UPPSALA
UNIVERSITET

UPTEC STS 19046
Examensarbete 30 hp
Oktober 2019

The GDPR Compliance of Blockchain

A study of legal and technical perspectives on regulating innovative technology

Karin Melin

UPPSALA
UNIVERSITET

Teknisk- naturvetenskaplig fakultet
UTH-enheten

Besöksadress:
Ångströmlaboratoriet
Lägerhyddsvägen 1
Hus 4, Plan 0

Postadress:
Box 536
751 21 Uppsala

Telefon:
018 – 471 30 03

Telefax:
018 – 471 30 00

Hemsida:
http://www.teknat.uu.se/student

Abstract

The GDPR Compliance of Blockchain: A qualitative study on regulating innovative technology

Karin Melin

This thesis aims to explore the compliance of blockchain technology and the GDPR. The GDPR was implemented for the EU member states in May 2018 with the purpose of harmonizing data protection regulation. However, the regulation is based on the notion that data is stored and processed in a centralized system. This causes an issue when it comes to distributed networks, and in particular with the distributed ledger technology (DLT), the underlying technology of blockchain.

For this thesis, a literature review has been conducted to investigate the problems of GDPR compliance for blockchain projects, and what technical solutions exist to make a blockchain solution more GDPR compliant. In addition, interviews have been conducted to investigate the technical and legal perspectives on the current and future situations of regulation and technology.

Compatibility problems mainly concern the immutability and transparency of a blockchain and examples of technical solutions that handle those problems can be found in the literature. Nevertheless, none of the discussed solutions are yet to guarantee full GDPR compliance. The technical and legal perspectives share ideas of the main compliance issues. However, differences such as interpretation of technical details can be identified, indicating problems to arise when regulating blockchains in the future. Further interdisciplinary work on guidelines for the GDPR is necessary for blockchain projects to be successful in complying with the regulation as well as to strengthen the technology neutrality of the GDPR.

Handledare: Edith Ngai
Ämnesgranskare: Christian Rohner
Examinator: Elísabet Andrésdóttir
ISSN: 1650-8319, UPTEC STS 19046

Populärvetenskaplig sammanfattning

I takt med att samhället i allt större utsträckning digitaliseras lagras också en allt större mängd personuppgifter i digital form. I maj 2018 trädde dataskyddsförordningen GDPR i kraft för att stärka och harmonisera alla de olika dataskyddslagar i EU som funnits sedan tidigare. GDPR syftar bland annat till att ge individer större kontroll över sina personuppgifter genom att ge dem rättigheter att få sina personuppgifter uppdaterade eller raderade hos företag eller organisationer. Det måste även finnas processer för att hantera ärenden som gäller personuppgifters förändring samt att roller som ansvarar för personuppgifter, så kallade personuppgiftsansvariga och personuppgiftsbiträden, måste utses. GDPR har en stor räckvidd, är omfattande och ska även vara teknikneutral. Dock är flera delar av GDPR grundade på premissen att data hanteras i en centraliserad IT-arkitektur. Detta leder till problem när personuppgifter förekommer i distribuerade nätverk med en decentraliserad struktur som grund. En decentraliserad struktur innebär även en decentraliserad makt för att hantera personuppgifter vilket är den makt GDPR i mångt och mycket fokuserar på ska begränsas. En teknologi baserad på ett distribuerat nätverk är distribuerad databasteknik (Distributed Ledger Technology (DLT) på engelska). DLT är den underliggande tekniken för en blockkedja, en typ av databas som blev känd i samband med att kryptovalutan Bitcoin presenterades år 2008. En blockkedja kan användas för en mängd olika tillämpningar och branscher och anses vara en framtidsteknik tack vare sina många användningsområden. Några exempel där blockkedjor används är för finansiella transaktioner, id-hantering eller värdeskapande genom klimatvänliga aktiviteter i token-baserade system.

En blockkedja är en typ av databas som möjliggör transaktioner att sparas i följd, tillsammans med en tidsstämpel. Blockkedjor där transaktioner av något värde utförs mellan individer, vilka innehar en digital plånbok kopplad till blockkedjan, skapar en krypterad adress som placeras i kedjan. Adressen som är kopplad till en specifik individ anses i dagsläget vara icke-anonymiserad vilket enligt GDPR gör att den ska tolkas som en personuppgift. Därmed blir dessa typer av blockkedjor mål för GDPR.

Data som placeras i en blockkedja kan anses vara permanent och är i vissa fall öppen att se för allmänheten. Dessa två egenskaper skapar problem i relation till GDPR på flera sätt. Särskilt rättigheten att ändra eller radera personuppgifter blir svår att uppfylla, men även frågan om vem som tar rollen som personuppgiftsansvarige respektive personuppgiftsbiträde kan bli svårlöslig. Den här uppsatsen ämnar till att undersöka problematiken att följa GDPR i samband med att implementera en blockkedje-lösning, samt att undersöka tekniska, respektive juridiska, perspektiv på nämnda problematik. Detta för att utreda hur framtiden för blockkedjor kan komma att påverkas av GDPR.

I en litteraturstudie har problemområden för blockkedjor i relation till GDPR undersökts. De problemområden som identifierades har diskuterats i relation till de två huvudtyper av blockkedjor som finns. Blockkedjor kan vara publika, öppna för vem som helst att läsa information från eller gå med i, eller privata, skapade för ett i förväg bestämt antal aktörer som vill dela information. Olika typer av tekniska lösningar för att följa GDPR framkom i litteraturstudien och presenteras i uppsatsen. Inga av de diskuterade tekniska lösningarna kan dock erbjuda en garanterad enlighet med GDPR men erbjuder generellt ett förhöjt skydd av persondata såsom adressen som sparas vid en transaktion.

I intervjuer med respondenter, med både tekniska och juridiska bakgrund, framgick det att de delar syn på flera av de problemområden som finns gällande blockkedjor och GDPR. Det framgick även att båda parter identifierar att det i dagsläget har infunnit sig en typ av dödläge. Projekt med blockkedjor är osäkra på hur de ska ta sig vidare eftersom det finns rättsliga tveksamheter. Samtidigt behöver vidare riktlinjer för GDPR tas fram från juridiskt håll för att kunna analysera blockkedjan ur ett rättsligt perspektiv. Ytterligare framkom hur det finns skillnader mellan grupperna angående hur man värderar egenskaperna en blockkedja ger gentemot risken att inte vara kompatibel med GDPR. Detta kan vara en indikator på att problem kan komma att uppstå när lag och teknik ska utvecklas framåt. Båda grupperna ser att tvärvetenskapliga insatser behövs för att komma vidare. Dels för att möjliggöra riktlinjer som vidgar GDPR:s neutralitet till teknik men även för att kunna utvärdera blockkedjor och dess tillämpningar på ett rättssäkert sätt.

Acknowledgements

This thesis, which will be the conclusion of my studies at Uppsala University, could not have been conducted without the help of a few people worth mentioning. First, I would like to thank my supervisor Edith Ngai for offering me valuable support and guidance in writing this thesis. In addition, I would like to mention all the people I have been in contact with, discussing the SimpliCITY project from both Uppsala and Salzburg. Your work is inspiring, and I am happy to have gotten an insight into the work of encouraging communities to become more sustainable.

Next, thank you to my academic supervisor, Christian Rohner, who has given me valuable insights and feedback as well as encouragement throughout the process. The conclusion of this thesis would have been less interesting without your input.

I would like to thank everyone which have been in contact with me, giving me tips on interview respondents or other leads for me to follow. And of course, a great thank you to all of the respondents participating in the interviews. Thank you for giving me your time and well formulated answers to abstract questions on technical details as well as educating me on legal terms. Thank you for the many wise discussions and thoughts on blockchain, privacy and climate change which might not have made it all the way into this thesis, but definitely made the process of writing it even more interesting.

Lastly, thank you to family and friends who have offered support and feedback during the writing of this thesis. It would not have been the same without you.

Karin Melin
Uppsala, October 2019

Table of Contents

1. Introduction **1**
1.1 Aim and Research Questions 2
1.2 Delimitations 2
1.3 Overview 3

2. Background **4**
2.1 Distributed Ledger Technology and Blockchain 4
2.1.1 Distributed, centralized and decentralized systems 4
2.1.2 Blockchain: a decentralized, distributed system 5
2.1.3 The details of a block 8
2.1.4 Cryptographic components 10
2.1.5 Types of blockchain 11
2.1.6 Cryptocurrencies 12
2.2 The General Data Protection Regulation (GDPR) 13
2.2.1 Personal data 13
2.2.2 Data controllers and processors 14
2.2.3 The six protection principles 14
2.2.4 Privacy by Design and Privacy by Default 15
2.2.5 Rights of the data subject 15
2.2.6 Reach of the GDPR 16
2.3 Incentive Projects 16
2.3.1 The SimpliCITY project 17

3. Method **18**
3.1 Choice of Research Method 18
3.2 Literature Review 19
3.2.1 Location of studies 19
3.2.2 Selection and evaluation 20
3.2.3 Analysis and synthesis 20
3.3 Interviews 21
3.3.1 Choice of respondents 21
3.3.2 Semi-structured interviews 23
3.3.3 Data analysis 24

4. Results **25**
4.1 The Compatibility of the GDPR and Blockchain 25
4.1.1 Personal data in blockchain 26
4.1.2 Permissionless blockchains 28
4.1.3 Permissioned blockchains 32

4.2 Making Blockchains More GDPR Compliant 33
4.2.1 A permissioned solution 33
4.2.2 Channels and private data collections in Hyperledger Fabric 34
4.2.3 Implementation of cryptographic primitives 35
4.2.4 Digital identities enabling Privacy by Design 38
4.3 The Relationship of Technology and Legislation 39
4.3.1 Technology need to adapt to slow-paced regulation 39
4.3.2 Consequences of a status quo 40
4.3.3 Test environments 41
4.3.4 The trust of blockchain 41
4.3.5 The future for the GDPR and blockchain 42

5. Discussion 45
5.1 The Non-Compliance of a Blockchain 45
5.2 The "More GDPR Compliant" Blockchain 45
5.3 The Technical and Legal Perspectives 47
5.4 Further Research 48

6. Conclusions 50
6.1 The Situation on GDPR Compliance for Blockchain 50
6.2 What are the Solutions? 50
6.3 Interdisciplinary Research is Needed 50

References 52

Appendix A 61

Appendix B 62

Appendix C 63

Appendix D 64

1. Introduction

In 2011 the World Economic Forum released the report "Personal Data: The Emergence of a New Asset Class". It said the following:

> The rapid rate of technological change and commercialisation in using personal data is undermining end user confidence and trust. Tensions are rising. Concerns about the misuse of personal data continue to grow. Also mounting is a general public unease about what "they" know about us (World Economic Forum, 2011).

Since then, several scandals have been uncovered where personal data has been gathered in an unethical way to foster outcomes hurting democracy. For example, the Cambridge Analytica-scandal which was revealed in early 2018 (Rosenberg & Frenkel, 2018). In May the same year, the General Data Protection Regulation (GDPR) was taken in force in the EU to update the previous regulations for protecting personal data. The GDPR covers a wide range of aspects of collecting and processing personal data. A year after its enforcement, over 65 000 incidents have been reported (Lindström, 2019).

The GDPR aims to harmonize data protection regulation for the EU member states and is stated to be technology neutral to cover all digital processing of personal data. However, the regulation is based on the notion that data is stored and processed in a centralized system. This causes an issue when it comes to distributed networks, and in particular with the distributed ledger technology (DLT), the underlying technology of blockchain. Blockchain, a technology made famous by the implementation of the cryptocurrency Bitcoin, is thought to have potential to revolutionize a broad spectra of sectors, from the financial sector to political institutions (Zhang et al., 2019). The technology offers an easy setup of cryptocurrencies, marketplaces, and service information and exclude intermediaries such as banks or tech companies (Lyons, 2018). The ability to store value on a blockchain can also contribute to put value on actions or resources not fitted into the regular economic system, such as environmental actions. This has caused a movement of incentive projects focusing on climate change and sustainable services. Using blockchain, which enables so called 'token economies', a buy and sell of environmental services is made possible (Denis Le Sève, 2018). This thesis is conducted in collaboration with the SimpliCITY project, one incentive project aiming to use blockchain for increasing usage of regional sustainable services.

Blockchain is believed to be a must-have technology for many companies and the investments during 2019 are estimated to $2.9 billion, an increase of 88 % from 2018's $1.5 billion (IDC, 2019). The fast-paced development of blockchain is however discouraged because of regulations. In a survey from 2019 with senior executives of companies investing in blockchain technology, 50 % said that privacy was their main regulatory issue when implementing blockchain in their business (Deloitte, 2019). The potential non-compliance

with the GDPR is an unresolved issue for many blockchain projects. By contrast, the GDPR and blockchain are sometimes discussed as sharing objectives, such as security and individual control (De Meijer, 2018). Would it be possible for the technology and regulation to overcome their contradictions to co-exist and work together towards these shared objectives?

1.1 Aim and Research Questions

The aim of this thesis is to investigate the compatibility of blockchain and the GDPR, clarifying the prerequisites of ensuring personal data protection in a blockchain environment. Technical innovation and legal regulation does not evolve without a context. Consequently it becomes relevant to explore the different perspectives on the situation of blockchain and the GDPR. This could help in finding answers to what the future of blockchain will look like. The perspectives from technical and legal sources were therefore chosen to study this further.

The research questions developed based on the aim of this thesis are:

- What are the main compatibility issues with the GDPR and blockchain?
- What technical solutions exist to make a blockchain solution more GDPR-compliant?
- What are the technical and legal perspectives on the future of blockchain with regards to the effects of the GDPR?

1.2 Delimitations

To answer the research questions, decisions have been made to exclude excess information and research. The GDPR is a broad regulation which contain many aspects of handling personal data. For this thesis, the important sections of the GDPR has been chosen to be explained and discussed based on what has been found in literature to be important. Particular aspects of GDPR, for example how there are special conditions for citizens under 13 years of age or the regulation about transferring data outside of EU, have not been included for discussion. This is due to the focus on the general compliance issues of the GDPR and blockchain technology rather than looking into a detailed solution.

When it comes to investigating the GDPR, Finck (2019) explains how it is important to note how blockchains is not one technology, but rather a class of technologies. This is an important difference. Hence, the question of compliance need to be analysed on a case-by-case basis rather than making general rules on what is allowed to do under the GDPR (Finck, 2019, p.i). This thesis will, however, discuss the GDPR and blockchain in general terms and touch upon the main compliance issues.

This thesis was conducted with the SimpliCITY project in mind, since their intended blockchain implementation is one potentially facing issues with the GDPR. A focus will be held on distributed ledgers used for transactions of some sort, where tokens or some other value in a wallet owned by a person is connected to the blockchain. For every unique

blockchain implementation, it is important to note that the following discussion might not be complete and that further evaluation of the situation need to be evaluated when it comes to a specific case.

Coming from a technical background, this thesis aim not to interpret regulation and the GDPR in detail. References to the original GDPR is consequently not included as literature. Instead, literature and respondents who are from a legal background is used to retrieve those interpretations. It is those sources which is used as the foundation for the discussion in contrast of the technical point of view. Multiple sources have been reviewed to get a cohesive overview of the GDPR, but it is possible that details stated in the regulation itself are not considered when looking at the broad discussions. Before implementing a suggested solution as described in this thesis, look further into the regulation and seek advice from an authorized party.

1.3 Overview

This introduction chapter will be followed by chapter two, containing a background of the blockchain technology and the GDPR. The second chapter also contain an introduction to incentive projects and describe SimpliCITY which is one such project. The third chapter explains the qualitative research methods used to conducts this thesis. Chapter four discusses the results found in the literature review and interviews when it comes to compatibility issues, technical implementations for making a blockchain more GDPR compliant, as well as a discussion on the perspectives on the situation from a legal and technical point of view. The results are discussed in chapter five and conclusions of the discussion are presented in chapter six along with proposed further research.

2. Background

In this chapter the Distributed Ledger Technology (DLT) will be explained in detail followed by an overlook over the main types of ledgers in section 2.1. Next, in section 2.2, an overview of the GDPR regulation and its implications is given. In section 2.3, a brief introduction to incentive projects is provided, followed by a presentation of the SimpliCITY project.

2.1 Distributed Ledger Technology and Blockchain

To understand the basics of distributed ledger technology and later blockchain, a first look into different networks is necessary.

2.1.1 Distributed, centralized and decentralized systems

A *distributed system* is a system where computers, *nodes*, work together to run an application, for example a database, distributed on each node in the network. All nodes in the network operate concurrently and have a shared state. If one node fails, it does not affect the system's uptime. To the end user, the application does not reveal that the system is running on multiple machines. The distributed nature of a system enable horizontal scaling, i.e. adding more computing power, and low latency (Kozlovski, 2018) which both are attractive features when it comes to computing. A distributed system, can have either a centralized or different degrees of a decentralized architecture.

Centralized network

In a centralized network, one main *server* is in possession of the necessary computing power, memory, and storage for the computing needed. The server accepts incoming connections from other nodes, *clients*, who all depended on this single source of service (Bashir, 2018), which is illustrated in Figure 1.

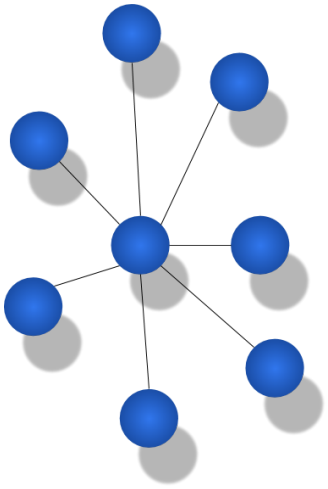

Figure 1. An example of a centralized system.

Further evolvement of this client-server structure led to the now widely used cloud computing architecture. Still based on the client-server structure, cloud computing enables clients to be more lightweight and get access to resources such as computational power and storage from a centralized server. However, the centralized network architecture presumes a trust between the main server, in this example the cloud distributor, and the clients using the system (Raj, 2019). If two clients want to communicate with each other in a centralized network they depend on that third-party which makes the third party a single point of failure.

Decentralized network

The opposite to a centralized architecture is a decentralized one. In a distributed and decentralized network, all nodes are connected to each other without a central point of connection, creating a *peer-to-peer network* (Singhal et al., 2018). In a peer-to-peer network, every node acts both as a server and a client (Cope, 2002), which is illustrated in Figure 2. In theory this makes the network more fault tolerant and have better attack and collusion resistance. Yet in practice, other factors than the architectural structure can cause these characteristics to be less dependable (Buterin, 2017). However, the dependency on a third-party which is inevitable in a centralized network, can be eliminated with a decentralized network.

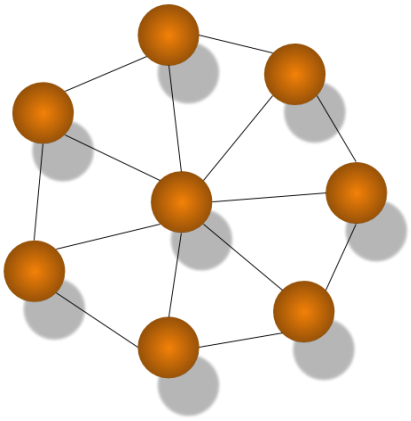

Figure 2. An example of a decentralized system.

2.1.2 Blockchain: a decentralized, distributed system

Distributed Ledger Technology

One example of a distributed and decentralized system is the distributed ledger technology (DLT) which provide this non-dependency of a third-party (Raj, 2019). A distributed ledger is a digital log of transactions stored in a distributed system. In DLT, the ledger is replicated across all nodes in the network (Singhal et al., 2018) and the transactions are validated by all nodes with the use of *asymmetric cryptography algorithms* (Zheng et al., 2017) which will be explained in section 2.1.4.

One type of a DLT is *blockchain*, a technology made famous by the implementation of the cryptocurrency Bitcoin (Raj, 2019). Bitcoin was introduced in 2008, by the person or group behind the name Satoshi Nakamoto. They published the article “Bitcoin: A Peer-to-Peer

Electronic Cash System" where they proposed a cryptocurrency, enabling a new way to carry out economical transactions without a third-party verifier. This also caused a break-through for the underlying technology of distributed ledger systems. Nakamoto called the ledger a *chain of blocks*. The term later transformed into the now used blockchain. In a blockchain, the ledger starts with a genesis block. New blocks are appended continuously and are linked together. A blockchain is, due to the linked blocks, considered to be nearly impossible to change, i.e to be immutable (Nakamoto, 2008).

Nodes which are included in the peer-to-peer network of a blockchain can be *Byzantine nodes* or *honest nodes*. If a node behaves in an arbitrarily fashion, which can be both intentionally malicious or just an unexpected behaviour, it is called a Byzantine node (Bashir, 2018). An honest node is a non-Byzantine node (Wang, W. et al., 2019), behaving as expected by the network. Since the ledger in a blockchain is distributed among the nodes in the network, they all share a process for keeping the ledger correct and up to date.

This distributed and decentralized system on which blockchain rely entails some positive properties: 1) there is no single point of failure, 2) it is democratic in its decision-making (Singhal et al., 2018), due to every node participating in a process to decide which blocks are added to the chain. This process is called consensus and will be explained later in this section.

However, the distributed character of the network also generates some challenges concerning coordination and connection of nodes in the network. These challenges can be illustrated by the *CAP Theorem*.

The CAP Theorem

The CAP Theorem, which was proved as a theorem by Seth Gilbert and Nancy Lynch in 2002, states that a tradeoff between properties of a distributed system is necessary. The CAP Theorem properties are the following:

- *Consistency*: Ensure all nodes have the latest version of the ledger.
- *Availability*: The system is accessible and accept requests and responses when required.
- *Partition Tolerance*: The system continues to operate correctly even if a group of nodes fails.

In a distributed system it is not possible to guarantee all three properties at the at the same time (Bashir, 2018). Or rather, a distributed system provides partition tolerance but not with both consistency and availability at the same time (Kozlovski, 2018) which is illustrated in Figure 3. For blockchain, the availability and partition tolerance are prioritised before consistency, which is achieved over time, called eventual consistency (Bashir, 2018).

Consensus mechanisms

Another aspect of blockchain being based in the context of a distributed and decentralized system is how to create agreement between nodes. Since each node stores one replica of the ledger and has the equal amount of authority as any other node, a *consensus* is needed to keep a common view of the blockchain over the whole network, i.e to decide which blocks are allowed to be added to the chain (Wang, W. et al., 2019). A block is validated if the transactions inside it are valid. This is checked by each node to some rules set by the creators of a specific blockchain (Botjes, 2017). Without a consensus, possible *Byzantine failures* may occur. Byzantine failures include malicious attacks, node mistakes (software inconsistency leading to unexpected *blockchain forks*, i.e. that the blockchain diverges into two paths), and connection errors (Wang, W. et al., 2019).

A blockchain can be said to achieve consensus if the following properties are fulfilled (Wang, W. et al., 2019):

- *Validity* (Correctness): All honest nodes activated on a shared state suggest an expansion of the blockchain by the same block and one honest node adopts the blockchain headed by the block in question and updates its state.
- *Agreement* (Consistency): An honest node confirms a new block header and then any honest node that updates its local blockchain view will update with the new block header.
- *Liveness* (Termination): All transactions originated from the honest nodes are eventually confirmed.
- *Total order:* Nodes that are honest accept the same order of transactions if they are confirmed in their local blockchain view.

A consensus is settled among the network by a consensus mechanism. There is a wide variety of consensus mechanisms, each with different characteristics and thereby suited to fit certain blockchains (Wang, W. et al., 2019). For example, Bitcoin uses the mechanism Proof-of-work (PoW) to avoid duplicate payments or misrepresentation of data. PoW verify that a node is honest by making the node do computational work, called *mining*, which require a lot of computational resources. If winning the race of mining the next block, the node is verified to contribute to add the next block to the blockchain (Zheng et al., 2017) and receives a reward of some currency. The reward gives an incentive for nodes to behave honest. The spent resources on mining will be lost if the block proposed contain a non-valid transaction and is denied by other honest nodes in the network (Singhal et al., 2018).

To be able to add a block containing non-valid transactions, a node, or a group of nodes, need to control more than 50 % of the total computational resources in the network to ensure that a block with a non-valid transaction is eventually added and accepted. With a majority of the nodes validating a non-valid transaction the byzantine chain fork will be accepted as the longest and will be followed by the rest of the network. This is called a 51% attack (Baliga, 2017). Other consensus mechanisms use other proofs than computational resources for a node

to show their honesty and for byzantine nodes to lose something if trying to add a block not validated by the rest of the network. A few types of consensus mechanisms together with their proofs can be seen in Table 1.

Table 1. Some examples of consensus mechanisms (Mattila, 2016; Wang, W. et al., 2019).

Name	**Abbreviation**	**What is needed as proof**
Proof-of-work	PoW	Burning computational power
Proof-of-stake	PoS	Holding tokens of the cryptocurrency
Proof-of-activity	PoA	Combination of PoW and PoS
Proof-of-burn	PoB	Destroying tokens of the cryptocurrency
Proof-of-validation	PoV	Deposit of tokens that is in danger of being destroyed
Proof-of-capacity	PoC	Storage capacity
Proof-of-importance	PoI	Active participation in the cryptocurrency
Ripple Protocol		Peer reputation (listed on a list as unique node)
Stellar consensus		Peer reputation (Quorum vote)
Practical Byzantine Fault-Tolerant Protocol	PBFT	Peer reputation (voting mechanism)

W. Wang et al. (2019) discuss how consensus mechanisms, for example PoW, depend on the majority of nodes in the network to be altruistic in forwarding information. This is needed to create convergence to the longest chain and to resolve the blockchain forks that are bound to appear. W. Wang et al. (2019) explain that the consensus mechanisms should strive to benefit decentralization and fairness, as well as have a balance between processing throughput and scalability of the network. When a consensus is reached, a block of transactions is added to the ledger.

2.1.3 The details of a block

As previously explained, a blockchain can be viewed as a linked list of blocks, creating a ledger. A block consists, in general, of a few components that are all working together to create the characteristics of the immutable and cryptographic secure blockchain. The two main parts of a block are the *block header* and the *block body*, which is shown in Figure 3.

The block header contains information about the block and the link to the previous block in the blockchain. The block body holds information about the *transactions* as well as some additional data (Bashir, 2018). The structure for the blocks can differ depending on the type of blockchain, but the header containing the linkage to the preceding blocks and the body with the transaction information are consistent for most blockchain types.

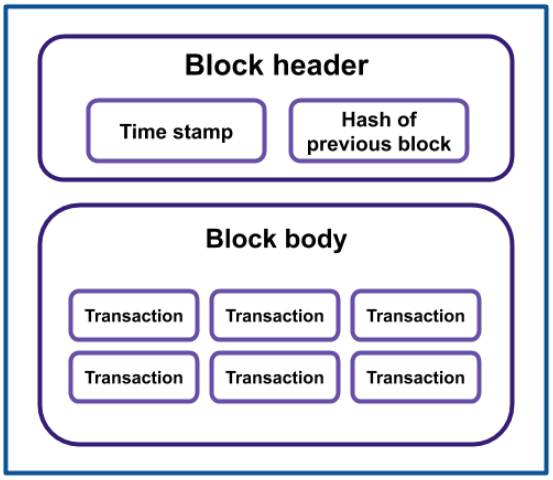

Figure 3. A simplified overview of a block with a header and a body.

Transactions

In most cases, a blockchain enables users, nodes, to send data such as tokens or coins to others and log the *transaction*. Transactions can be said to be the atomic data structure of a blockchain and are invoked either by a *smart contract*, a program which run on the blockchain, or by nodes wanting to send tokens. A transaction makes use of *cryptographic hashing* and *asymmetric encryption* to ensure a secure process (Wang, W. et al., 2019).

Smart contracts

The concept of smart contracts was introduced in 1994 by Nick Szabo but was first implemented with the emergence of blockchains (Raj, 2019). A smart contract is a transaction protocol, written in code, which can execute itself when conditions set in the contract are met (Lauslahti, 2017). When implemented in a blockchain, a smart contract can enforce transactions in a fully autonomous way. Smart contracts give blockchain solutions both flexibility and power, which is why it is desired to use them. The automation eliminates conditions to be checked by an intermediary and ensure that the instructions stated in the contract executes consistently (Raj, 2019). The usage of smart contracts in blockchains causes new types of platforms being built, for example in the finance and banking sector. Although, since smart contracts are a new type of concept which is not recognized by current contract law they do raise questions of how to regulate them (Lauslahti, 2017).

2.1.4 Cryptographic components

To ensure the blockchain to be cryptographic secure and immutable, cryptographic hashing is used. Cryptographic hashing is a function that maps a arbitrary-length set of data to a fixed-length output, called hashing the data (Wang, W. et al., 2019). The process is illustrated in Figure 4.

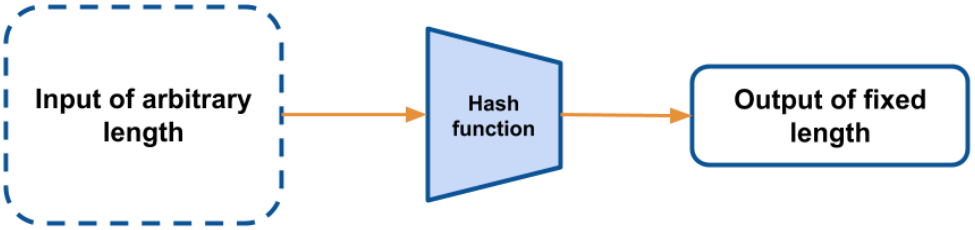

Figure 4. A hash function which process any output to a fixed-length output.

A minor change in the data gives a completely new hash if hashed again. If the function used is a secure hash function, the hashed data is infeasible to recover. It is also highly unlikely to get the same output from any two different inputs. In a blockchain, each block is hashed, and the hashed value is then put into the next-coming block. This hash stored in the block header is called *hash pointer* (Wang, W. et al., 2019) which can be seen in Figure 5.

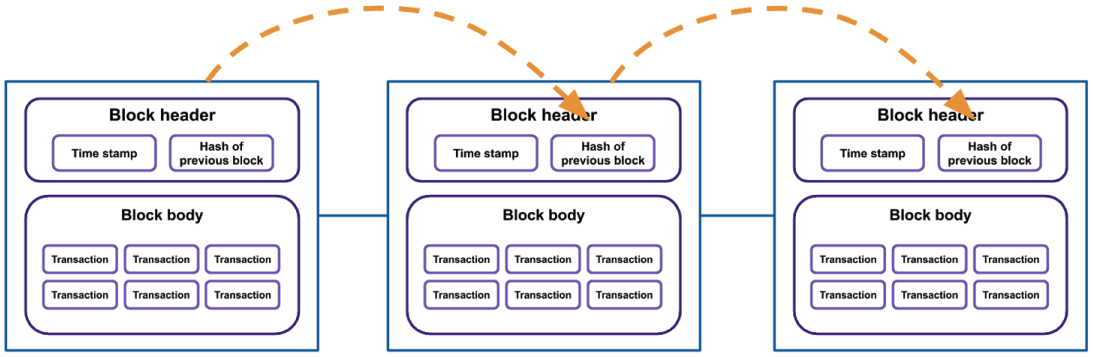

Figure 5. Overview of blockchain consisting of three blocks.

Since each block contains a hash of the entire blockchain up to that point, changes made in a previous block would require all hashes in the blocks thereafter to be recalculated which would require an immense amount of computational power. A blockchain is therefore often considered to be immutable (Bashir, 2018). Depending on how many hash pointers are allowed to point to a predecessor, a blockchain can be either a linear linked list, a tree of blocks or a Directed Acyclic Graph (DAG). The common view of a blockchain is however the linear type (Wang, W. et al., 2019).

Asymmetric encryption

Usually, a transaction is made from one sender node to a receiver node and logged in the blockchain. The identity of the nodes involved in the transaction are hidden by using asymmetric encryption and asymmetric keys. Each node has a generated pair of a *public* and a *private key*. The hashed public key of each node is the pseudo-identity of the node, the *address* of the node (Wang, W. et al., 2019).

The process of encrypting a transaction consists of a few steps. First, the sender hashes the data they want to send. The sender then uses its private key together with a digital signature algorithm and generate a *digital signature* (Singhal et al., 2018). A digital signature is usually a fixed-length string which is used as part of the input for the transaction. The digital signature is used to sign the data sent in the transaction. When receiving the transaction, the receiver can use the sender's address to check whether the data was sent from the sender by using a verification function. The verification function uses the signed data and the signature itself to determine if the signature was generated by the specific signature algorithm of the sender's private key, which can be verified using the sender's address. The receiver then hashes the data received themselves. If the signed hashed data is the same as the receivers hashed data, the transaction is considered as valid (Wang, W. et al., 2019). The process is illustrated in Figure 6 below:

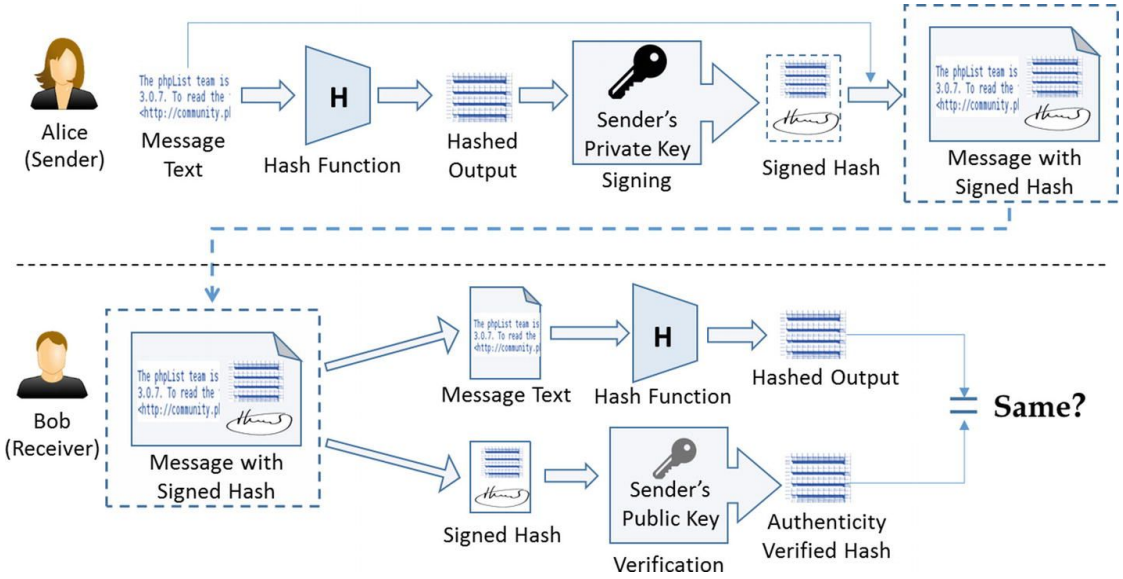

Figure 6. The Digital Signature Algorithm (DSA) used to sign a transaction and to verify it on a receiver side. Adapted from Singhal et al. (2018).

2.1.5 Types of blockchain

One can divide blockchains mainly into two types: *permissionless* (public) or *permissioned* (private) blockchains. In addition, a third alternative combining the two is sometimes discussed as a *consortium blockchain*.

A permissionless blockchain treats all nodes impartially. It is completely open and transparent, and it is the foundation for cryptocurrencies such as Bitcoin. A permissionless blockchain secures the system when the network is trustless, since the recorded transactions cannot be modified and thus fulfill the immutability property. It does however face a problem of scalability (Raj, 2019) due to its characteristic of being open for anyone in the world to join (Buterin, 2015). The earlier discussed consensus mechanisms used to ensure agreement between untrusted nodes usually requires a lot of computational resources that lead to wasted energy which on an aggregated level can become significant amounts (Mattila, 2016). To resolve this, permissioned blockchains is an alternative.

A permissioned blockchain has an access control, providing certain access rights to certain nodes in the network. The use case for a permissioned blockchain is a trusted network for when a ledger needs to be shared, e.g. within an organisation (Raj, 2019). In a permissioned blockchain, the nodes are known which make the blockchain not as dependent on consensus mechanisms as permissionless blockchains. In return they do not ensure immutability or censorship-resistance (Mattila, 2016).

To try to achieve the benefit of both types of blockchain one can create a blockchain with mixed elements from the both types. In a consortium blockchain, the consensus process is managed by a set of pre-selected nodes to get the benefits of the permissioned blockchain. The right to read or validate within the blockchain might be public but the consensus is settled by a smaller number of nodes who can reach an agreement more efficiently. This causes the solution to be cheaper and eliminate the risk of getting a consensus decision by a malicious majority of unknown nodes. Although, it is on the expense of the self-management of a permissionless blockchain. One can achieve many hybrid-solutions of blockchains and which one to be the right one depends heavily on which industry the system is created for. In simpler words: it depends on the application (Buterin, 2015).

2.1.6 Cryptocurrencies

Since Bitcoin entered in 2009, the currencies based on blockchain, so called cryptocurrencies, has become many. Cryptocurrencies all depend on blockchain technology and its consensus mechanisms to avoid the *double-spending problem* (Dierksmeier & Seele, 2018). The double-spending problem occurs if the same digital currency is used more than once i.e. when a transaction of a coin has been registered and another transaction with the same coin is added to the blockchain. With the consensus mechanism, each transaction is validated by the network. A double transaction of the same coin will therefore be denied, and the double-spending is avoided (Singhal et al., 2018). Cryptocurrency is retrieved either by exchange of regular currency or by taking part in the mining process and earn coins in exchange for computer power or previous equity (Dierksmeier & Seele, 2018). Making use of a permissionless blockchain, anyone can join and support the cryptocurrency of their choice. Each user then has a digital wallet connected to the cryptocurrency of use.

Cryptocurrencies enable transactions between a sender and a receiver to be made without a third party. This non-existing intermediary has made cryptocurrencies accused of fostering criminal activities (Janze, 2017). The largest cryptocurrencies today are Bitcoin, Ethereum and Ripple (CoinMarketCap, 2019) but new currencies are developed every day. Tech giants like Facebook are investing in creating their own version (Statt, 2019).

2.2 The General Data Protection Regulation (GDPR)

In 1995 the Data Protection Directive 95/46/EC was implemented by the European Community (now the EU) for the member states to transpose into national law. The directive aimed to protect individuals with regard of the free movement and processing of *personal data*. With its implementation, it also caused a fragmentation of the data protection across the EU. As a counteract and an enhancement of this directive, the GDPR has been taken in force since May 25th, 2018. The GDPR equalize the rules for data protection across the member states. By doing so, the obstacles for the flow of personal data is removed (Voigt & Bussche, 2017, p.1-2).

According to the GDPR, an organisation handling personal data has several the requirements to fulfill. The organisation needs to record its processing activities. A processing activity can be for example to collect, store, adapt, use, erase or destruct personal data (Calder, 2018, p.23). designate a Data Protection Officer and when feasible also a Data Protection Management System as well as make a Data Protection Impact Assessment. Data protection need to be ensured by design and by default. The GDPR also contains appeals to implement technical and organisational measures of safeguarding personal data. Further, the regulation introduces a breach notification report duty, of 72 hours after the getting knowledge of the breach. The rights that each individual, the *data subject*, have against the data processing entities are also stated (Voigt & Bussche, 2017, p.3-4).

Since the regulation is constructed to cover the general area of data protection, a wide range of topics are reviewed. The GDPR consists of 99 articles and 173 recitals, hence it is critical for any organisation to fully understand the terms (Pyle et al., 2018). The fines for violations can be up to 20 million EUR, or a 4% of the total worldwide annual turnover, whichever is higher (Voigt & Bussche, 2017, p.31). In the sections that follow, some of the most important highlights of the regulation will be explored.

2.2.1 Personal data

What kind of data is in fact classified as personal data? In article 4 of the GDPR, personal data is described as any kind of information related to an identified or identifiable individual. Data thus becomes personal data when some kind of identification of a data subject is possible. However, identification of an individual might be possible by combining different data that is not in itself classified as personal data. It is also not stated in article 4 who need to be identifying the data subject, but the additional information needed to identify an individual

must be “easily accessible” and if there is no chance that the controller or processor could retrieve the information it is considered not to be identifiable (Voigt & Bussche, 2017 p.12).

2.2.2 Data controllers and processors

The GDPR assigns the task of determines the purpose of the data processing to a *data controller*. The controller also determines the means of processing. A controller can be a legal person but might also be a public authority, agency or another kind of entity. More than one controller can exist, so called joint controllers, although it today is somewhat unclear what it requires to be appointed as a joint controller. A sentencing court in Germany stated that it is important to keep the term joint controller broad to properly ensure the data subject’s right. The evaluation should, as everything else, be done on a case-by-case basis (Finck, 2019, p.39-41). Another role is the *data processor*. The data processor is a legal person, public authority, agency or entity that process data on behalf of the controller. Often the controller and the processor are one and the same, but a controller can also have multiple processors in addition to itself (Calder, 2018, p.19).

2.2.3 The six protection principles

In article 5 of the GDPR, six protection principles can be found. These can work as a guide for organisations in how they should manage personal data. The six principles are the following (IT Governance Privacy Team, 2017):

1. *Lawfulness, fairness and transparency:* The data subject has right to be informed about how the data will be processed and the processing must follow this description. The processing must be for a purpose stated in the regulation.
2. *Purpose limitation:* The purpose of the data processing must be stated and limit the data used to that purpose.
3. *Data minimisation:* No data besides what is strictly required should be acquired and processed.
4. *Accuracy:* The data needs to be accurate and, where it is possible, also updated.
5. *Storage limitation:* If data is no longer needed for the previous stated purpose, the data should be erased.
6. *Integrity and confidentiality:* Data must be handled in a manner to ensure appropriate security for personal data. That includes protection against unauthorised, unlawful processing.

Another principle that could be argued to be the ‘seventh’ can be found in clause 2 of article 5 (IT Governance Privacy Team, 2017):

7. *Accountability:* A data controller should be responsible for ensuring compliance with the earlier stated six protection principles.

2.2.4 Privacy by Design and Privacy by Default

An incentive from the GDPR is to include data privacy in digital systems in larger extent. This is enforced by the concept of Privacy by Design and Privacy by Default from article 25

GDPR. Here it is emphasized that developers are bound to have data minimisation as a focus when implementing new digital systems as well as use *pseudonymisation*, i.e. replacing data with a pseudonym to hinder the data subject to be directly identifiable (Data Protection Commission, 2019), of personal data as much as possible. Only personal data critical for the set-out purpose of the system should be collected. Technical and organisational structures should be adapted according to minimisation of data. Privacy by Default implies that the default technical option should be protecting the personal data in the furthest extent. A user should not have to opt-out but rather opt-in on storage of their personal data (Voigt & Bussche, 2017 p.63). Since the regulation cover implementations of technical solutions, it can be the manufacturer of the solution that is affected of the regulation, even though they are not responsible for the collection or processing of data in the system itself. Violating article 25 can give fines up to 10 million EUR or 2% of the total worldwide turnover (Voigt & Bussche, 2017 p.64).

2.2.5 Rights of the data subject

The data subject has always had a range of rights, but they are now increased due to the GDPR. Generally, it is to give individuals more control and understanding of what is done to the data that belongs to them. If rights are infringed, compensation can be requested from a data controller or a data processor (Calder, 2018, p.33).

Consent

One important addition of the GDPR is the consent. The data controller is obliged to demonstrate that the data subject has given consent for the processing of the personal data. Consent means a freely given and unambiguous indication of that the data subject agrees to the processing. The GDPR does not provide formal requirements for what is counted as consent, rather it is the controller who bears the burden of proof that a consent has been given. This becomes specifically relevant when the consent is given online (Voigt & Bussche, 2017 p.94). A consent is not counted as inactivity or a pre-ticked box. It is also important that the data subject has given consent to all of the processing that is going to be executed of the controller or the processor. This is a shift from opt-out to opt-in. Further, a consent can, due to the GDPR, always be withdrawn. This puts pressure on developers to build robust systems which can easily change from consent to a non-consent (Calder, 2018, p.34).

The right to access, to be rectified and to be forgotten

The controller is also obligated to be able to provide the data subject with a copy of their personal data and information about how it is processed and what third parties are involved when handling the data. The data subject has the right to access this information within a month after requesting it, with an extension of two months if necessary, with a valid reason given from the controller. If the information is requested electronically it should be received electronically (IT Governance Privacy team, 2017, p.193).

The data subject also has the right to rectification stated in Article 16 GDPR. This means that, with the burden of proof of inaccuracy or incompleteness of personal data, a data subject has the right to demand a rectification and complete incomplete data. It is a balance of interests between the data subject and the data controller to whether a rectification is necessary and that has to be solved from case to case (Voigt & Bussche, 2017, p.155).

Further, the right to be forgotten is a right making it possible for the data subject to demand erasure of personal data from the data controller, stated in Article 17. The data erasure can be valid when there is no purpose for processing data, even if it once where the case, as it become when a data subject withdraw consent. The erasure is also valid if an issue with the legality to process the data arises and there the controller fails to prove that there is a legal reason for processing the data (IT Governance Privacy team, 2017, p.195-196). With the right to be forgotten, the personal data that has reached other third parties or processors must also be erased, which can become complicated when the personal data is electronical. The GDPR states that an attempt of removing the data taking into account the available technology and costs take "reasonable steps" to erase the data (IT Governance Privacy team, 2017, p.197).

2.2.6 Reach of the GDPR

The scope of the GDPR is relatively broad even though it is initiated by the EU for mainly the member states. The GDPR is stated to apply to activities by a controller or a processor within the EU, irrespectively to if the processing of personal data is done within the EU or not. It also applies to where personal data of EU data subjects are processed, even though the controller or the processor might not be established in the EU themselves. The regulation applies regardless the need of payment from a data subject for a service or offering of goods. Another appliance of the GDPR is when a behaviour of a data subject is monitored and processed when the behaviour take place within the EU. In this case it also does not matter where the controller or processor are located (Finck, 2019, p.8-9).

2.3 Incentive Projects

In any virtual community it is possible to gather people with a shared goal and influence them into a certain behaviour. This influencing is done by using incentive mechanisms. The incentives can be of various kinds, both monetary and non-monetary, giving all kinds of rewards from payments, premiums or coupons to social recognition or performance appraisal. Incentive mechanisms are based on user's participation and encourage them to strive towards new objectives (de Melo Bezerra, 2015). Incentive mechanisms can also be used in blockchain-based platforms. Many projects aiming to motivate people to become more sustainable are using incentive mechanisms in combination with blockchain technology to reward their users. Sustainable choices, such as producing solar power (SolarCoin, 2019) or reducing CO2-emissions (EnergyCoin Foundation, 2019) is two examples of how tokens can be rewarded on a blockchain for each specific purpose. With the same logic as a national currency, tokens can later be spent on something else (Dapp, 2018). One project which could

have the potential of using blockchain to influence sustainable choices in everyday life is the SimpliCITY project.

2.3.1 The SimpliCITY project

In October 2018, the SimpliCITY project was initiated by the Salzburg Research Forschungsgesellschaft as a project to increase visibility of regional sustainable services (RSUS) by aggregating such services on a digital platform. As a part of developing methods and tools for making a community engaged and aware of a "sustainable city lifestyle", the project aims to enable international collaboration and knowledge transfer in scaling initiatives for smart cities. The project alludes to result in a proof-of-concept for a digital marketplace with aggregation of multiple RSUS divided in mainly three categories: bike mobility, social inclusion and local consumption. Enhanced with incentive and reward mechanisms the project hope to introduce a sustainable lifestyle for citizens (SimpliCITY, 2019). The platform will help with connecting cities, consumers, and producers when trying to reach long-term goals in sustainability (Salzburg Research, 2019).

SimpliCITY will connect a wide range of stakeholders. Currently the main project partners are the Municipality of Salzburg, Salzburg Institute for Regional Planning and Housing (SIR), Uppsala University, Uppsala Municipality, as well as the tech-company Polycular which is situated in Austria (Salzburg Research, 2019). The project is funded by Austrian Research Promotion Agency, Federal Ministry Republic of Austria Transport, Innovation and Technology, and Vinnova. Further, the target groups for the project outcome is city administration, service providers, association initiatives and of course, citizens, which are thought to be the main users of the platform (SimpliCITY, 2019).

Salzburg in Austria was chosen as a first pilot city. After advertising for another city with similar demographic variables, Uppsala in Sweden was chosen as the second city for pilot tests (Rosén, 2019). The plan is to have up to five more follower cities that will replicate the project in the extent suited the specific city's needs.

The Salzburg part of the project will build an initial platform which will be listing information about local services as well as start focusing on integrating actual services in the bike mobility category (Stabauer & Schrempf, 2019). A range of bike mobility services are in the pipeline for the city of Salzburg, including bike upcycling and earning tokens for making trips by bike rather than by car. The SimpliCITY platform is thought to use blockchain technology in creating a token system for incentive mechanisms to encourage citizens to become more sustainable. Citizens should be able to earn tokens for completing challenges or taking part of local services. The tokens can then be used for example to make their voice heard in the platform, in the form of a voting system on changes to be made in the city by the municipalities or getting discounts and deals at local service providers.

3. Method

In this chapter it will be described what research methods have been used to conduct this thesis. The choice of qualitative research methods is motivated in section 3.1, followed by a description on how the literature review was conducted in section 3.2. This is followed by a review of the respondents as well as the structure of interviews and how the interviews was later analysed in section 3.3.

3.1 Choice of Research Method

This study aims to explore the compatibility of blockchain technology and the GDPR. Corbin & Strauss (2015, p.5) describe how a qualitative method is suitable when researching areas that are not yet researched thoroughly and when there is a need for a holistic approach of the phenomena researched. The subject of blockchain technology and its compliance with the GDPR is relatively new, whereas the lack of volume in current research created a need for a qualitative approach. A literature review was chosen as the main method to explore the existing current research. The current existing technical solutions to handle personal data in a blockchain was also researched in a qualitative manner to discover variables, in this case technical implementations, which can later be tested by quantitative methods. This is also something Corbin & Strauss emphasizes as yet another reason to choose qualitative methods.

The aim for this study is also to explore different perspectives on the impact which the GDPR might have on blockchain development. Corbin & Strauss (2015, p.5) list investigating different perspectives as another reason to choose a qualitative research method. Interviews were chosen as the second method to complement results of the literature review, but also to explore the different perspectives of compliance. Qualitative research is regarded as an approach to find categorizations and build theory from collection and analysis of data. Bryman (2016) suggests an outline of six steps when conducting qualitative research, these can be seen in Figure 7.

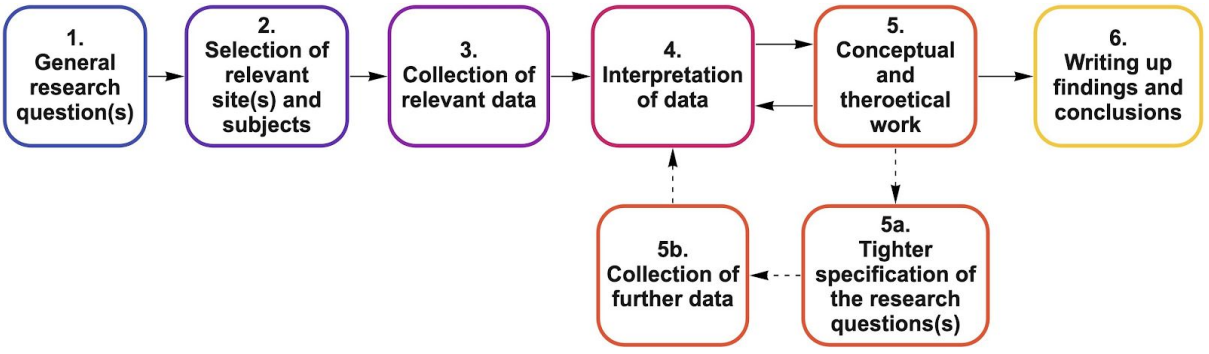

Figure 7. The suggested steps of qualitative research. Adapted from Bryman (2016, p.379).

Starting out with general research questions and selecting relevant subjects to study as well as collecting and interpreting data, help create the foundation for conceptualizing theories. Iterating the process by further collecting data, after re-specifying the research questions, create a process where the first general concepts are narrowed down and researched in a structured manner (Bryman, 2016, p. 378-381).

This study makes use of a broad literature study and complementary interviews. Interviews was conducted both with people within the area of blockchain as well as people with insights in the impacts of the GDPR. Initially the research questions had wider definitions. Concepts of compatibility emerged from practices and current knowledge of what could be found in literature along with the responses of the participating respondents. In line with the method of Bryman, these insights were iterated which later curated new research questions. Further data collection was based upon the new research questions. This process was repeated twice until the current research questions were reached and they were furthered explored to create the outcome of this thesis.

3.2 Literature Review

One type of literature review is the systematic review. Saunders, Lewis and Thornhill (2016) describe the systematic review as a process for reviewing literature by locating existing literature, evaluating its contribution and analysing the findings. This allow the reviewer to draw conclusions about what is and what is now known related to stated research questions. The systematic review is suitable to use when there is an uncertainty about the possible policies and their evidence, as well as when a general overall picture of the research evidence is lacking. The type of review also helps to create an accurate overview of the current research and associated methods which further can help develop new methods. Saunders, Lewis and Thornhill refer to some reviewers using the systematic review method to follow a few steps shown in Figure 8.

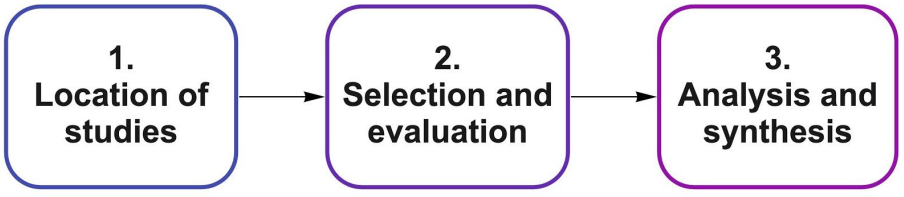

Figure 8. Steps of a systematic review. Adapted from Saunders, Lewis & Thornhill (2016, p.108-111).

3.2.1 Location of studies

Saunders et al. (2016) emphasize the importance of the review to be critical. To avoid the review becoming just a listed notation of previous work a critical mindset needs to be kept throughout the reviewing. To keep focus on what relevance the chosen literature brings and how it goes against or confirm other literature a thematic approach can be used. By choosing

themes that cross-over several authors work one can compare and contrast work of research and in such a way find interesting similarities and issues (Saunders, Lewis & Thornhill, 2016, p.80). To maintain a critical review, the literature researched for this thesis has been focusing on two different domains of literature. Firstly, literature with a technical focus has been researched to explain blockchain technology, find technical solutions which increases privacy and to review how the GDPR are interpreted from a technical point of view. The second domain is the legal, where literature on the GDPR itself as well as its implications on new technologies have been researched. Location of studies has been done through searches for relevant keywords, from recommendation by respondents or in conversation with possible respondents. Also, much literature has been found in references from previous literature.

3.2.2 Selection and evaluation

The selection of literature has been done with the aim of find as many perspectives as possible. Academic research was combined with written material of other sorts, such as reports, magazine articles, books, video-material from seminars and blog posts to generate an overview of the current situation. Each source was evaluated based on other references and relevancy to the subject at hand. Another aspect evaluated was the accuracy of the literature. Since the subject investigated is relatively new, it became important to retrieve as relevant material as possible and the selection of literature was based heavily on that notice.

3.2.3 Analysis and synthesis

Aligned with Saunders et. al., Booth et.al. (2016) discuss the matter of a systematic research review. Booth et. al. mean that a systematic research synthesis can help to answer research questions and evaluate different results while identifying gaps of knowledge that need to be further researched. It can also be evaluated whether the results are consistent or not throughout the literature. The reviewer should also keep evaluating and highlight weaknesses in the research found to keep the critical approach (Booth et al., 2016, p.11). For this thesis, each literary source has been interpreted as one perspective on a complex situation. Critical points of divergence or agreement in perceptions of the same situation have been in focus when researching and when generating the argumentation in the results.

The literature review should also aim to place work into context to see how it contributes to the field of study, find new ways to interpret previous research identify and resolve conflicts in previous studies. According to these steps the gaps existing can be identified and make path for what need to be researched further (Booth et al., 2016, p.14). The research questions in this thesis aim to lift the discussion of compliance and put what is said in literature in a context of the different perspectives of technology and legislation. This aim to bring further points to be researched, which would fulfill the criteria that Booth present to be necessary for the literature review.

3.3 Interviews

To complement the literature, qualitative interviews were chosen as a complementary research method. Taylor et al. (2016) state that in-depth interviewing is suitable when there is a relatively well-defined research area and a broad range of people or settings to be understood (Taylor et al., 2016, p.104-106). For this thesis, the research questions are well-defined, and the area of interest is limited to how the GDPR is affecting blockchain projects. People with different experiences needed to be interviewed, since the people with knowledge about the GDPR are not necessarily the same people knowledgeable of blockchain technology. Different insights and perspectives needed to be combined to generate the full picture and the research had to be done in a limited amount of time.

Interviews with involved parties of the SimpliCITY project were conducted to explore details of the project since it is used as a case study for this thesis. Second, interviews with technical respondents were conducted to get insights in the practical usage of blockchain technology in projects for various use cases. Thirdly, interviews with people with an insight in legislation and law was conducted to get further knowledge about the details and the case-by-case approach of the GDPR.

3.3.1 Choice of respondents

To find respondents suitable for this study, a few key people were identified. Using what Bryman & Bell (2016) calls Snowball sampling, more relevant respondents were found throughout the process of research. Snowball sampling implicates that a small group of participants are chosen, and these participants suggest other participants with relevant experience or knowledge (Bryman & Bell, 2016, p.415). The first respondents for this study were found through personal contacts, through the case study project representatives and through the literature review. These respondents led to further contact with relevant participants. 11 interviews were held in total. 4 concerning the SimpliCITY-project, 3 with technical experience, 2 with legal experience and 2 involved with an ongoing blockchain project. One additional respondent involved with an ongoing blockchain project did not have time for an interview but answered a few questions by email. The respondents for this thesis can be seen in Table 2 below:

Table 2. Respondents for interviews.

Respondent	Business/ Organisation	Interview form and location	Duration	Domain
Christoph Wögerbauer	Polycular	Skype	30 min	SimpliCITY/ Technical
Johan Rosén	Uppsala kommun, SimpliCITY	Telephone	45 min	SimpliCITY
Johan Rubbestad Lilja	Uppsala kommun, SimpliCITY	In-person, Uppsala	60 min	SimpliCITY
Petra Stabauer & Bernhard Schrempf	SimpliCITY	Skype	50 min	SimpliCITY
Okan Arabaci	IBM	In-person, Kista	60 min	Technical
Yonghui Jin	Ericsson	In-person, Kista	75 min	Technical
Daniel Olsson	Truesec	In-person, Stockholm	60 min	Technical
Brian Mulder	EnergyCoin	Skype	35 min	Blockchain project/ Technical
Anne-Marie Pronk	EnergyCoin	Skype	40 min	Blockchain project
Brigitte Devos	Buck-e	Email	-	Blockchain project
Andreas Kotsios	Uppsala University	In-person, Stockholm	120 min	Legal
Amelia Wallace & Ellinor Mörner	IT-advokaterna	In-person, Stockholm	80 min	Legal

The number of interviews chosen to conduct for this thesis was not a set goal from the beginning. Kvale & Brinkmann (2014, p.156) answer the question about how many interviews is needed for a qualitative study is as many needed to answer the research question and what the purpose of the study is. Since the aim of the interviews was to get different perspectives to the one situation, a variety of respondents was more important than the total

number. More than one person from each domain was desirable to be able to draw moderate general conclusions. Kvale & Brinkmann (2014, p.157) discuss the fact that if a few respondents enable a greater possibility of thoroughness and to find details of comparison not possible in the same way with a large group of respondents which was something suited to this thesis.

3.3.2 Semi-structured interviews

It is important for the interview setting to create a secure and free situation for the respondent. Magnusson & Marecek (2015, p.46) explain how the interview situation should encourage participants to discuss and narrate their thoughts in their own words and not be imposed by the interviewer. Before interviews with each domain, an interview guide was made which made in total four different guides which each can be found in Appendix A-D. Questions were added to each guide with the aim of discussing the respondents area of expertise or project as well as their view on data privacy and future of blockchain. Interview questions were formulated to get information which would answer the research questions but also to cover some extra ground for eventual alterations in them.

Working from an interview guide, the questions of a meaning-making interview should flow from one topic to another, so called semi-structured interviewing. In a semi-structured interview, the questions are open-ended to allow the respondent to answer rich and complex, as well as making room for the interviewer to adapt the discussions accordingly. Follow-up questions can be used to help participants fill out their answers (Magnusson & Marecek, 2015, p.46-54) and questions that are not in the interview guide can be included if the interviewer notice something important from the respondent (Bryman, 2016, p.468). With this in mind, several of the interviews covered topics not included in the interview guide. When a respondent drifted away from the interview guide the discussion could sometimes lead into details not previously known and was therefore relevant to discuss. If the sidetrack was not relevant, the respondent was asked a new question, since it is the role of the interviewer to keep guiding the respondent back to the original interview guide to cover the wanted topics (Magnusson & Marecek, 2015, p.62-63).

The interviews held for this thesis was, when possible, conducted in-person. When time or place was a restraint, interviews were held by telephone or skype. Before the two interviews with the legal respondents, a brief with information was sent out beforehand on request from the respondents. The brief contained a quick overview of the problem areas of the GDPR and blockchain technology together with questions relevant to discuss before the interview. This was done to get a head start on the discussion and prepare for discussing details rather than explaining the concepts of the regulation at the actual interview. For two of the skype interviews with representatives from EnergyCoin as well as for the respondent of Buck-e project, questions were sent out in beforehand. This was done because the interviews were held on skype with a shorter time frame than the rest of the interviews. Prepared questions made the interview more concise and time efficient.

The majority of the interviews held were recorded and transcribed to later be analysed. Transcribing interviews is a time-consuming process but is needed to be able to interpret from a person's own words and to be able to code the material in a later stage (Magnusson & Marecek, 2015, p.73). The interviews were transcribed shortly after they were held.

3.3.3 Data analysis

The transcribed interview was coded to be used in the results of this thesis. Coding interviews mean that the transcribed material is categorized based on keywords. Sections of the interviews are tied to a certain subject which is relevant to discuss and the point of view from the respondent is in this way portrayed as fairly as possible. This type of data analysis is commonly used in the Grounded Theory, which was introduced by Glaser & Strauss in 1967. In grounded theory, qualitative analysis is done to retrieve the experience of the respondent and develop new theories in an inductive way (Kvale & Brinkmann, 2014, p.242). This thesis aims to explore the borderland of law and technology with respect to the GDPR and blockchain and to investigate what might happen in the future. The grounded theory could help develop such a theory, whereas it became relevant to use this type of data analysis of the interview-material. When coding this type of material there is always a reducing experiences and perspectives into certain categories which Kvale & Brinkmann (2014, p.243) lift as a criticism of the analysis. However, reducing material into categories is also the factor which facilitate overviewing and comparing which is desirable for the purpose of this thesis.

4. Results

This chapter contain an overview of the compatibility issues of the GDPR and blockchain, going into detail on what could be personal data on a blockchain and how the situation of GDPR compliance differ for permissionless and permissioned ledgers in section 4.1. Next, in section 4.2, technical solutions to make the blockchain more GDPR compliant are discussed. This is followed by a review of the relationship of technology and regulation in section 4.3.

The first two sections of this chapter will have an analytical approach in investigating the research questions. Reviewing and analysing literature as well as explain concepts and solutions will be in focus. The third and last section will be focusing on the results from interviews to explain the current situation from a different angle. Together the two parts will make up the foundation for the following discussion in the next chapter.

4.1 The Compatibility of the GDPR and Blockchain

In July 2019, the Panel for the Future of Science and Technology (STOA) of the European Parliament produced a guide on the GDPR and blockchain technology. In the report, Dr Michèle Finck discusses in a broad perspective the compatibility and uncertainties with blockchain's compliance with the GDPR. The report is therefore suited to use as a guide for this section. Another important entity when it comes to the GDPR is the Commission nationale de l'informatique et des libertés (CNIL). Based in France, the commission presented a report in 2018, regarding the GDPR and blockchain. The report was one of the first to address the compliance in a broader sense (CNIL, 2018a). In addition, this section will discuss other articles and thesis' handling the subject, as well as include perceptions from interview respondents with technical knowledge and respondents with a background in law. The aim is to give as many perspectives as possible to the issue at hand.

Finck (2019) discusses a few general divergences of blockchain and the GDPR which can be seen in Figure 9. The first compliance issue which can be identified is how the GDPR has been initiated based on the idea that every storage of personal data has one dedicated data controller. Blockchain on the other hand is a distributed database, contributing to a decentralization of responsibility to all parties involved in the network. The second, according to Finck, is how the GDPR assumes data to be modifiable or erasable. Blockchain technology on the other hand, depend on the ledger to be immutable. Furthermore, the distribution of data on a shared ledger opposes the purpose limitation, data minimisation and storage limitation principles stated in the GDPR. All of these divergences need to be considered when constructing a blockchain solution. Hence, the design process executed by blockchain architects become very important (Finck, 2019, p.i). It is also in the initial design process it is possible to apply the Privacy by Design concept of the GDPR (CNIL, 2018b).

Figure 9. Factors making blockchain not compliant with the GDPR (based on Finck, 2019).

For the GDPR to be relevant, the data stored on a distributed ledger need to be considered as personal data. Therefore, a look into what data is personal data and where it would be possible for that data to end up in a blockchain, is needed.

4.1.1 Personal data in blockchain

In most blockchains there are two sections of data that could be considered as personal data: 1) The transactional data and 2) The metadata, such as addresses connected to the sender and the receiver, and a timestamp of the transaction (Bacon et al., 2017). If both of these are considered not to be personal data, all blockchains could be considered to be fully compliant with the GDPR. Is it possible to anonymise personal data so that it can be stored in compliance with the GDPR, if needed, as a proof of a transaction? And also, are addresses considered to be personal data?

Anonymisation of personal data

The GDPR discusses both pseudonymisation and anonymisation of data. Pseudonymisation is considered to be personal data, while anonymised data is considered to be of no concern of the GDPR. Processed personal data can be classified as anonymous if the likelihood that it will identify a person is removed. The identifiability then become the focus for the regulation. The GDPR take plausibility into consideration when evaluating the possibility of identifying a person based on processing personal data. The plausibility is dependent on the available technology and the cost of the activities needed to identify a person (Finck, 2019, p.19). Classifying personal data as anonymised is further complicated by the fact that predictions on what additional data can come to be available in the future can never be certain (Finck, 2019, p.20). On the other hand, Kotsios, who is a PhD in civil law at Uppsala University and one of the legal respondents, appoints that the legislation must evaluate only the current situation, not regulate for future uncertainties on data availability (2019).

It is hard to anonymise data, and it is yet unclear whether the GDPR will demand that personal data should be completely anonymised. The CNIL writes in its report that personal data should be stored on a blockchain in the format that least impact the freedom of the data subject and recommends to store personal data primarily as commitment schemes, for example as non-interactive zero-knowledge proofs (which will be discussed later), and secondly cryptographically hashed. However, the CNIL also recognizes that these measures might be questionable with the compliance of the GDPR in the future (CNIL, 2018b). The

question on anonymisation and pseudonymisation need further evaluation by authorities since they largely affect the situation for blockchain in relation to the GDPR. Uncertainties regarding anonymisation also become relevant when analysing the address used in transactions.

Addresses as identities

The question on whether a public key, or rather the address, which is the hashed value of a public key, can count as personal data and accordingly complicate that it is stored on a blockchain is an ongoing discussion. First and foremost, this is only relevant if the address is linked to a data subject (Kotsios, 2019). In the case of cryptocurrencies and tokenization systems, this is often the case. Since the address on a blockchain is linked to a wallet owned by a data subject it creates a link to the data subject which could be used to identify them. From a legal point of view this is seen as personal data:

> You must be able to identify someone to be able to send them a transaction. So, there will at least be elements of personal data in that blockchain (Wallace & Mörder, 2019).

From several interviews with technical respondents it shows that the general idea is that an irreversible address should not be classified as personal data (Arabaci, 2019; Jin, 2019). The argument is that since the data is irreversible the encrypted storage should be considered compliant with the GDPR. The CNIL discusses the subject in somewhat the same manner, pointing to the fact that an address is data that cannot be minimised further and is essential to the function of the blockchain (2018b) even if they do not state in clear text that a solution with a hashed public key is compliant with the GDPR. A similar conclusion can be shown in other studies, for example in Onik et al. (2019).

In contrast to this opinion and in the light of the earlier discussed identifiability, Finck (2019) raises the issue of incidents where an address has, in combination with other information, in fact become an identifier. Hence, an address should be seen as a pseudonymised data since it can be reversible, or at least help to reveal the identity of a person (p.27). Other studies support the view of first and foremost Bitcoin, but in general most blockchains handling addresses, to be pseudonymous and have issues with privacy (De Filippi, 2016; Genkin et al., 2018; Henry et al., 2018; Arabaci, 2019; Goldfeder et al., 2018).

There are several scenarios involving an address which could lead to identifying a person where the public address is used as fragmental information. Some of the scenarios found in the literature are:

1. Storing additional information about the users off-chain, for example in a ‘Know Your Customer’ (KYC) verification, risking the information to be combined with an address (Finck, 2019, p.27; Schmelz, 2018).

2. Several transactions made by the same address can reveal a pattern, leading to possible reveal of the person behind the address (Bacon et al., 2017; Finck, 2019, p.27; Zheng et al., 2017).
3. A non-secure connection between the blockchain and the user on a network level. Could lead to collection of network information such as IP-addresses, which can qualify as personal data in itself (Schmelz, 2018), or give away access patterns, which could reveal an identity of a user (Henry et al., 2018).
4. Identifying a node based on which other nodes it is connected to (Zheng et al., 2017).

To try to approach a case-by-case analysis when looking at the compatibility of the GDPR and blockchain, the next sections will go into detail of the two main types of blockchains: the permissionless and the permissioned. It does not matter if an address exists in a permissionless or permissioned blockchain, it counts as pseudonymised data in both cases (Finck, 2019, p.28) and could therefore be enough for the GDPR to be relevant for any blockchain solution. As a consequence, other properties of a blockchain will further complicate the compliance with the GDPR. Next, it will be discussed how the transparency, immutability and the handling of personal data, illustrated in Figure 10, become important if personal data is in fact a part of a blockchain.

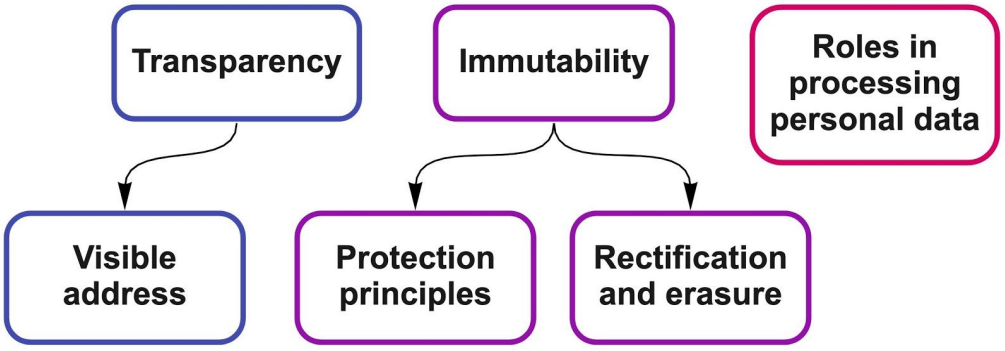

Figure 10. Problem areas with personal data in a blockchain.

4.1.2 Permissionless blockchains

Transparency

Even though it is still up for debate, some parties view the permissionless blockchain as the original and true blockchain due to its characteristics (Arabaci, 2019). For this type of blockchain, transparency is a key feature and no gatekeepers guard who is allowed to view or join the network, making the information stored available to anyone interested (Finck, 2019, p.5). According to the Privacy by Design concept, if personal data is stored in plain text on the blockchain it would not protect the privacy rights of data subjects and therefore would not be compliant with the GDPR (Wirth & Kolain, 2018). The CNIL formulates it as being strongly recommended not to store personal data in plain text on any type of blockchain (CNIL, 2018). Furthermore, so called blockexplorers, search engines for blockchains, can be

used to search information from a public ledger. This further decreases the privacy of a permissionless blockchain in a significant way (Finck, 2019, p.5).

Immutability

The trait of being immutable is often one of the key reasons to use a blockchain (Arabaci, 2019). In a permissionless blockchain, the immutability is secured based on consensus mechanisms involving all nodes in the network, making it very expensive to alter the blockchain. It also gets harder for each block added since more power is needed for every hash that need to be recalculated. Even if it is not considered to be practically feasible to alter a blockchain, it still is theoretically possible, which Finck (2019) highlights.[1]

The immutability trait of the permissionless blockchain become problematic in the eyes of the GDPR in a few ways which are discussed by Bacon et al (2017). One is the data subjects right to rectification and erasure of data. As stated, it is theoretically possible to make the whole network perform a 'hard fork' of the blockchain and in that way deviate from the data that would be desirable to change or erase. In practice, it would however be very hard to make a large number of nodes to agree to a shift. The method would demand too much of the network to be realistic, at least for a public blockchain. Another measure is to change any data by updating the information with a new transaction. This would however not remove the actual data from the ledger (Bacon et al., 2017).

The method of storing only the hash of the personal data and alter or erase the data off-chain is one way to try to comply with the right to rectification. This method makes the data inaccessible, which according to the CNIL is closer to the effects of data erasure (2018b). The term 'erasure' is by some parties interpreted in this context as something less literal. Kotsios (2019) lift how the Information Commissioner's Office in the UK have discussed deletion of personal data in a more subtle way as "putting data beyond use". Finck (2019) also discusses the erasure and lift the alternative of removing data off-chain and keeping the hash as a scenario that need to be evaluated according to the means which could potentially reveal an identity. Finck conclude that further regulatory guidance is needed (p.32) whereas the rectification and erasure of personal data is still an unsolved issue for permissionless blockchains.

It also matters in what purpose the personal data is stored. In a report from the project The land registry in the blockchain - implementation test, which started in Sweden during 2016 with the aim of putting land registry on a blockchain, the erasure of data is discussed regarding personal data which is handled by a government authority. In the reports it is stated

[1] To put the theory in a practical perspective, it was in November 2018 estimated to cost around $1.4 billion to create a 51% attack on Bitcoin, which is the largest cryptocurrency to date: https://cryptoslate.com/analysis-bitcoin-costs-1-4-billion-to-51-attack-consumes-as-much-electricity-as-morocco/

that the requirement of erasing data can be lifted if the data is handled with regard to the interest of the society, or as an exercise of authority (Kairos Future, 2018).

Another aspect of the immutability in combination with the distribution of data become a topic for discussion when looking at the protection principles 2) Purpose limitation, 3) Data minimisation and 5) Storage limitation of the GDPR. The outcome, however, depend on the interpretation of the regulation.

The purpose limitation requires personal data stored by an entity to be relevant and limited to the purpose for its storage. If that implication of further processing of personal data is included in the original purpose it might be considered to be alright to process personal data in a distributed manner, such as in a permissionless blockchain. However, even if this is stated in the original purpose, the problem could still arise with the GDPR allowing a data subject to revoke their original consent at any time (Finck, 2019, p.67). Consent is needed for any handling of personal data according to the GDPR. Although it is not a formal requirement to inform data subjects of how the processing of their data will be in a certain technical architecture it is mandatory to inform them about the risks of the processing. According to Finck it consequently become important to inform data subjects if using a blockchain when asking for consent (2019, p.64). It is also included in protection principle 1) Lawfulness, fairness and transparency, which entitle the data subject to the right to be informed of how their data will be processed and the processing must follow this description.

The data minimisation can be debated related to the interpretation of the requirement to the quantity of data or to the quality of data (Finck, 2019, p.68). If the regulation is based on the quantity of data, blockchains can become problematic due to the replications of data among the participating nodes. If based on the quality of data, it points to how no data, other than certain categories that are absolutely necessary, should be stored. The data should also be pseudonymised or anonymised when possible. Finck mean that the quality of data-approach is unlikely to be the case due to how Article 25(2) GDPR expresses how the obligation depend on the amount of data, extent of processing, period of storage and accessibility. This also is a case which need regulatory guidance to be addressed in the right way (Finck, 2019, p.68).

About the storage limitation principle, Finck raises the question of when data on a ledger become obsolete. It might be a possibility of interpreting data as necessary for subsequent transactions or as earlier discussed with the Swedish project, as necessary for the purpose which is the interest for the society. However, if it is not, the immutability of a permissionless blockchain might create a problem in the perspective of the GDPR (Finck, 2019, p.69).

All of these issues are also discussed in a summary from a workshop during an event in Amsterdam in 2018, concerning the GDPR and blockchain. Finck was one of the participants in the final workshop together with other speakers on the theme. The summary states that the

Terms of Service and Privacy Policy could include the user to waive their right of erasure or rectification. However, if this possibility would exist, several further questions could arise, regarding how much a user can waive from and how other situations could take advantage of this. The panel concludes that the subject should be addressed for each implementation, but a general code of conduct would be beneficial (Ferrari, 2018). In contrast, the legal respondents for this thesis is highly skeptical to the possibility of the right to waive due to the inconsistency of the data subject's right that would cause (Kotsios, 2019; Wallace & Mörner, 2019).

Data controller and data processor

Another aspect of the permissionless blockchain is the ownership of the ledger and the data stored. As discussed previously, the GDPR enforces that it must be stated which parties have access to which personal data and which of these parties can be considered to be controllers and processors. The controller is the entity deciding the purpose and the means of processing and the processor is the entity handling the data on behalf of the controller (section 2.2.2).

It is not sufficient to state the roles on paper and thereafter hold them accountable. Rather, it will have be evaluated due to which activities linked to personal data an entity has been involved with (Finck, 2019, p.58; Kotsios, 2019). The CNIL however, points out that the role of data controller should be identified and appointed, at least in the case of having possible joint controllers. The reason would be that the data subject should know where to turn to when wanting to issue their rights (CNIL, 2018b).

In a permissionless blockchain where the stored data is distributed, and the management of the ledger is decentralized it become relevant to look closer at what entities decide the means and the purpose of handling personal data. Bacon et al. (2017) suggest that one can look at the situation of means and purpose on either a macro-level or a micro-level. On a macro-level, the means become the software and hardware used in the blockchain. Nodes and miners participate in the chain for the purpose of facilitating the system which would make them controllers. On a micro-level, each user decides the purpose of the data sent over the chain and choose to use the blockchain as the means. Other nodes and miners only facilitate access to the chain which make them likely not to be controllers. Bacon et al. view the micro-level perspective to be most appropriate due to the GDPR's aim of protect specific processing of personal data (p.41-42). On the contrast, a data subject using a blockchain do not have control over how long the data will be stored or when (or if) it will be deleted. Despite this, the consensus in literature can be viewed as the user of a blockchain to be seen as at least a joint controller (Finck, 2019, p.49). The CNIL has another interpretation, more following the macro-level perspective. It suggests that participants, i.e. nodes which are writing and sending data on the chain, can be considered to be controllers. The participants need to be a natural person and the data processed need to be for a non-personal activity or the participant is a legal person handling personal data on chain. In addition, the CNIL states that miners and users using the blockchain for personal matters would not be considered as

controllers (CNIL, 2018b). Furthermore, any entity gathering data from a blockchain and process it further would be classified as a controller (Bacon et al., 2017, p.45).

Because of the design of a distributed ledger and the current state of the law it is cumbersome to determine both which entities are controllers, but also which entities should be seen as processors (Finck, 2019, p.55). Data processors handle data under the instruction of a controller. Users of a blockchain can be seen as both controllers and processors due to storing the ledger on their own computers. The role of processor also depends on the implementation and entities surrounding the blockchain and should be solved on a case-by-case basis (Finck, 2019, p.57). Based on the micro-level perspective from Bacon et al. (2017), nodes and miners only process data on demand from a user making a transaction. They could however, if taking a more active role in handling transactions, also be considered as controllers (p.45).

Problems arising for the permissionless blockchain are not completely solved if instead using a permissioned blockchain. Some discussions are still valid but there are a few alternative prerequisites which will be discussed further.

4.1.3 Permissioned blockchains

Transparency

A permissioned blockchain is not open to read as a permissionless blockchain which make the permissioned blockchain different. It is commonly used as a solution internally in a company or in a joint venture for multiple companies which need a common solution (Finck, 2019, p.5), especially when handling data not open to be viewed by other than the appointed parties, which is the case for the technical respondent Jin (2019). A consortium is chosen and take the role as gatekeepers (Finck, 2019, p.5). When looking at privacy, the CNIL recommends to always evaluate to use a permissioned blockchain in favour of a permissionless when designing a solution, since it is easier to safeguard the data involved and govern where it would be located and spread (CNIL, 2018b). Overall, transparency is generally easier to control with a permissioned ledger. If not all parties involved in the permissioned blockchain have the same access rights, transparency can be limited. The limited transparency then asks for a higher degree of trust from the parties with less influence (Bacon et al., 2017).

Immutability

Since permissioned blockchains only have known participants, requests to alter or erase data can be handled in a more ordered manner than in a permissionless blockchain. The parties involved can pursue rectification or erasure of data by jointly re-hash the blockchain (Finck, 2019, p.73). They can also decide to delete the blockchain entirely if every party involved also delete their local replica of the shared ledger. Jin (2019) emphasizes that deletion of any once shared ledger should be obligated and stated in an agreement before initiating a shared ledger between parties.

Data controller and data processor

In the case of a permissioned blockchain it is overall easier to interpret who is data controller and data processor (Finck, 2019, p.52). The one entity, or the group of entities, involved in creating the blockchain can generally be seen as the controller or joint controllers (Finck, 2019, p.44-45). Bacon et al. (2017) discuss the macro- and micro-level perspectives on a permissioned ledger with regard to the entities involved being controllers in the macro-level perspective, both when setting up the platform as well as when they utilise it. In a micro-level perspective, with regard to the single transaction of data, the participants might be regarded as controllers when making a transaction but merely as a processor when acting like a node or miner in the private network (p.43).

4.2 Making Blockchains More GDPR Compliant

Despite how the transparency and immutability of a blockchain is in contrary to the directions given in the GDPR, there are ways which have been developed to make blockchains more GDPR compliant. The mitigations discussed further can be seen as 1) design choices when creating a new blockchain or 2) as mitigation principles when dealing with an already existing blockchain. An overview of the solutions discussed can be seen in Figure 11 below.

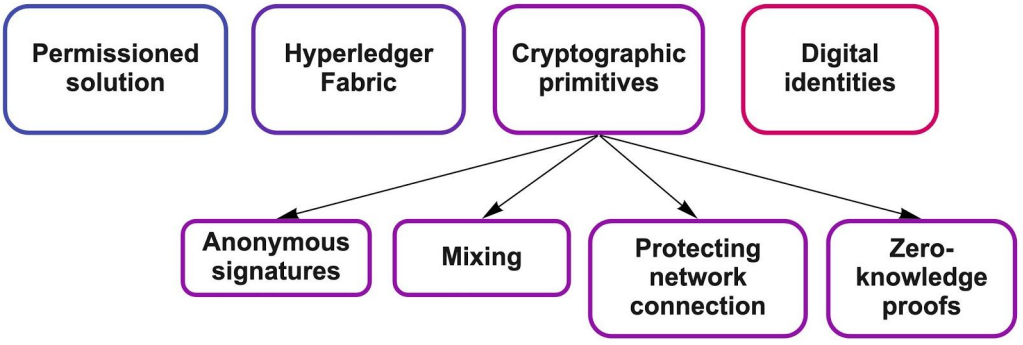

Figure 11. Technical solutions to enhance GDPR compliance.

4.2.1 A permissioned solution

A key point in the borderland of the GDPR and blockchain technology are the permissioned blockchains. The CNIL discusses how permissioned blockchains should be favoured to resolve the issues of transparency, immutability and roles that come with the GDPR. This is also the suggested design in a study made by Onik et al. (2019). The authors present a management system based on the roles of user, controller and processor, all separate nodes in a private network. The controller divides the data from the user and hash the data which could be classified as personal data and put it into the blockchain as a hash. The actual data is stored on local databases connected to each node. Rectification and removal of data is then enabled by creating an updated hash on the chain in a consensus mechanism among the nodes participating. The authors conclude how hashes kept on the chain is then not viable for the GDPR. The changes can be verified by cross-checking the hashes by each node, most importantly by the user node. This causes the system to be transparent in how data is used

among the nodes and a user can also easily file a claim for compensation through a smart contract if data not handled correctly (Onik et al., 2019).

The key for compliance with this solution would be that it is a permissioned chain with known participants. However, as previously noted when discussing permissioned blockchains: the difficulties of rectification and erasure of data is still not uncomplicated (Bacon et al., 2017, p.24). Also, if the ledger would become public or misused, a compliance issue with the GDPR could be prevailing.

When talking about permissionless blockchains, Bitcoin and Ethereum are usually mentioned. Ethereum is a programmable blockchain, meaning applications can be built on top of it, launched in 2015 with a native cryptocurrency. The default version of them both is the cryptocurrency which is open and public. However, both of them give the opportunity of using their software to set up a permissioned ledger and run it among known nodes. If using a permissioned version, the local cryptocurrency on the permissioned base cannot be used equivalent to the public cryptocurrency (Lewis, 2016) but it would enable more control over the data processed on the blockchain.

From the technical interviews it is evident that a permissioned blockchain would be the way to go to build a GDPR compliant system, but also for the sake of the use case. Jin (2019) describe:

> A permissioned blockchain is the same as a permissionless but included that not everyone has access to it. The aim is that it should not be open to all. (...) And we cannot disclose any data to any outsider.

Arabaci (2019) lifts the fact that a permissioned blockchain is useful because members can be excluded if they do something wrong. He and Jin also points to the platform Hyperledger Fabric which can offer some features helping to make a blockchain more GDPR compliant.

4.2.2 Channels and private data collections in Hyperledger Fabric

Hyperledger Fabric, established under the Linux Foundation, is a blockchain platform offering a permissioned blockchain setup which can be used in a range of industries and use cases. The blockchain can be modified with custom consensus mechanisms, smart contracts in general-purpose programming languages and do not require a native cryptocurrency to perform transactions (Hyperledger Fabric Docs, 2019a)

In Hyperledger Fabric, a blockchain can be shared among known parties. The ability to set up private channels, a network overlay, enable only chosen parties to share a path of information exclusive to them, even when working on shared network (Hyperledger Docs, 2019b). To go one step further, one can choose to use a private data collection within a channel. Creating multiple channels might create administrative overhead and private data can instead be used

to separate which parties can view which data. A private state database, a SideDB, is set up which need authorization to be viewed. Authorized peers share the SideDB while a hash of the private data is shared on the ledger whenever a transaction is made. The shared ledger is visible to all peers of the channel, but the private data collection is only visible to those authorized to the private data collection (Hyperledger Docs, 2019c). The structure of this architecture can be seen in Figure 12.

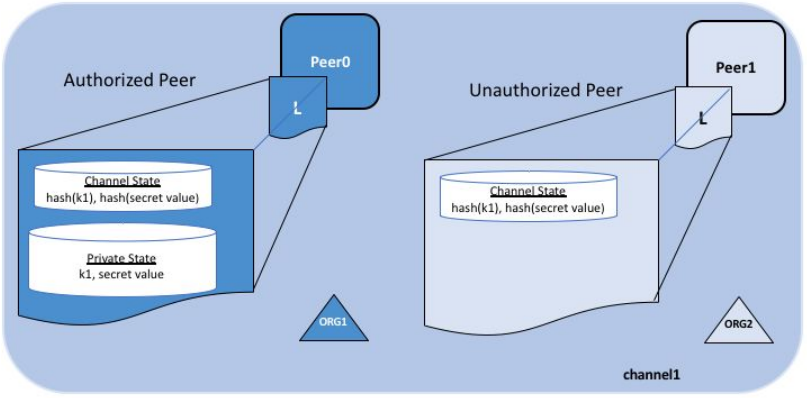

Figure 12. Peers view different data in the same channel enabled by personal data collections. Adapted from Hyperledger Docs (2019c).

Private data collections also include the opportunity to erase sensitive data stored on the SideDB, keeping only the hash stored on the ledger as evidence of the once existing data. The data can be erased either periodically, or if it has not been updated in a certain number of blocks (Hyperledger Docs, 2019c). There is also a possibility to keep the data just until a business process is done (Hyperledger Docs, 2019c).

4.2.3 Implementation of cryptographic primitives

Using a permissioned blockchain is one way to create a more GDPR compliant blockchain solution. However, when using any type of blockchain solution it might be desirable to protect the users even further. As earlier explained, several cryptographic building blocks are used to maintain trust and verification for blockchains such as hashes and asymmetric digital signatures. To avoid issues with pseudonymity, like in Bitcoin, further cryptographic primitives can be, and has been, incorporated in alternative cryptocurrencies and other types of blockchain solutions to enable privacy in further extent (Zhang et al., 2019). Since the main issue for the GDPR is the identifiability, implementations which could have an impact of protecting the address, and therefore lower the identifiability of a user, will be discussed in this section. Some cryptographic primitives concerning the privacy of a user are anonymous signatures and zero-knowledge proofs. These are not unique to blockchains but can be implemented in blockchain solutions. Another popular addition to a blockchain is mixing which help with the trace an address is leaving on a public blockchain. The last subject of discussion is technical solutions to secure the network connection which was one of the discussed issues with identifiability in a blockchain.

Anonymous signatures

Anonymous signatures provide anonymity for any entity using a digital signature to sign data, which is usually done by the sender in a transaction on a blockchain. One type of anonymous signature is the ring signature, proposed in 2001. A ring signature suggests that instead of signing a message by only one user, a group of users are included in a signature. In that way, the unique user who signs the message cannot be distinguished but just which group who signed. A group can be formed of any constellation of users which enable ring signatures to be used in all sorts of blockchain setups. The drawback of using a ring signature is that the user is anonymous and cannot be revealed even when it would be necessary (Zhang et al., 2019). This could instead be an advantage for malicious nodes (Wang, L. et al., 2019). Cryptocurrencies built on the CryptoNote protocol, for example Monero, which are using ring signatures to ensure privacy for its users, have this anonymity. It is also possible to include it for Ethereum since 2015 (Zhang, 2019). Ring-Coin is another example of a cryptocurrency using ring signatures (Wang, L. et al., 2019).

Group signature is a similar type of signature where one user can sign the message with a group's joint signature and hence hide their own identity. Although, for the group signature, there is one appointed group manager which could reveal the identity of the user if it behaves like a byzantine node. This manager-dependent system is therefore in need of a permissioned blockchain (Zhang et al., 2019) where the number of parties are limited.

Zero-knowledge proofs

Another solution to disguise a user's address is by using zero-knowledge proofs (ZKP). ZKPs was introduced in 1985 and was then a completely new concept in cryptography. The concept enables a way to prove correctness of a statement without disclosing any extra information (Wu & Wang, 2014). A type of non-interactive ZKP are the Zero-Knowledge Succinct Non-Interactive Argument of Knowledge, also called a zk-SNARK. A zk-SNARK enables verification of a transaction while protecting the privacy of both receiver and sender (Zhang et al., 2019). In short, a function C, which takes two inputs, x and y, return either true or false. X is a public input which is known, and y is secret. The prover, which want to be verified without disclosing any information, must, given a specific input x, give y such that C(x,y) is true. In the actual algorithm a secret variable is used, which if leaked, create a possibility of faking the proof. Hence it can be complicated to implement zk-SNARKs in a blockchain solution (Lundkvist, 2017). One project which have succeeded to incorporate zk-SNARKS in a cryptocurrency is Zcash.

Zcash is based on the Zerocash protocol, an extension of Bitcoin (De Filippi, 2016), where the zk-SNARKs are encoded into a few of the consensus rules (Zcash.com, 2019). It is also possible to use zk-SNARKs on the Ethereum blockchain, something EY were first to launch in 2018 (EY, 2018). L. Wang et al. (2019) lift the zk-SNARKs as the only cryptographic primitive to fully secure both the sender and the receiver's privacy while also hiding the content of the transaction. They do point out that using zk-SNARKs is lowering the

efficiency of the blockchain. There are also examples of zk-SNARKs being used in Hyperledger Fabric (Lavrenov, 2019).

Another implementation using zero-knowledge proofs is Hawk, a framework for developing smart contracts which are privacy preserving. Hawk protects the amounts included in the transaction from the public and also protect the sender and receiver from each other. The party's transaction goes through Hawk which uses zero-knowledge proofs to verify the correctness to the blockchain. The only party needed to be trusted is the so-called manager which is an in beforehand appointed node. The manager will however be penalized if not acting according to protocol (Kosba et al., 2016).

Mixing

Cryptocurrencies like Zerocash and Monero are incorporating privacy from the start. But if using Bitcoin or something similar, there are ways to improve on privacy using services like mixing, sometimes called tumbling. One such solution was CoinJoin which in a centralized route mix transactions. In this way the trace from sender to receiver in a transaction is concealed from viewers of the blockchain. The downside with such a service is the single point of failure of CoinJoin itself, since the disguised information is still visible to them. Several improvements with a penalty for the third party if disclosing any information has been tried in services like Mixcoin or TumbleBit. TumbleBit, which was the final evolvement, is however decreasing the efficiency of transactions in a large extent. Other solutions exist, for example CoinShuffle++ which make the parties perform multiparty computations to hide their traces, eliminating the intermediary support (Henry et al., 2018). Another option is confidential transactions, masking the amount of transactions using a cryptographic primitive called additively homomorphic commitments. This solution was introduced by Blockstream for the blockchain Elements and can be used in combination with the mixing services to improve on privacy (De Filippi, 2016). When using any of these services, a certain amount of trust needs to be put into the service itself since it is an additional risk for information to leak (Zhang et al., 2019). Yet another, more manual alternative, is for users to generate a new key, and therefore a new address, for each new transaction as a way to mimic the mixing-services (Bacon et al., 2017).

Protecting the network connection

There is however another issue not resolved by the previously reviewed methods of disguising the address. The possibility of linking an IP-address to an address in a blockchain (Bacon et al., 2017) is a situation which could potentially lead to identifying a user. This risk is something certain cryptocurrencies, for example Zcash, assume that users use to protect themselves. Services exist to provide this type of protection, for example Tor (Henry et al., 2018). Tor is a proxy, enabling routing to hide the address of the user on the network level. However, the implementation of Tor when using, for example Bitcoin (which Zcash is in extension built upon) can cause a target for attacks on the interaction between Tor and the blockchain itself (Genkin et al., 2018).

4.2.4 Digital identities enabling Privacy by Design

Wirth & Kolain has in a study from 2018 looked into how a technical solution should be designed if following the GDPR regulation as a base requirement. The authors make use of the Privacy by Design concept from the GDPR to create an architectural blueprint for a compliant blockchain solution. The solution focuses on how to enable the data subject to consent to each handling of their personal data, with a structure where personal data is stored off-chain with a hash pointer to it in the blockchain (Wirth & Kolain, 2018). This type of implementation is one way of creating a digital identity.

A digital identity is a way to gather all credentials for a user to identify themselves into one digital solution instead of getting an identification certificate from an outside party, for example a bank providing an id-card. With cryptographic signatures it is possible to verify that the identification credential used is authentic and with a self-sovereign identity (SSI) a user could generate and control unique identifiers themselves. The user can then allow parties interested in verifying the user to view and process a credential in order to authorize themselves. The SSI cannot be taken away from a user by any authority, making it a unique and revolutionizing way to handle personal data. A decentralized identifier (DID) is needed to setup the cryptographic public and secret part of the digital identity and the DID would be created by the user. This give the user freedom but also a responsibility on securing their digital identity (Lyons et al., 2019). Blockchain technology is thought to be helping with constructing an SSI in multiple ways and several blockchain platforms exist providing a foundation for developing an identity management framework, for example Blockcert or Hyperledger Indy (Soltani et al., 2018). Blockchain is thought to provide support for creating and store DID's, manage access control and consent to data use as well as notaries credentials (Lyons et al., 2019).

Wirth & Kolain's design is an example some of these properties. A third party wanting to use personal data can in the proposed design request the data using a smart contract. The data subject will be notified and can consent to the third-party getting access to their data while at the same time update their personal data with a new timestamp, causing the hash of the data to change. This cause the third party to have a unique set of the personal data which can be verified to be rightfully used if questions arise. This process with updating the personal data is also done through a smart contract. In addition, a new key pair is also generated with each request. This protect the personal data even if one key pair would be compromised (Wirth & Kolain, 2018). Wirth & Kolain (2018) further discuss the interpretation possibility of data on a blockchain (mainly a hash or an address) to be pseudonymous and the proposed blueprint from Wirth & Kolain do not address the uncertainty with the joint controllership or how the right to rectification and erasure would be handled. The authors point to using a permissioned solution where it is possible to resolve these issues (2018).

Olsson (2019), a senior architect for digital solutions at Truesec, also discusses digital identity based on blockchain as a solution to separate identity management from banks:

> Today we are fortunate in this country. We can use the BankID which work great. But it is a bank, a bank organisation who owns the ID. I do not like that idea. (...) We should not have a bank-ID but rather a Sweden-ID. (...) But that lead us into these kinds of questions. Ok, it is an ID. Is blockchain the best way to protect that ID? I think we could do it.

"But it is still a personal data. It will still be possible to connect it" said Mörner when the subject was briefly discussed with IT-advokaterna. Further clarification on how a digital identity data would be considered in the eyes of the GDPR is needed as well further evaluation of how other technical solutions can mitigate the possibility of being in the scope of the GDPR (Lyons et al., 2019).

4.3 The Relationship of Technology and Legislation

In the previous sections it can be seen that there are still a lot of uncertainties when it comes to the GDPR and blockchain technology. Technical solutions trying to settle some of them are in the second half of 2019, still evolving. However, the regulation has not been implemented for long and there is little legal assessment to use as guidance when wanting to construct a GDPR compliant blockchain solution. In this section, the perspectives on blockchain from a technical and legal point of view will be discussed, with the aim of trying to find indications on how the prerequisites surrounding the innovation of blockchain will affect the technology itself.

4.3.1 Technology need to adapt to slow-paced regulation

The tension between law and technology can be described as originating from different pace of evolvement. Several interviewed respondents mention how technology is developing in a fast pace while regulations and legal processes are slow evolving. This put pressure on regulation to keep up with technological innovations at the same time as innovation and new technical solutions are depending on the interpretation of regulations. For the sake of the GDPR this implies that the regulation as it is today will stay intact for a foreseeable future and will be something that blockchain projects will have to take into consideration.

> The data protection directive was from 1995, so it took 20 years to change that. (...) A major regulation which has that much work put in, (...) you would never change it in just 5 years or so. You will only interpret it in different ways (Kotsios, 2019).

Finck (2019) agrees that the GDPR will be kept in its current version, although increased legal certainty and further guidelines are needed. Blockchain put some of the central concepts of the GDPR to the test, which Finck emphasizes as necessary for the evolvement of the GDPR. Since it is supposed to be technology neutral it must be able to withstand challenges

from a more fast-paced data-economy (Finck, 2019, p.iv). And the technology neutral trait of the GDPR is necessary, or the regulations would become obsolete in 2-3 years (Kotsios, 2019). The current version of the GDPR is here to stay but regulatory guidelines are needed to avoid the current status quo between law and technology (Finck, 2019, p.iv).

4.3.2 Consequences of a status quo

The status quo situation is something which junior Associates at IT-advokaterna, Mörner and Wallace (2019), also discuss. They agree on technology being fast-moving and regulation coming from a slow-evolving process. Wallace lifts how the authorities might, due to this difference of pace, lack knowledge of the technology itself and therefore put off taking a stand or look further into the uncertainties. Mörner adds that it might be other aspects of the GDPR that have been required to look into before the specific technology of blockchain. In an email correspondence made with The Data Protection Authority (DPA) of Sweden it is confirmed that there are currently no specific guidelines of the GDPR when it comes to blockchain projects. Instead, they refer to follow the GDPR in general (Hallström, 2019).

The current status quo situation could however lead to additional problems in the long run, Mörner says:

> I think the risk which arise by postponing dealing with [the status quo] is that they panic in the end and make everything illegal. (...) Because [technology] moves too fast and they feel out of control. And that is because they have not familiarized themselves with it enough.

Coming from a legal background Wallace reflects on how the authorities need to take the first steps to resolve this: “A first step could be to actually look into this situation a bit more. (...) How can we build the road ahead to avoid companies to feel restrained because of the law?”

And companies do feel restrained by the law. In a report from PwC from 2018, 600 executives were interviewed about their organisation’s involvement in blockchain technology. 48 % said regulatory uncertainty was their main barrier to future adaptation of blockchain solutions (PwC, 2018). The situation indicated by the 48 % is also coherent to the stories from the respondents involved in blockchain projects. “For example, why did solutions like a permissioned, private blockchain appear? It goes against the original idea [of blockchain]. It is because there are regulations” (Jin, 2019).

Pronk (2019), one of the board members for the cryptocurrency EnergyCoin, hopes that authorities will keep an open mind when evaluating blockchain-initiatives: “I really hope that the municipalities are not afraid and do not want to regulate everything. Let us show how this works, what people can do with [the cryptocurrency], and then regulate things.”

Examples of how technology have had to adapt to legislation before can be seen in the many technical solutions aiming to make a blockchain more GDPR compliant. Personal data is something hard to handle in blockchain projects. The incentive project Buck-e, using a blockchain to reward, claim they had to change a lot of their software due to the GDPR since they collect personal data of their users (Devos, 2019). Mulder (2019), another one of the board members of the cryptocurrency EnergyCoin, describe a project of which he knows which have a hard time to launch since their blockchain solution involve sensitive health data which is hard to handle. Arabaci (2019) encountered the problem when choosing a project involving a blockchain for his master thesis: "We was not completely sure if it would work or not. But later we realized that alright, some of this data is too sensitive for [a blockchain]."

4.3.3 Test environments

Consequently, it is an unanimous opinion from both the technical and legal point of view that further certainty about the GDPR is needed, and that the authorities are the more critical part in providing this. At the same time, Wallace welcomes initiatives from both authorities as well as from technical solutions meaning it could provide some testing ground for the regulation to spark from. One method for moving forward is to provide testing environments, where companies could try their technical solutions under the watch of authorities. Olsson (2019) agrees that possibility to test technical solutions is highly important and suggests that projects not involving direct personal data should be a first test implementation. The earlier discussed project of putting land registry on a blockchain with the Swedish authority Lantmäteriet in the lead, Olsson sees as a success story for testing the technology under controlled forms. Mörner & Wallace also applauds the same initiative with the motivation that these types of tests are necessary for the future of blockchain as well as for the GDPR. However, Mörner says the risk of developing solutions within test environment might create an artificial safety zone, ignoring problems with regulation which could become a problem later in the process:

> If you cannot go outside of the created environment, then what does it matter? (...) You might create incredible things, but outside of the test environment everything might be regulatory uncertain. I believe you have to keep moving forward. (...) Dare to push it a bit (Mörner, 2019)

4.3.4 The trust of blockchain

Blockchain solutions in general are questioned, both in research but also by the legal respondents. Concerns exists on how the value of blockchains will be enough to excel over the possible issues with, for example, privacy.

> What is it that you don't trust? Is it other actors? And if you change something, what are the consequences? (...) If you have [a blockchain solution] you have to accept some risks. And evaluate, is it worth it? Is it that much better of a solution? (...) But like I said, I come from another background (Kotsios, 2019).

The technical respondents are in general in favour of blockchain solutions, to the right use case. It is often the immutability and the enforced trust which they emphasize as something lacking in other solutions. Vitalik Buterin describes, in the blog post "On Public and Private Blockchains" (2015), how the benefit of having "no authority to do so", referring to the inability to change the history of the transactions, can build valuable trust for a system. If even the system developers are excluded from the right to alter the blockchain, the security is established as permanent. Mulder agrees and discusses how it might look in reality:

> [Employees] do manual work to the database, they can change records by coincidence or intentionally. If someone from the marketing side does not like the numbers, well change the numbers. (...) I'm a technician but I don't trust technical solutions like they are today. Blockchain is at least one way to solve that problem of trust. Without you having to trust the other side, but basically through the technology (Mulder, 2019).

4.3.5 The future for the GDPR and blockchain

Mörner and Wallace expect that a convergence in regulatory guidelines will eventually develop among the EU member states, based on sentences and guidelines from each national authority. Kotsios believes that the outcome of interpreting the GDPR in relation to blockchain is dependent on the legal approach to the technology itself:

> The question is always: what is the goal with a specific policy? If you believe that [blockchain] will be important and become the new internet (...), then they will promote it. And interpret it in favor of [technology]. Otherwise they will say that it does not work. If you cannot erase data - that's it. It is like any other regulation, for better or worse (Kotsios, 2019).

From a technical point of view, it is suggested that blockchain technology will prove itself to be a greater solution than many others when looking at privacy, at least when it comes to the more controllable permissioned blockchain. Olsson, who has a background in working with IT in the public sector, believes that blockchain can indeed provide security and privacy in greater extent than current practices do. He emphasizes how it is needed to be understood from a legal perspective that private blockchains divide the risk of misconduct of collected personal data. The current common solution of centralized storage of personal data is not an optimal situation:

> [We have seen that before]. To have all data in a vault and then give the key to the enemy, it is not a good solution. Would you choose to use blockchain technology, distributed between multiple municipalities or collaborations, then the risk would be shared, and it would be even harder to steal the shared secret (Olsson, 2019).

Arabaci is also more confident that permissioned blockchains will be used in the future:

> It takes quite a lot for blockchain to be worth its while. (...) It will be very valuable in situations where there is a need or when it is feasible to use blockchain. But the phenomenon of cryptocurrencies (...), if it happens, it is still a long way to go (Arabaci, 2019)

What is shown in these answers is how the permissionless blockchain is viewed as a more unsure case for future development in relation to the GDPR. And a majority of the technical respondents testify that it is important to consider several aspects before implementing a blockchain solution at all, the GDPR being just one of many.

> People are thinking very traditional about technology. And they see [blockchain] as something new, something great. But if you really want to get the value out of it I think you should start (...) thinking about how it affects the traditional structures and this traditional thinking about data. About integrity, about privacy. What does it solve? And what does in not solve? (Mulder, 2019).

Another answer to the GDPR compliant blockchain might be the digital identity. Olsson, which have a technical background, believes digital identities based on blockchain could give great value for example with handling personal medical records. The ability to give permission to view data and then be able to retract it, Olsson means would be a much more private way to handle individual's personal data, than how it is handled today out in reality. But he also sees the difficulty of the processing around the technology:

> It will be a great challenge to get everyone on board in how the process will look like. (...) If someone give their consent to be in this blockchain with my identity, how do they assure that an exit is possible? (...) And maybe it will not be possible of a full data erasure. They will not be searchable but might be purged in there in some way. (...) How do they assure that it is only them which can verify themselves and their identities not to be misused? (...) I believe that if you want this to work, there must be an immense amount of trust to the entities maintaining the system (Olsson, 2019).

For some use cases, especially where a blockchain is used to tokenize actions and choices, personal data might not be necessary to collect at all.

> [The blockchain solution we are developing] is only a layer to transact or transform value. (...) You have the savings, you transform it into something else and if you look at that, then what do we need? Only where its coming from, how much, and then you do the calculations. (...) So, there is nothing like a birthday needed, or an address or whatever would point you to a person (Mulder, 2019)

Kotsios, even though he come from a legal perspective, is thinking in the same way:

> You can get output in ways you could not get before. If you think about that we in a few years have collected such large amounts of data already, it might be possible to (...) make personalized adverts, for example, without using any personal data. There will be AI-solutions creating a bunch of scenarios which will be able to give you what you want from just something like your name. (...) There are many talking about personal data being general data. It will be something we won't need. But you know, no one knows for certain what will happen in the future (Kotsios, 2019).

5. Discussion

In this chapter, the findings of the results are discussed. First, in section 5.1, a brief explanation of the main issue of the GDPR and blockchain are discussed. Next, in section 5.2, permissioned blockchains and digital identities are evaluated in relation to GDPR compliance. This is followed by a discussion in section 5.3, on what the differences in perspectives might imply for the future of blockchain. In section 5.4 an overview of what other possible problems blockchain technology is facing and what further research could be interesting.

5.1 The Non-Compliance of a Blockchain

Even though blockchain and the GDPR are sometimes described as having similar goals in enhancing security and help individuals taking control over their personal data, the uncertainties of their compatibility are several. If something classifying as personal data end up in a blockchain used for transactions where at least one party is a data subject, questions like how the data is processed, by whom it is processed, and how it will be possible to rectify or erase the data remain in general terms uncertain.

As long as no personal data is stored in a blockchain one can be certain to have a GDPR compliant solution, and no further work need to be done. The main reason for blockchains to even fall under the GDPR is the link to a person, through some sort of service needing to verify that a person is in possession of an asset, in most cases some kind of wallet. To be able to carry out transactions, the parties involved will have to be verified in some way. The question then narrows down to how that link can be disguised enough to be outside of reasonable means to be identified but still allow a transaction from a non-byzantine user.
How can an address move from pseudonymous data to as close to anonymous data as possible? For any initiative using transactions made by an individual user, this uncertainty will need to be taken into consideration. For blockchains with a connected digital wallet to be able to be designed, guidelines concerning the GDPR applicable for these types of blockchains need to be generated. Otherwise the uncertain situation will keep projects to be held back and hesitate developing solutions on digital problems using blockchain. Since the reach of the GDPR is wide it affects not only initiatives backed up by big companies but also smaller projects with a non-profit agenda like SimpliCITY.

5.2 The “More GDPR Compliant” Blockchain

The technical solutions lifted in this thesis are the ones most commonly discussed today and are generally focusing on this one problem. Currently, a guaranteed GDPR compliant blockchain does not exist, only methods to enhance privacy and data protection to create a “more GDPR compliant” blockchain. This is much due to the technical solutions to not have been tested or discussed by a legal authority. Some comments about cryptography can be

found in legal literature and the GDPR advise personal data and its processing to always be encrypted. However, with the address being an encrypted message put in a blockchain and still raise questions on data protection from legal sources, encryption has proven to not resolve the issue entirely. Some technical solutions are for each individual user to implement themselves. This is not going to be enough for the GDPR, according to the legal respondents. It is up to the implementers of the blockchain solution to enable the data subject their rights, not for each user to be liable to protect their own privacy.

Solutions with a permissioned blockchain, where the handling of personal data in a blockchain environment can be done between a selected number of entities, resemble the traditional centralized storage of data. The fact that almost all legal parties in this thesis promote them over permissionless blockchains might indicate that there will be less problems concerning the GDPR in the future if choosing to go with a permissioned solution for a blockchain project. It is also a favoured option according to technical articles and technical respondents, at least for when the use case allows it. Both because of regulatory uncertainties but also due to technical advantages of efficiency. However, some questions still linger with a permissioned solution, which is discussed by both groups. For example, how to handle rectification and erasure of personal data. Overall, a permissioned blockchain is more GDPR compliant than a permissionless blockchain. But is it fully GDPR compliant?

Permissionless blockchains have in general more problems with GDPR compliance, for example due to the transparency of the ledger. The possibility of mapping addresses to additional information is enhanced when the ledger is open to view and search through. Many technical solutions to disguise the address have been developed for the usage of cryptocurrencies since they have been revealed to be less private that it was first claimed. Some have added cryptographic primitives to enhance the cryptocurrency itself but for some public ledgers and cryptocurrencies it is still up to each user to protect themselves with technical solutions. As discussed, legal voices say this is insufficient as data protection.

The digital identities, which could be a blockchain based solution, are discussed by some sources as the future of controlling and process personal data. This is mostly brought forward by technical sources since the technology is the first crucial point. However, using blockchain to enable digital identities still does raise questions in relation to personal data and the GDPR which in much resemble the discussion of using blockchain for anything else personal data related. If digital identities would however be regulated and possible to use, it would change the situation. It would then be possible to not store personal data in relation to the blockchain, but rather ask for permission to use it occasionally when needed. The pseudonymous address might be more leaning towards being anonymous. With personal data not being stored in addition to the blockchain itself it might make blockchain projects more acceptable for the GDPR in general.

5.3 The Technical and Legal Perspectives

When analysing answers from the respondents and reviewing the literature, it become evident that there are many similarities in how the situation of blockchain and the GDPR is interpreted. In general, legal and technical sources agree on what the main problem areas for the GDPR are, concerning the transparency and immutability of a blockchain. They also agree on how the difference in pace of technical innovation and regulation is causing a status quo which is now putting both developers and authorities in difficult positions on how to move forward.

Both groups suggest that test environments or pilot projects involving both parties would be an idea to move forward even if there is a hesitation from the legal side on how this would affect the solutions built under this supervision. The fear of developed solutions created in a safe environment could be legally uncertain outside of this environment is put forward from the legal side. It is suggested from a legal point of view that projects should reach out to data protection authorities when developing a project in larger scale to receive consultation in how to be compliant. Several blockchain projects mention how they have done research of GDPR before implementing a solution and sometimes the changes needed to the project have been strict. However, since the DPA of Sweden does not have specific regulations for blockchain solutions it might be hard for projects to get the right support.

The roles of the controller, joint controller, and processor in a blockchain is something interpreted differently by entities within the legal sphere. Based on the literature but also the legal respondents it is clear that this is a highly case-by-case factor that need to be evaluated for each situation based on what activities each entity is involved with. Clarity for the general situation can however be improved to help blockchain projects do right from the start, in particular when there is the possibility of look at the problems from the micro- or macro-levels discussed by Bacon et al. (2017).

A slight difference can be shown when it comes to acknowledging the address as something which could be put into a blockchain. From a technical point of view the address linked to an individual is something that is almost required to be allowed by the GDPR, since it is an important piece to enable functionality in blockchain projects. From a legal perspective the address is something compromising the GDPR compliance and is still up for debate on whether it will be allowed or not. Although, differences can be seen in each group. For example, CNIL, which as a legal entity also approaches the address as something approved to put into a blockchain even if it put a disclaimer with it needing to be disguised with commitment schemes.

A small detail but which have been evident in the literature review is how legal sources often point to the fact that a blockchain, theoretically, is not immutable while technical sources name it as just “immutable”. This detail might also be an indicator of how the perceptions of

the technology differ and how it is important that both parties get a shared consensus on the properties on blockchains when evaluating them in relation to the GDPR.

Another difference is the level of enthusiasm for blockchain as a useful technology. The technical respondents, since involved in blockchain projects, are positive to the value blockchain could bring while the legal respondents are more hesitant since they see what the many risks are. As Mörner says, this difference of opinion might risk the authorities to not engage in the technology. This could in extension lead to a lock-down on the situation with personal data in blockchain projects which would harm not only blockchain projects but also the evolvement of the GDPR. Finck discusses how the questions raised in the GDPR-blockchain relationship will not be unique to the situation with blockchain. The uncertainties should therefore be used to clarify the regulations further to enhance its technology neutral trait. The answer to resolve these uncertainties, according to Finck and other legal sources, is further research and especially interdisciplinary research to make the technology and law to work together. Technical respondents are not foreign to the thought of working together on this matter, saying that the legal side need to understand a few things about the technology to let it be used for their purposes.

From both a technical and legal point of view, it is raised thoughts on how personal data might not be considered to be the asset it has been declared to be the last years, in the future. Not all organisations or companies developing blockchain-based solutions see data as something valuable which they want to collect but rather something necessary evil to handle in relation to their project. This show how data minimisation of personal data is not necessary only from a legal concern but can be a joint objective to work towards from both sides.

But it is not only the technical enthusiasts and legal authorities which will have an impact on the future of blockchain technology. Kotsios discusses how much money have been invested in the technology and how that might lead to interpretations from a legal side to be done in favor of blockchain in the future. A regulatory limitation of blockchain as a technology would compromise the investments put into blockchain projects. Investments of projects come from both venture capitalists but also from the EU and national municipalities. SimpliCITY being one example. Interest from both tech giants, venture capitalists, the blockchain community as well as municipalities and taxpayers might play a role for certain outcomes when evaluating the technology against the GDPR and data protection. No one can for certain know how this will play out. However, it can be concluded that the context around blockchain is complex, with many groups beside developers and legal authorities to take into consideration.

5.4 Further Research

The lack of research on this current subject has become visible when conducting a literature review. New research develops continuously but even more would greatly benefit the regulatory guidelines on how to handle novel technologies in a technology neutral regulation

such as the GDPR. A few areas, already mentioned by Finck, that the EU would need to investigate further to expand the technology neutrality of the GDPR are how a pseudonymised address should be handled, what roles is relevant to have for each party involved in a blockchain and how the problem of inability of rectification and erasure of personal data should be resolved.

For both permissioned and permissionless blockchain-based token systems, other regulation than the GDPR might be relevant to take into consideration. One example is if a blockchain-based token system classifies as a financial instrument. In 2018, the Swedish National Bank stated that cryptocurrencies, like Bitcoin, is viewed as assets rather than a currency and is therefore regulated thereafter (Söderberg, 2018). Some countries, both inside and outside of the EU, have regulated cryptocurrencies themselves and exchange of them, while others have waited. A global regulatory framework has been requested to be prepared for both opportunities and challenges cryptocurrencies might bring (Viens, 2019).

A subject interesting to investigate further would be what actors are involved in determining the future of blockchain and what their respective interests are. This would enhance the knowledge about what interdisciplinary, or rather multidisciplinary, work would be needed to make the network of actors around blockchain technology agree, and to ensure the technology is used under the right regulations.

Legal respondents question the aim of using a blockchain when personal data is involved, much due to the many uncertainties of the future of privacy demands on the technology. Even though the technical respondents mostly see the worth of blockchain, many of them also question using blockchain for some use cases. The choice to use a blockchain instead of, for example a database, should be evaluated on how a blockchain would affect the system it is used in over time. Even though a permissionless solution is better from a GDPR perspective, factors such as how frequently one will read/write to the blockchain as well as what system environment is used affect the pros and cons of using a blockchain instead of a similar architecture such as a distributed database (Bergman et al., 2019). Also, problems with scaling is a recurring issue discussed in articles as something which need thorough research before implementing a blockchain in a larger scale. Solutions aiming to protect the address, which have been seen in this thesis being one of the most important issues of GDPR compliance, is often decreasing the efficiency of processing in a blockchain. This might prevent privacy solutions from being reasonable to use in reality. These questions might need to be resolved concurrently as the privacy question and will in return affect how the situation for processing of personal data in blockchains will be like in the future. Data protection must also be included in the overall technical enhancement of blockchain technology.

6. Conclusions

It is in this chapter presented what conclusions can be drawn from the previous discussions. Firstly, the current situation is presented in section 6.1. Section 6.2 summarize the technical mitigations which exists while section 6.3 lift what the potential future could look like for blockchain and the GDPR. Lastly, suggested further research on the theme of this thesis is presented in section 6.4.

6.1 The Situation on GDPR Compliance for Blockchain

The main compatibility issues with the GDPR have much to do with the existence of an address included in the transactions of many blockchain solutions. The address can be seen as pseudonymous and is therefore classified as personal data in the eyes of the GDPR. When personal data is incorporated into a blockchain several issues arise. The transparency and immutability traits of a blockchain force the data to be visible, distributed and not possible to rectify or erase on the data subject's command. Distributing data over a network also comprise several of the protection principles of the GDPR. Furthermore, questions around what roles different participants of a blockchain solution will be interpreted to possess is not yet resolved by legal parties. The uncertainties are putting both blockchain developers and data protection authorities in a status quo situation, hard to move forward from.

6.2 What are the Solutions?

Technical solutions disguising the address with non-interactive zero-knowledge proofs, mixing and cryptographic primitives exists and some cryptocurrencies make use of them to enhance privacy. No technical solutions are yet to be seen as fully GDPR compliant, due to lack of trials and regulatory guidelines. In general, it is more compliant to make use of a permissioned ledger than a permissionless. Digital identities, making one's personal data possible to authorize for certain purposes and later retracted, could be an implementation working closer to a GDPR compliant solution. This is not however up for specific blockchain projects to incorporate but need to be handled on a higher level in consensus with the rest of the society and might still raise questions in relation to the GDPR.

6.3 Interdisciplinary Research is Needed

The technical and legal perspectives on the situation of the GDPR and blockchain have similarities in identifying problem areas between technology and regulation as well as how the current status quo situation is restraining on both sides. Differences can be seen in how the severity of putting a pseudonymised address in a blockchain is rated and also in the general belief in blockchain technology providing value when risking data protection worthwhile. The differences indicate a split view on how the technology would stand in a trial and could be helped by regulators getting a more comprehensive overview of the technology they are regulating. Interdisciplinary work is needed to enable blockchains and

their implementations to be regulated in the appropriate way by authorities. This opinion is shared between both legal and technical sources. A convergence of the regulation is on the horizon, but with a 48 % of blockchain projects today claiming to be unsure how to proceed due to uncertainty with regulation, initiatives are needed sooner than later.

References

Bacon, J., Michels, J. D., Millard, C., Singh, J. (2017), *Blockchain Demystified*, Legal Studies Research Paper No. 268, Queen Mary University of London, School of Law. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3091218 (2019-07-01).

Baliga, A. (2017), *Understanding Blockchain Consensus Models*, Persistent. Available online: https://pdfs.semanticscholar.org/da8a/37b10bc1521a4d3de925d7ebc44bb606d740.pdf (2019-10-22)

Bashir, I. (2018), *Mastering Blockchain*, 2nd edition, Packt Publishing. E-book. Available online: https://learning.oreilly.com/library/view/mastering-blockchain-/9781788839044/ (2019-05-20).

Bergman, S., Asplund, M., Nadjm-Tehrani, S. (2019), *Permissioned blockchains and distributed databases: A performance study, Concurrency Computation*, Practice and Experience. Available online: https://doi.org/10.1002/cpe.5227 (2019-10-15).

Booth, A., Sutton, A., Papaioannou, D. (2016), *Systematic Approaches to a Successful Literature Review,* 2nd edition, Los Angeles: SAGE Publications.

Botjes, E. (2017), Pulling the Blockchain apart.. The transaction life-cycle, Ignition at Medium. Available online: https://medium.com/ignation/pulling-the-blockchain-apart-the-transaction-life-cycle-7a1465d75fa3 (2019-10-22)

Bryman, A. (2016), *Social Research Methods*, 5th edition. Oxford: Oxford University Press.

Buterin, V. (2017), *The Meaning of Decentralization*, Medium. Available online: https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274 (2019-05-22).

Buterin, V. (2015), *On Public and Private Blockchains*, Ethereum Blog. Available online: https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/ (2019-05-23).

Calder, A. (2018), *EU GDPR - A pocket guide*. 2nd edition, Ely, Cambridgeshire, United Kingdom: IT Governance Publishing. Available online: https://doi.org/10.2307/j.ctv6cfnkk (2019-06-10).

CNIL (2018a), Blockchain and the GDPR: Solutions for a responsible use of the blockchain in the context of personal data. Available online: https://www.cnil.fr/en/blockchain-and-gdpr-solutions-responsible-use-blockchain-context-personal-data (2019-07-01).

CNIL (2018b), Blockchain. Available online: https://www.cnil.fr/sites/default/files/atoms/files/blockchain.pdf (2019-07-01).

Coinmarketcap.com (2019), *Top 100 Cryptocurrencies by Market Capitalization*. Available online: https://coinmarketcap.com/ (2019-09-15).

Cope, J. (2002), *What's a Peer-to-Peer (P2P) Network?* Computerworld. Available online: https://www.computerworld.com/article/2588287/networking-peer-to-peer-network.html (2019-06-03).

Corbin, J., Strauss, A. (2015), *Basics of Qualitative Research*, 4th edition, Thousand Oaks, California: SAGE Publications.

Dapp, M. M. (2018), *Toward a Sustainable Circular Economy Powered by Community-Based Incentive Systems*. In: Treiblmaier H., Beck R. (eds) Business Transformation through Blockchain. Palgrave Macmillan, Cham. Available online: https://doi.org/10.1007/978-3-319-99058-3_6 (2019-10-14).

Data Protection Commission (2019), *Guidance Note: Guidance on Anonymisation and Pseudonymisation.* Available online: https://www.dataprotection.ie/sites/default/files/uploads/2019-06/190614%20Anonymisation%20and%20Pseudonymisation.pdf (2019-09-05).

De Filippi, P. (2016), The Interplay between Decentralization and Privacy: The Case of Blockchain Technologies, Journal of Peer Production (7). Available online: https://ssrn.com/abstract=2852689 (2019-09-03).

Deloitte (2019), *Deloitte's 2019 Global Blockchain Survey - Blockchain get down to business.* Available online: https://www2.deloitte.com/content/dam/Deloitte/se/Documents/risk/DI_2019-global-blockchain-survey.pdf (2019-10-15).

De Meijer, C. (2018), *Blockchain versus GDPR and who should adjust most*, Finextra. Available online: https://www.finextra.com/blogposting/16102/blockchain-versus-gdpr-and-who-should-adjust-most (2019-10-15).

de Melo Bezerra, J., Massaki Herata, C., Randall, D. (2015), *A Conceptual Framework to Define Incentive Mechanisms for Virtual Communities*, Journal of Universal Computer Science, vol. 21. Available online: https://pdfs.semanticscholar.org/fb99/ef592042581169b3eeb8ca83262babf79c84.pdf (2019-10-14).

Denis Le Sève, M., Mason, N., Nassiry, D. (2018), *Delivering blockchain's potential for environmental sustainability,* ODI. Available online: https://www.odi.org/sites/odi.org.uk/files/resource-documents/12439.pdf (2019-10-15)

Dierksmeier, C. & Seele, P. (2018), *Cryptocurrencies and Business Ethics*, Journal of Business Ethics, 152 (1), pp. 1-14. Available online: https://doi.org/10.1007/s10551-016-3298-0 (2019-08-20).

EnergyCoin Foundation (2019), Avoided CO2 as means of exchange. Available online: https://www.energycoinfoundation.org/en/ (2019-06-24).

EY (2018), EY launches the world's first secure private transactions over the Ethereum public blockchain. Available online: https://www.ey.com/en_gl/news/2018/10/ey-launches-the-world-s-first-secure-private-transactions-over-the-ethereu-public-blockchain (2019-09-05).

Ferrari, V. (2018), *EU Blockchain Observatory and Forum Workshop on GDPR*, Data Policy and Compliance, Institute for Information Law Research Paper No. 4. Available online: http://dx.doi.org/10.2139/ssrn.3247494 (2019-07-03).

Finck, M. (2017), *Blockchains and Data Protection in the European Union*, Max Planck Institute for Innovation & Competition Research Paper No. 18-01. Available online: https://doi.org/10.2139/ssrn.3080322 (2019-06-17).

Finck, M. (2019), *Blockchain and the General Data Protection Regulation: Can distributed ledgers be squared with European data protection law?* Panel for the Future of Science and Technology, European Parliamentary Research Service. Brussels: Scientific Foresight Unit. Available online: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634445/EPRS_STU(2019)634445_EN.pdf (2019-07-01).

Genkin, D., Papadopoulos, D., Papamanthou, C. (2018), *Privacy in decentralized cryptocurrencies*, Communications of the ACM 61:(6) pp. 78-88. Available online: https://doi.org/10.1145/3132696 (2019-07-03).

Goldfeder, S., Kalodner, H., Reisman, D., Narayanan, A. (2018), *When the cookie meets the blockchain: Privacy risks of web payments via cryptocurrencies*, Proceedings on Privacy Enhancing Technologies, (4):179–199. Available online: https://doi.org/10.1515/popets-2018-0038 (2019-09-03).

Henry, R., Herzberg, A., Kate, A. (2018), *Blockchain Access Privacy: Challenges and Directions*, IEEE Security & Privacy 16:(4). Available online: https://doi.org/10.1109/MSP.2018.3111245 (2019-07-03).

Hyperledger Fabric (2019a), *Introduction*, Hyperledger Fabric Docs. Available online: https://hyperledger-fabric.readthedocs.io/en/release-1.4/whatis.html (2019-09-17).

Hyperledger Fabric (2019b), *Hyperledger Fabric Functionalities*, Hyperledger Fabric Docs. Available online: https://hyperledger-fabric.readthedocs.io/en/release-1.4/functionalities.html (2019-09-17).

Hyperledger Fabric (2019c), *Private data*, Hyperledger Fabric Docs. Available online: https://hyperledger-fabric.readthedocs.io/en/release-1.4/private-data/private-data.html (2019-09-17).

IDC (2019), *Worldwide Blockchain Spending Forecast to Reach $2.9 Billion in 2019, According to New IDC Spending Guide*. Available online: https://www.idc.com/getdoc.jsp?containerId=prUS44898819 (2019-10-15).

IT Governance Privacy Team (2017), *EU General Data Protection Regulation (GDPR): An Implementation and Compliance Guide*, 2nd edition. Ely, Cambridgeshire, United Kingdom: IT Governance Publishing. Available online: https://doi.org/10.2307/j.ctt1trkk7x (2019-06-11).

Janze, C. (2017), *Are Cryptocurrencies Criminals Best Friends? Examining the Co-Evolution of Bitcoin and Darknet Markets*, Twenty-third Americas Conference on Information Systems, Boston. Available online: https://www.researchgate.net/publication/326405862_Are_Cryptocurrencies_Criminals_Best_Friends_Examining_the_Co-Evolution_of_Bitcoin_and_Darknet_Markets (2019-08-20).

Kairos Future (2018), *Fastighetsköp och lagfart genom en blockkedja – governance och juridik*. Available online: https://www.lantmateriet.se/contentassets/8d2b5d7647634c02a329b01e46e61071/publikation-swe-fastighetskop-och-lagfart-genom-en-blockkedja--governance-och-juridik-2018.pdf (2019-09-14).

Kosba, A., Miller, A., Shi, E., Wen, Z., Papamanthou, C. (2016), *Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts*, 2016 IEEE Symposium on Security and Privacy (SP). Available online: https://doi.org/10.1109/SP.2016.55 (2019-09-03).

Kozlovski, S. (2018), *A Thorough Introduction to Distributed Systems*, freeCodeCamp. Available online: https://www.freecodecamp.org/news/a-thorough-introduction-to-distributed-systems-3b91562c9b3c/ (2019-06-20).

Kvale, S., Brinkmann, S. (2014), *Den kvalitativa forskningsintervjun*, 3rd edition, Lund: Studentlitteratur AB.

Lauslahti, K., Mattila, J., Seppälä, T. (2017), *Smart Contracts – How will Blockchain Technology Affect Contractual Practices?* ETLA Reports No 68. Available online: https://pub.etla.fi/ETLA-Raportit-Reports-68.pdf (2019-07-28).

Lavrenov, D. (2019), *Securing a Blockchain with a Noninteractive Zero-Knowledge Proof*, Altoros.com. Available online: https://www.altoros.com/blog/securing-a-blockchain-with-a-noninteractive-zero-knowledge-proof/ (2019-09-22).

Lewis, A. (2016), *A gentle introduction to Ethereum*, Bitsonblocks.net. Available online: https://bitsonblocks.net/2016/10/02/gentle-introduction-ethereum/ (2019-09-14).

Lindström, K. (2019), *Så många GDPR-incidenter har anmälts efter ett år med nya lagen*, ComputerSweden. Available online: https://computersweden.idg.se/2.2683/1.719289/gdpr-incidenter-anmalda-eu (2019-10-15).

Lundkvist, C. (2017), *Introduction to zk-SNARKs with Examples*, ConsenSys Media. Available online: https://media.consensys.net/introduction-to-zksnarks-with-examples-3283b554fc3b (2019-09-03).

Lyons, T. (2018), *Blockchain Innovation In Europe*, The European Union Blockchain Observatory and Forum. Available online: https://www.eublockchainforum.eu/reports (2019-10-15).

Lyons, T., Courcelas, L., Timsit, K. (2019), *Blockchain and the Digital Identity,* European Union Blockchain Observatory & Forum. Available online: https://www.eublockchainforum.eu/reports (2019-10-15).

Magnusson, E., Marecek, J. (2015), *Doing interview-based qualitative research : a learner's guide*, Cambridge: Cambridge University Press.

Mattila, J. (2016), *The Blockchain Phenomenon – The Disruptive Potential of Distributed Consensus Architectures*, ETLA Working Papers No 38. Available online: http://pub.etla.fi/ETLA-Working-Papers-38.pdf (2019-05-28).

Nakamoto, S. (2008), *Bitcoin: A Peer-to-Peer Electronic Cash System.* Bitcoin.org. Available online: https://bitcoin.org/bitcoin.pdf (2019-05-22).

Onik, M., Kim, C., Lee, N., Yang, J. (2019), *Privacy-aware blockchain for personal data sharing and tracking,* Open Computer Science, 9(1), pp. 80-91. Available online: https://doi.org/10.1515/comp-2019-0005 (2019-07-03).

PwC (2018), *Blockchain is here. What's your next move? - PwC's Global Blockchain Survey 2018*. Available online: https://www.pwc.com/blockchainsurvey (2019-09-25).

Pyle, E., Bertran Manyé, L., Swerdloff, J., Sharp, L. G., Irvin, R. E., Koziol, J., Jownani, S., Holt, S., Goodloe, V. (2018), *Decoding GDPR*, Judicature, 102(1), pp. 58-66. Available online: https://www.edrm.net/resources/gdpr-resources/decoding-gdpr/ (2019-06-09).

Raj, K. (2019), *Foundations of Blockchain*, Packt Publishing. E-book. Available online: https://learning.oreilly.com/library/view/foundations-of-blockchain/9781789139396/ (2019-05-20).

Rosenberg, M., Frenkel, S. (2018), *Facebook's Role in Data Misuse Sets Off Storms on Two Continents*, The New York Times. Available online: https://www.nytimes.com/2018/03/18/us/cambridge-analytica-facebook-privacy-data.html (2019-10-17).

Salzburg Research (2019), *SimpliCITY.* Available online: https://www.salzburgresearch.at/en/projekt/simplicity/ (2019-06-26).

Saunders, M., Lewis, P., Thornhill, A. (2016), *Research methods for business students*, 7th edition, Harlow: Pearson Education.

Schmelz, D., Fischer, G., Niemeier, P., Zhu, L., Grechenig, T. (2018), *Towards Using Public Blockchain in Information-Centric Networks: Challenges Imposed by the European Union's General Data Protection Regulation*, 2018 1st IEEE International Conference on Hot Information-Centric Networking (HotICN). Available online: https://doi.org/10.1109/HOTICN.2018.8606000 (2019-09-04).

SimpliCITY (2019), *Project Description*. Available online: https://www.simplicity-project.eu/en/projectdescription/ (2019-06-26).

Singhal, B., Dhameja, G., Sekhar Panda, P. (2018), *Beginning Blockchain: A Beginner's Guide to Building Blockchain Solutions*, Apress. E-book. Available online: https://learning.oreilly.com/library/view/beginning-blockchain-a/9781484234440/ (2019-05-20).

Solarcoin (2019), *Introducing SolarCoin*. Available online: https://solarcoin.org/ (2019-06-24).

Soltani, R., Nguyen Trang, U., An, A. (2018), *A New Approach to Client Onboarding Using Self-Sovereign Identity and Distributed Ledger*, IEEE. Available online: https://ieeexplore.ieee.org/document/8726515 (2019-10-15).

Statt, N. (2019), *Facebook confirms it will launch a cryptocurrency called Libra in 2020*, The Verge. Available online: https://www.theverge.com/2019/6/18/18682290/facebook-libra-cryptocurrency-visa-mastercard-digital-currency-calibra-wallet-announce (2019-08-22).

Söderberg, G. (2018), *Är Bitcoin och andra kryptotillgångar pengar?* Ekonomiska kommentarer - Sveriges Riksbank vol 5. Available online: https://www.riksbank.se/globalassets/media/rapporter/ekonomiska-kommentarer/svenska/2018/ar-bitcoin-och-andra-kryptotillgangar-pengar.pdf (2019-10-14),

Viens, A. (2019), *Mapped: Cryptocurrency Regulations Around the World*, Visual Capitalist. Available online: https://www.visualcapitalist.com/mapped-cryptocurrency-regulations-around-the-world/ (2019-10-14).

Voigt, P., von dem Bussche, A. (2017), *The EU General Data Protection Regulation (GDPR): A Practical Guide.* Cham: Springer International Publishing. E-book. Available online: https://doi.org/10.1007/978-3-319-57959-7 (2019-06-09).

Wang, L., Shen, X., Li, J., Shao, J., Yang, Y. (2019), *Cryptographic Primitives in Blockchains*, Journal of Network and Computer Applications 127. Available online: https://doi.org/10.1016/j.jnca.2018.11.003 (2019-09-20).

Wang, W., Thai Hoang, D., Hu, P., Xiong, Z., Niyato, D., Wang, P., Wen, Y., In Kim, D. (2019), *A Survey on Consensus Mechanisms and Mining Strategy Management in Blockchain Networks*, IEEE Access Vol 7. Available online: https://ieeexplore.ieee.org/document/8629877 (2019-05-22).

Wirth, C., Kolain, M. (2018), *Privacy by BlockChain Design: A Blockchain-enabled GDPR-compliant Approach for Handling Personal Data*. In: Prinz, W., Hoschka, P. (Eds.), Proceedings of the 1st ERCIM Blockchain Workshop 2018, Reports of the European Society for Socially Embedded Technologies. Available online: http://dx.doi.org/10.18420/blockchain2018_03 (2019-08-22).

World Economic Forum (2011), *Personal Data: The Emergence of a New Asset Class.* Available online: https://iapp.org/media/pdf/knowledge_center/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf (2019-10-17).

Wu, H., Wang, F. (2014), *A Survey of Noninteractive Zero Knowledge Proof System and Its Applications*, The Scientific World Journal. Available online: https://doi.org/10.1155/2014/560484 (2019-09-03).

Zcash (2019), *What are zk-SNARKs?* Zcash.com. Available online: https://z.cash/technology/zksnarks/ (2019-09-03).

Zhang, R., Xue, R., Liu, L. (2019), *Security and Privacy on Blockchain*, ACM Computing Surveys (CSUR), 52 (3). Available online: https://doi.org/10.1145/3316481 (2019-09-20).

Zheng, Z., Xie, S., Dai, H., Chen, X., Wang, H. (2017), *An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends*, IEEE: 2017 IEEE International Congress on Big Data (BigData Congress). Available online: https://ieeexplore.ieee.org/document/8029379 (2019-06-14).

Interviews and E-mail

Arabaci, O: Technical Consultant at IBM. Interview 2019-06-13.

Devos, B: Business developer at Cipal Schaubroeck. E-mail 2019-09-17.

Hallström, E: Lawyer at Data Protection Authority of Sweden. E-mail 2019-09-20.

Jin, Y: Technology Specialist at Ericsson. Interview 2019-08-08.

Kotsios, A: PhD in Civil Law at Uppsala University. Interview 2019-09-11.

Mulder, B: Board Member at EnergyCoin. Interview 2019-09-19.

Mörner, E. & Wallace, A: Junior Associates at IT-advokaterna. Interview 2019-09-25.

Olsson, D: Senior Architect at Truesec. Interview 2019-09-18.

Pronk, A-M: Board Member at EnergyCoin. Interview 2019-10-02.

Rosén, J: Strategic Advisor at Uppsala Kommun. Interview 2019-06-12.

Rubbestad Lilja, J: Business Developer at Uppsala Kommun. Interview 2019-07-16.

Stabauer, P. & Schrempf, B: Board Members at SimpliCITY. Interview 2019-08-20.

Wögerbauer, C: Developer at Polycular. Interview 2019-05-31.

Appendix A

Interview questions for the SimpliCITY respondents

Respondents

- Can you describe your role in the SimpliCITY-project?
- How did you get involved?

SimpliCITY

- Could you describe SimpliCITY?
- What are the involved actors for the organisation?
- What was included in the pre-study?
- What services are thought to be relevant?
- Bike mobility
    - What services?
    - What actors?
    - Who are the users?
    - What would the platform enable in detail?
    - What would be the future evolvement?
- Social inclusion & local consumption?
    - What services more specifically?
    - Would the services be regional?
- Would the tokens be of a local kind or able to be used as a cryptocurrency or regular currency?
- Which role do Polycular have?
- Who will have ownership over the platform?
- Who will be responsible for maintenance?
- What are your view on the benefits with such a platform?
- What are your view on the risks with such a platform?

Users

- Who is the typical user?
    - How would they be involved?
- Have there been discussions on which data will be collected on the users?
    - What discussions related to privacy/GDPR have you had?
- Have you been looking at other projects doing with the same goal?

Appendix B

Interview questions for the technical respondents

Intervjuperson

- Vill du berätta lite kort om dig själv?
- Vilka projekt eller uppgifter är du involverad i just nu?
- Vilken erfarenhet har du av blockkedjor och dess tillämpningar?

Blockkedjor

- Hur arbetar företaget du arbetar på med blockkedjor?
- Vilka exempel på aktuella projekt med blockchain tycker du är intressant?
- När passar en blockchain bättre än andra typer av databaser?
- Vad är din uppfattning om publika blockkedjor relativt privata blockkedjor?
    - Vilka typer av blockkedjor tror du är de som har mest möjligheter?
    - Vilka implementationer är lämpliga?
- Vilka utmaningar står företag som bygger blockkedjor inför?
- När tror du blockkedjor är redo att implementeras i “vardagliga” lösningar?

Personuppgifter/GDPR

- Vad är ditt intryck av GDPR?
- Har projekt med blockkedjor påverkats av GDPR? Isåfall hur?
- Hur diskuteras frågan kring persondata och integritet för blockkedje-lösningar?
- Hur diskuteras anonymisering och pseudonymisering?
- Vilka utmaningar med personuppgifter tror du är de viktigaste?
- Vad tror du är det viktigaste för lagstiftningen att förstå angående blockkedjor?
- Vilken roll kommer blockkedjor att ha i framtiden?

Appendix C

Interview questions for the legal respondents

Intervjuperson/er

- Vill du/ni berätta kort om dig/er själv(a) och vad du/ni gör?
- Vilka aspekter av data integritet har du/ni tittat på?

Personuppgifter & blockchain

- Vad är en personuppgift?
- Hur avgör man vilka fragment som skulle kunna vara en personuppgift?
- Hur ser du på situationen med kryptovalutor och blockchain i stort?
    - Är exempelvis Bitcoin anonymt?
- Hur troligt skulle du säga att det är att kunna härleda en identitet baserat på att kartlägga transaktioner?
    - Ibland uppmanas användare att använda nya adresser, vilket ansvar kan läggas på individnivå?
    - Hur ser lagen på pseudonymisering/anonymisering?
- Hur ser lagen på kryptering av data?
    - Vilka tidsperspektiv har man?
- Vad är den största risken med att samla persondata som företag?

GDPR

- Skulle du kunna beskriva syftet med GDPR?
- Kan GDPR vara teknikneutral?
- I förhållande till en blockkedja, vad tror du om:
    - Oklarheter i hur länge personuppgifter får sparas?
    - Oklarheter i personuppgifter ska raderas?
    - Vem som ska anses vara ägande/processande av persondata?
- Vad tror du om förslag att en användare kan avsäga sig rätten till radering av persondata?
- Vad tror du om framtiden för GDPR?
    - Hur kommer ändringar/tillägg/tolkningar ske?
    - Hur står den sig mot nya tekniker?

Appendix D

Interview questions for the blockchain project respondents

Respondent

- Tell me a bit about yourself and your role in the project?
- Could you describe the project?

The project

- Why did you choose to use blockchain?
- Did you do any research of the GDPR in relation to the project?
    - How did that affect your technical solution for the project?
- Do you collect any personal data?
    - How is it stored?
- How do you follow up on data privacy in relation to your project?
- What benefits do you see with the project?
- What risks do you see with the project?
- How do you believe the future will look like:
    - For your project
    - Blockchain technology/cryptocurrencies as a whole?
- What are the most important things legal people need to notice about blockchain?