into three main families:

- **Lottery-based** consensus mechanisms in which the leader is chosen via a deterministic but random process, consensus is determined by the chain with the most work or stake, and finality happens over time. For the sake of simplicity we can refer to this overall approach as Nakamoto consensus. Relevant Nakamoto consensus protocols are Proof-of-Work (PoW), for instance as used in Bitcoin, and Proof-of-Stake (PoS);[5]
- **Voting-based** protocols involving (usually a small number of) known nodes in which a leader is selected by voting, random selection, round robin, or similar processes, consensus is achieved through voting and, as a result, finality is generally immediate (when a block is appended, it is immediately immutable). These are generally categorised as Byzantine Agreement Consensus protocols;
- **Hybrid protocols** that aim to exploit the advantages of each of the two. These include, for instance, Delegated Proof-of-Stake (dPoS) and Proof-of-Authority (PoA).

How do these mechanisms differ in terms of security? There are a number of things to take into account when evaluating this question.

First, it is good to understand the environment these consensus mechanisms must operate in:

- **Network model:** Is the environment the wide-open, anything-goes Internet in which there is no guarantee that a message sent will be received in a timely manner, or a private network that features access controls and can be managed by a network administrator?
- **Trust assumptions:** Can anyone join the network as a node so that the nodes cannot necessarily trust each other; or are the nodes authenticated first and so are known to each other, meaning trust can be enforced via off-chain methods like contracts and the legal system?
- **Network environment:** Is this network in a hostile environment, with a risk that malicious nodes may join the network, or that honest nodes can be co-opted?
- **Network size:** How big is the network? Is there a large, ever-changing number of nodes, or rather a small, fixed set?

Second, security in a distributed network is often judged in terms of different tradeoffs. Among those that are relevant to the blockchain security discussion are:

- **Blockchain trilemma.** The blockchain trilemma posits that blockchains are restricted to two of the following three properties: scalability (performance in terms of speed and volume), decentralisation and or security. If a blockchain is to be

[5] See Proof-of-Work is not a Consensus Protocol: Understanding the Basics of Blockchain Consensus, Stefan Beyer, 1 April 2019.

SECURITY ISSUES IN BLOCKCHAIN

highly decentralised and highly secure, it will come at the cost of scalability. If it is highly performant and highly decentralised, it will not be secure. Similarly, if one is willing to accept a degree of centralisation, it is possible to build highly secure and performant blockchains.[6]

- **CAP theorem.** Distributed networks are also often evaluated in terms of the CAP theorem. Here the properties are a) "consistency", meaning the degree to which all nodes have the same information at the same time; "availability", meaning the degree to which the network is available such that any node that sends a message will get a response (if not always an up-to-date one); and "partition tolerance", meaning the degree to which the network is able to continue running even when a certain number of nodes fail. Sometimes these concepts are expressed with different terminology. When talking about the underlying algorithms, the term "liveness" is used to mean the ability of the algorithm to continue functioning under different circumstances, and "safety" to refer to its ability to keep delivering correct results.

Different protocols exhibit different mixes of these properties.

Lottery-based protocols using Nakamoto consensus work well in hostile environments with no trust and with large numbers of nodes that may or may not be available at any given time. They are, however, not very performant. In terms of the blockchain trilemma, they favour decentralisation and security at the expense of scalability.[7] In terms of the CAP theorem, they favour availability and partition tolerance at the price of consistency. Finality in such approaches is also probabilistic: it does not occur immediately, but rather over time as more blocks are added to the chain.[8] That is because it becomes increasingly difficult to alter a block the farther down the list it is. Yet, because there is always a small but theoretical chance this could happen, this is referred to as probabilistic consensus.[9] These approaches are considered good for public blockchains because they work with large numbers of often unidentified nodes.

Because voting-based protocols like PBFT rely on heavy message exchange between nodes to come to consensus, they tend not to scale well, but with a limited number of nodes they can offer high performance and immediate transaction finality. In terms of the blockchain trilemma, they favour decentralisation and performance over scalability. Voting-based consensus protocols are considered highly suitable to private blockchains with a limited number of nodes and reliable network conditions.

Each of these consensus mechanisms has its cyber security vulnerabilities as well. PoW is well known for its susceptibility to the 51% attack: if a single entity gains more than half of the computing power (known as hashing power) on the network, it could in theory rewrite the transaction history in its favour. In practice, in large, high-vale networks like

[6] See Scalability, Interoperability and Sustainability of Blockchains, EU Blockchain Observatory & Forum, 6 March 2019.

[7] Bitcoin, for example, manages about seven transactions per second.

[8] This can be referred to as "eventual consistency". The Bitcoin protocol stipulates that the longest chain is the correct one, the one around which the network has reached consensus, because that is the one into which the most work has been put. Blocks, however, are not considered irreversibly added to the blockchain until a certain number of new blocks have been appended on top of them. In Bitcoin it is often said that transactions are final after six blocks. This is, however, just a convention, not a mathematically proven fact. Most PoS consensus protocols rely on the longest chain as well.

[9] Some PoS protocols add another layer that attempts to bring immediate finality. For the sake of clarity we are discussing PoS without such additions.

Bitcoin, this is prohibitively expensive. However, smaller PoW networks are susceptible, and such attacks have been known to occur.[10] PoW is also prone to selfish mining, where miners increase their chance of a reward by pre-mining blocks. Byzantine approaches require at least two-thirds plus one of nodes to be honest to work, meaning they are vulnerable if one-third of the nodes are malicious and cooperating. PoS protocols can also be susceptible to what are known as long-range attacks, collusion among delegators or poor randomness.

Despite these security issues, overall we can consider blockchain consensus mechanisms to be secure in the sense that they can be relied on to work as designed. The most obvious example of this is the Bitcoin network, which has never been compromised.

ARE DIGITAL ASSETS SECURE?

While blockchains are generally secure, digital assets held on blockchains unfortunately are often not. This is a major problem as many blockchain platforms are designed to handle transactions and store value, making them a preferred target of cyber criminals. It can also make errors extremely costly in the literal sense: a lost private key can mean an irrecoverable loss of funds. According to one report, some USD 1.7 billion were lost or stolen on blockchains in 2018.[11]

There are two main types of vulnerabilities affecting digital assets on blockchains. One comprises vulnerabilities in the smart contract technology that underpins most digital assets. The other comprises issues traceable to traditional cyber security vulnerabilities not specific to blockchains.

Smart contract vulnerabilities and fixes

The term “smart contract” as used here refers to code deployed on a blockchain that contains instructions for carrying out some process based on certain conditions. Because they are stored on an immutable blockchain, once deployed smart contracts can in theory not be stopped. This makes them useful for a wide variety of purposes, from making self-executing business agreements that all parties can trust will execute as written to creating tokens and digital assets.

Smart contracts can make blockchains very powerful, but they also open a Pandora’s box of security vulnerabilities – the more sophisticated a smart contract programming language is, the more possibilities there are for bugs. At our Cyber Security workshop a smart contract auditor reported that, of the 22 contracts he had audited in the previous year, none was error free, an indication of how difficult it is to get smart contract code right the first time.[12] These errors can have serious consequences. Famous incidents involving smart contracts include The DAO hack, in which USD 70 million was stolen, and the Parity Multisig bug which resulted in USD 300 million being lost forever.

These incidents are illustrative of different types of smart contract vulnerabilities. In the DAO incident, a hacker used a relatively well-known exploit referred to as “re-entrancy” to siphon funds out of a smart contract that had been deployed in an effort to create an autonomous

[10] Ethereum Classic 51% Attack – The Reality of Proof-of-Work, CoinTelegraph, 10 Jan 2019.

[11] Cryptocurrency thefts, scams hit $1.7 billion in 2018: report, Reuters, 29 January 2019.

[12] Cyber Security Workshop Report, EU Blockchain Observatory & Forum, 3 April 2020.

venture capital fund. Because it was a bug in the logic of the contract, the anonymous hacker even claimed that what he or she was doing was perfectly legitimate, as the smart contract allowed it. In the end, the Ethereum community had to agree to roll back the blockchain to its pre-exploit state and return the funds to investors. This caused a major schism in the community by breaking the law of immutability on the chain.[13] Despite being a well-known vulnerability, smart contracts are still falling prey to re-entrancy attacks today.[14] In the Parity incident, an anonymous user who was examining the Parity wallet library accidentally activated a kill switch in the wallet. The problem here was that the architecture was based on a library contract used by all wallets. Because of a permissioning issue anyone could make themselves the owner and execute the kill switch. Here the vulnerability was lack of adequate checks and controls in the code. Unfortunately, there is no way to undo the command, and the funds have been lost forever.[15]

Other vulnerabilities are related more to the environment smart contracts run in than the contracts themselves. In an exploit known as “front running”, attackers make use of the fact that on many blockchains, all transactions are public in the memory pool before they are confirmed. That in essence gives an attacker a view of upcoming transactions, enabling a kind of insider trading for clever and quick operators. Updates to the underlying blockchain protocol, which in essence change the rules of the game for smart contracts, can also cause vulnerabilities. The Ethereum Constantinople update was postponed because researchers found that the new version would break certain existing smart contracts.[16]

New issues are emerging related to the combination of different financial primitives that are being deployed on blockchains. For example, we recently saw exploits involving flash loans in decentralised finance applications.[17] Some systems combine different protocols such that, while each in itself is secure, the combination is not. It is therefore not enough to look at the code to be sure the smart contract delivers what it promises; it is also important to ensure it interacts safely with other smart contracts or if it can be gamed. These are issues not generally seen in traditional cyber security, and require different types of skills, for example a good understanding of financial instruments and/or data science.

Another security issue affecting digital assets and smart contracts is the difficulty of uniquely identifying these assets. To take one example, both Bitcoin Gold and the now inactive BitGem, carried the same BTG ticker symbol on many exchanges. The situation can be more complicated after a fork, as when Ethereum and Ethereum Classic split. The ambiguity caused by unclear references can create confusion in the market that could pose a danger to investors (for example, if a smart contract references the wrong cryptocurrency). One way to mitigate this is through a universal token numbering system.[18]

[13] The Story of the DAO – Its History and Consequences, Samuel Falkon (Medium), 24 December 2017.

[14] Uniswap/Lendf.Me Hacks: Root Cause and Loss Analysis, PeckShiel, 19 April 2020.

[15] What Caused the Accidental Killing of the Parity Multisig Wallet & How to Detect Similar Bugs, Bernhard Mueller, HackerNoon, 8 November, 2017.

[16] Security Alert: Ethereum Constantinople Postponement, Ethereum Blog, 15 January 2019.

[17] The DeFi ‘Flash Loan’ Attack that Changed Everything, Coindesk, 27 February 2020.

[18] For one such proposal see Unique Referencing and Identification in the Token Universe: Cross-Chain, Worldwide, and Fork-Resilient, Philip Sandner, 7 April 2020.

SECURITY ISSUES IN BLOCKCHAIN

The main vulnerabilities in smart contracts are, however, related to their code. So how can we make smart contract code more secure? There are several ways. The best way is to conduct security audits, which can detect flaws in the code but also in the business logic (for example, unintended consequences). Full, manual audits can be very expensive and take a lot of time. For this reason, many smart contract developers either do without them, or have the audit done once the contract is written. Unfortunately, at that point it can be extremely expensive to fix the problems.

The good news is that tools are being developed for automated audits, a process known as program analysis. This allows developers to test for vulnerabilities as they are writing the code. Program analysis, however, can only check for known vulnerabilities and logic issues. Because smart contract technology is relatively new, and is evolving quickly, there is no comprehensive or standardised library of best practice available, although there are those who are working on such libraries.[19] Another way to avoid problems with smart contracts is to not try to do too much with them. Many issues can be traced to complexity. Staying with template contracts that are known to be safe, or even developing smart contract languages with limited functionality (thus reducing the potential for error), could lead to safer code and better outcomes.

Traditional cyber security risks

Most of the risks to digital assets on blockchain have nothing at all to do with blockchains, and everything to do with traditional cyber security. These off-chain security issues include vulnerabilities in databases, in websites, in APIs and – a particular concern in blockchains – in key management. Such non-blockchain components can make up 80-90% of a blockchain platform. Yet even though most of these vulnerabilities are well known, as are the steps needed to mitigate them, they are often overlooked.

Specific vulnerabilities can include insecure authentication mechanisms, the erroneous use of cryptography, or the exposure of private keys. The latter is particularly problematic, as it is generally private keys that control access to funds. There can also be vulnerabilities in the operating system of the host on which the node is run, or vulnerabilities in the network (for example, susceptibility to DDoS or eclipse attacks). As in any other software product, blockchains are also vulnerable to malicious behaviour by developers, for example backdoors introduced by disgruntled employees.

As is the case with cyber security in general, the biggest vulnerability is the human being. Laziness and carelessness, for example by using easy-to-crack passwords or taking other security shortcuts, are the bane of every cyber security specialist. This is no different in blockchain. Humans are also prone to social engineering attacks, like phishing or man-in-the-middle attacks. We have had many examples of bogus Internet sites posing as the legitimate website of an ICO, and collecting funds from unwary investors. Humans can leak information, either accidentally or deliberately. Other issues have to do with tradeoffs. Decentralised application developers are often strongly focused on user experience out of the understandable desire to make this nascent

[19] See for example the SWC Registry.

SECURITY ISSUES IN BLOCKCHAIN

technology easy to use and so build a user base. But ease of use can often mean reduced security.

Mitigating these problems is a question of general security hygiene and best practice. The good news is that much of the “what to do” part is well known. The issue is rather discipline. For the blockchain world in particular, it can also pay to understand just to what extent a blockchain platform is vulnerable to such non-blockchain security issues. While it can be easy to focus just on the blockchain parts, neglecting the other security aspects can be dangerous.

Blockchains and privacy

After the general security of blockchain technology, many users and potential users are concerned about whether data – particularly personal data – on a blockchain is private. This is an understandable concern in any transaction platform. We trust our banks to keep our transaction data private because that is their job, and that is the law. But can we trust a public blockchain?

IS DATA ON A BLOCKCHAIN PRIVATE?

There is a misconception among many in the general public that blockchains are anonymous transaction platforms. One frequently heard criticism about Bitcoin is that it enables criminal activity through anonymous cryptocurrency transactions.

The truth is rather different. While you do not have to identify yourself in any way to make a Bitcoin transaction, you do have to cryptographically sign the transaction with a technique known as public/private key cryptography. This creates an indelible link between a public key appended to the transaction, which acts an identifier and in a certain sense as a user name, and a private key, which acts in a way like a password and is, hopefully, known only to you. The record of all transactions is, however, public. Using clever analytical techniques, it is possible to connect the dots in such a way as to trace public keys back to individuals. A whole new discipline called blockchain forensics[1] has arisen to do just this kind of analysis, and has already been the undoing of many criminals under the false impression that cryptocurrency is an anonymous transaction medium. In fact, it is pseudonymous, not anonymous, and these are two very different things.

Another misconception is that data on a blockchain is encrypted. This is not necessarily the case. Cryptocurrencies use cryptography to function, but the transactions in Bitcoin, for example, are not encrypted. It is possible to inspect the transaction amounts and public keys of all entries in the ledger. Indeed, that is the point. Furthermore, as we pointed out in our GDPR paper,[2] even if data is encrypted, no encryption is 100% foolproof. History has shown that most cryptographic functions are eventually cracked. Many people also worry that as quantum computing becomes a reality, all existing cryptography will immediately be vulnerable, and that the transaction history of every transaction on every blockchain up to that point will be exposed for all to see.[3]

Yet even today data encrypted using reversible encryption – that is, encryption that can be unscrambled by the possessor of the requisite encryption key – is only as safe as the safety of the key. And, again as we point out in our GDPR report, under certain circumstances it is even possible to recover the original data from a hash, even though this process is supposed to be irreversible.[4]

There are many other issues here too, the details of which are beyond this paper.

[1] Forensics and Bitcoin, Forensic Focus, 16 January 2015.

[2] Blockchain and the GDPR, EU Blockchain Observatory and Forum, October 2018.

[3] Of course, along with quantum algorithm-breaking we will see the advent of quantum-resistant cryptography, which means that data will be safe. The problem is data that has already been encrypted by non-quantum methods is exposed.

[4] Blockchain and the GDPR, EU Blockchain Observatory and Forum, October 2018.

BLOCKCHAINS AND PRIVACY

The important thing to keep in mind when thinking about data privacy on a blockchain is that in general, because blockchain ledgers are by their nature shared – that is, public – the data on them is exposed to some extent.

ARE THERE WAYS TO MAKE DATA ON A BLOCKCHAIN PRIVATE?

The short answer is: yes. One way to do this is to use a private, permissioned blockchain. This is a blockchain that has access controls on it, and functions a bit like an Intranet. Because the identities of the actors on the network are known, there is more leeway for data encryption or restricting access to the data on the ledger – for example by the use of dedicated channels for individual transactions. Of course, here data is still exposed to traditional cyber security vulnerabilities, but at least the blockchain-specific vulnerabilities of a public ledger are mitigated.

There are also a number of data obfuscation techniques available for use on public blockchains. In our GDPR paper we discussed, for instance, the use of third-party indirection services that bundle transactions together and submit them on users' behalf, masking single transactions in a sea of bits and bytes. Ring signatures, where multiple parties sign a transaction in such a way that an outsider can be sure that one of the signers is legitimate, but not which one, can help mitigate the traceability of public keys. Other data obfuscation techniques include homomorphic encryption and secure multi-party computation.

One of the most intriguing technologies we highlighted in the GDPR paper was zero-knowledge proofs (ZKP). Since then there has been a lot of excitement, and a lot of work, done around zero-knowledge techniques, and at our security workshop we were treated to a detailed presentation on the state of the art in this realm.[5]

ZKPs, which were invented by Micali and Goldwasser almost 20 years ago, are a mathematical way to prove that you know something without revealing what exactly you know. The details of how this works involve very complex maths, but the advantages are easy to see. Chief among these is the ability to present an irrefutable claim without exposing the data behind it. The classic (if somewhat overused) example is the ability to prove the claim "I am over 18 years old" without revealing your birthday.

One intriguing element of this from a data privacy perspective is the ability to prove claims of legitimacy without compromising privacy. Instead of a transaction being completely anonymous, a party to a transaction could – in theory – append proofs to the transaction that he or she is a) a citizen of a certain country, b) lives in a certain region, c) has no criminal record, d) has passed the requisite AML/KYC checks at a certain bank, and e) is indeed in possession of enough funds to make the transaction – all without revealing their name, address, or size of their bank account.

Zero-knowledge technology can, however, also be used to create truly anonymous transactions. In this case, the proofs are used to verify to the platform that the transaction is legitimate, but no other data is revealed at all. This allows for a provably correct yet

[5] Cyber Security Workshop Report, EU Blockchain Observatory & Forum, 3 April 2020.

completely anonymous transaction ledger. There are already so-called privacy-preserving blockchains in existence that make use of this technology for this purpose.[6] That said, the technology is still quite new, and needs refining. Yet most experts believe that within five years it will become commonplace. This could catalyse a privacy revolution – or nightmare, depending on your point of view.

IS IT A GOOD THING TO HAVE ANONYMOUS BLOCKCHAINS?

Which brings us to a final question. Is truly anonymous transacting a good thing? For many this might seem on the surface of it to be the case. Yet the subject is a matter of debate. While as we saw Bitcoin is not anonymous, contrary to what some believe, the idea that a truly anonymous cryptocurrency could be a boon to criminals is certainly not far-fetched.[7]

There has always been a tension between the need for privacy and transparency in society at large, and this tension is directly reflected in blockchain with its public record-keeping coupled with its decentralising, citizen-centric ethos. We want to make it clear that we take no side in this debate, and proffer no opinions. However, anyone interested in the issue of blockchain and privacy, will likely want to take this aspect of the discussion into account as well.

[6] See for instance Aztec, Monero and ZCash.

[7] Cryptocurrency proponents like to remind people, however, that by far the most preferred medium of exchange for criminals and terrorists remains banknotes.

Blockchain for security

In this last section, we take a quick look at the question of how we might use blockchain to enhance cyber security.

CAN WE USE BLOCKCHAINS TO ENHANCE THE SECURITY OF DATA?

Considering that above we have written that data on blockchains is often not private, and that blockchain ledgers, because they are public, are exposed to various kinds of forensic analyses, it might seem counterintuitive to ask if blockchain can be used to enhance data security.

Yet, as was discussed in our security workshop,[1] there are potential uses for this technology in data security contexts. One way to analyse this in terms of what is often referred to in data security discussions as the CIA triad, in which the acronyms stands for confidentiality, integrity and authenticity.

- **Confidentiality.** In data security, confidentiality means the ability to keep data private and safe from unauthorised access or use. Blockchain could, in theory, help with confidentiality by protecting data against unintentional, unlawful or unauthorised access, disclosure or theft. Blockchains could also be used to support confidentiality through managing access rights or registries, for instance in self-sovereign identity use cases.
- **Integrity.** This has to do with the ability to keep data secure from tampering or loss, and ensuring that it remains consistent over its lifecycle (that is, that everyone has the same data). Blockchains can support integrity through things like timestamping and notarisation, for example in document certification. Because blockchain makes things transparent, it is good for creating audit trails, and so could be useful in investigating fraud after the fact.
- **Authenticity.** This has to do with ensuring that the origin of the data can be verified. Blockchain technology is quite well suited to the task of recording the provenance of information. This is one of the properties that has made blockchain very interesting for supply chain and other track and trace use cases, whether involving information about physical objects or transactions. We can imagine the idea being extended to track and trace all kinds of data.

WHAT USE CASES ARE THERE FOR BLOCKCHAIN IN SECURITY?

Along with data security, there are other potential security-oriented use cases for blockchain. Among those that we discussed during out security workshop were:

On the network security level, blockchain could be useful in situations where decentralisation helps security, for example managing domain names and making it difficult to hack a domain name server and take over an identity. It might also be possible to manage certificate authorities in a decentralised way on a blockchain, including perhaps devising a reputation system for such authorities, adding a level of security.

[1] Cyber Security Workshop Report, EU Blockchain Observatory & Forum, 3 April 2020.

SMART CITIES: CONVERGENCE IN ACTION

More generally, blockchains could be used to secure network logs, creating an audit trail that could help fight tampering. Such audit trails as well as proof-of-provenance can be useful in software validation too, ensuring that software is authentic and has not been altered during transit. This could be very useful when working with IoT devices or in such things as over the air updates of vehicle software, including in autonomous vehicles.

Blockchain could provide access logs for medical records so people know who has accessed their information.[2] Blockchain could also in theory be used for track and trace on information to help combat things like fake news. In a similar way, blockchain could help keep information entering a network more accurate by providing a consensus mechanism for oracles (data sources for blockchain, for example stock prices or weather data).[3] Clear provenance of information could also help fight social engineering attacks like phishing by making it easy to verify from whom a message has really come. EUIPO, the EU Intellectual Property Office, has been working with blockchain to fight counterfeiting and tax evasion and to ensure authenticity of software.[4]

Another important use case for blockchain with data security aspects is in the area of identity. Again, as we have written elsewhere, blockchain is being looked at to support decentralised identity, including self-sovereign identities. By giving individuals control over the storage and access to their personal data, as well as enabling them to collect and use verifiable credentials in different contexts, decentralised identities could lead to a much safer and more secure Internet.

[2] This is already being tried in Estonia, which also provides a trusted and secure means for doctors to send patient prescriptions directly to pharmacies, with the whole process auditable by the patient.

[3] See Convergence of Blockchain with AI and IoT, EU Blockchain Observatory and Forum, 21 April 2020.

[4] Using blockchain in the fight against counterfeiting - EUIPO launches a Forum to support concrete solutions in that field, European Union Intellectual Property Office, 7 February 2020.

Conclusion and recommendations

As we have seen, the topic of cyber security is core to blockchain technology. For this reason, ensuring that blockchain technology and the platforms that are built on it are secure and behave as expected is of great importance to furthering adoption. Below we provide some recommendations for ensuring that blockchains and digital assets can be safely deployed and used.

1. Disclosure of protocol and smart contract vulnerabilities.

It is important to incentivise responsible vulnerability disclosure in blockchains. Considering that blockchain protocols often hold large amounts of value, there can unfortunately be strong economic incentives to not acknowledge problems that could affect the economics of a blockchain network. The community has begun to come together to develop clear communications and resources about vulnerabilities, for instance with the SWC registry.[1] Such efforts should be supported and expanded.

2. Recommendation or requirement for smart contract audits.

Developers and other actors in the blockchain space should leverage the tools that are out there that let you formally verify smart contracts or make informed security decisions based on mathematical facts you can discover from these protocols and contracts. This could potentially become a regulatory requirement, although over-regulation could be a barrier to innovation (see next point).

3. Certification.

In lieu of regulation, policy makers could consider various kinds of security certifications for blockchain protocols, smart contract platforms, and perhaps smart contracts themselves. A good set of quality certificates could be a middle ground between safety and innovation.

4. Education and best practice.

One issue facing blockchain is the fact that not enough people understand cryptography and how to use it properly. There are incidents of blockchain projects not using it properly or developing their own cryptography, which does not necessarily work as intended. We recommend efforts to increase education, expertise and dissemination of best practice in this area. The same can be said of smart contract technology. It is equally important that people understand how big the consequences can be of errors in immutable contracts and with blockchain protocols.

5. Regulatory landscape.

As for blockchain and privacy, the regulatory and policy issues here are by now well known, particularly the tensions between blockchain and the GDPR. We continue to recommend that policy makers look at the standard for anonymisation of personal data, particularly in light of new technologies like ZK proofs.

[1] https://swcregistry.io/

Appendix – Blockchain Terminology

What is a blockchain?
Blockchain is one of the major technological breakthroughs of the past decade. A technology that allows large groups of people and organisations to reach agreement on and permanently record information without a central authority, it has been recognised as an important tool for building a fair, inclusive, secure and democratic digital economy. This has significant implications for how we think about many of our economic, social and political institutions.

How does it work?
At its core, blockchain is a shared, peer-to-peer database. While there are currently several different kinds of blockchains in existence, they share certain functional characteristics. They generally include a means for nodes on the network to communicate directly with each other. They have a mechanism for nodes on the network to propose the addition of information to the database, usually in the form of some transaction, and a consensus mechanism by which the network can validate what is the agreed-upon version of the database.

Blockchain gets its name from the fact that data is stored in groups known as blocks, and that each validated block is cryptographically sealed to the previous block, forming an ever-growing chain of data. Instead of being stored in a central location, all the nodes in the network share an identical copy of the blockchain, continuously updating it as new valid blocks are added.

What is it used for?
Blockchain is a technology that can be used to decentralise and automate processes in a large number of contexts. The attributes of blockchain allow for large numbers of individuals or entities, whether collaborators or competitors, to come to a consensus on information and immutably store it. For this reason, blockchain has been described as a “trust machine“.

APPENDIX – BLOCKCHAIN TERMINOLOGY

The potential use cases for blockchain are vast. People are looking at blockchain technology to disrupt most industries, including from automotive, banking, education, energy and e-government to healthcare, insurance, law, music, art, real estate and travel. While blockchain is definitely not the solution for every problem, smart contract automation and disintermediation enable reduced costs, lower risks of errors and fraud and drastically improved speed and experience in many processes.

Glossary

The vocabulary used in the context of blockchains is quite specific and can be hard to understand. Here are the essential concepts you should know in order to navigate this breakthrough technology:

- **Node:** A node is a computer running specific software which allows that computer to process and communicate pieces of information to other nodes. In blockchains, each node stores a copy of the ledger and information is relayed from peer node to peer node until transmitted to all nodes in the network.
- **Signature:** Signing a message or a transaction consists in encrypting data using a pair of asymmetric keys. Asymmetric cryptography allows someone to interchangeably use one key for encrypting and the other key for decrypting. Data is encrypted using the private key and can be decrypted by third-party actors using the public key to verify the message was sent by the holder of the private key.
- **Transaction:** Transactions are the most granular piece of information that can be shared among a blockchain network. They are generated by users and include information such as the value of the transfer, address of the receiver and data payload. Before sending a transaction to the network, a user signs its contents by using a cryptographic private key. By controlling the validity of signatures, nodes can figure out who is the sender of a transaction and ensure that the transaction content has not been manipulated while being transmitted over the network.
- **Hash:** A hash is the result of a function that transforms data into a unique, fixed-length digest that cannot be reversed to produce the input. It can be viewed as the digital version of a fingerprint, for any type of data.
- **Block:** A block is the data structure used in blockchains to group transactions. In addition to transactions, blocks include other elements such as the hash of the previous block and a timestamp.
- **Smart contract:** Smart contracts are pieces of code stored on the blockchain that will self-execute once deployed, thus leveraging

APPENDIX – BLOCKCHAIN TERMINOLOGY

the trust and security of the blockchain network. They allow users to automate business logic and therefore enhance or completely redesign business processes and services.

- **Token:** Tokens are a type of digital asset that can be tracked or transferred on a blockchain. Tokens are often used as a digital representation of assets like commodities, stocks and even physical products. Tokens are also used to incentivise actors in maintaining and securing blockchain networks.
- **Consensus algorithm:** Consensus algorithms ensure convergence towards a single, immutable version of the ledger. They allow actors on the network to agree on the content recorded on the blockchain, taking into consideration the fact that some actors can be faulty or malicious. This can be achieved by various means depending on the specific needs. The most famous consensus algorithms include proof-of-work, proof-of-stake and proof-of-authority.
- **Validator nodes:** Validator nodes are specific nodes in a network that are responsible for constituting blocks and broadcasting these blocks with the network. To create a valid new block they have to follow the exact rules specified by the consensus algorithm.

Learn more about blockchain by watching a recording of our Ask me Anything session.