UNIVERSITY OF TARTU
Institute of Computer Science
Innovation and Technology Management Curriculum

Ilona Pavlenkova

GDPR and Blockchain Solutions

Systematic Literature Review

Master's Thesis (20 ECTS)

Supervisor(s): Fredrik Payman Milani

Tartu 2020

GDPR and Blockchain Solutions

Abstract

Within the last years, there has been an increased attention for the development of Blockchain technologies and solutions. Blockchain represents one of the major headways' technologies of the past decade, giving large groups of people and companies a possibility to reach agreement on record information without a central authority. Partially this can be achieved because data once stored on a Blockchain is immutable. The General Data Protection Regulation (GDPR) was developed in the European Union (EU) with the aim to standardize the privacy regulations across Europe and introduce the changes on how the personal data should be processed. The GDPR consists of a series of articles and chapters that require, among other things, provision a consent to collect and store data, identification of data controller, provision of encryption, and anonymization of data. This thesis explores the connection between Blockchain and GDPR. In particular, the thesis presents a decision framework that can help developers and decision-makers in their design of the GDPR-compliant blockchain projects. The framework, presented in the thesis, is developed based on Systematic Literature Review of existing studies on the interplay of technology and regulation. The provided findings will help the Blockchain development teams to consider the GDPR-related factors when designing and implementing the blockchain projects.

Keywords: Blockchain, GDPR, Compliance, Systematic Review

CERCS: P170

Isikuandmete kaitse üldmäärus ja Plokiahela lahendused

Lühikokkuvõte

Plokiahela tehnoloogiate ja lahenduste arendamine on viimastel aastatel on olnud kõrgendatud tähelepanu all. Plokiahel on viimase aastakümne suurim tehnoloogiline edasiminek, andes suurtele inimgruppidele ja ettevõtetele võimaluse sooritada tehinguid ja kokkuleppeid ilma keskse organita. Osaliselt on see saavutatav seetõttu, et andmed mis on kord plokiahelas salvestatud on muutumatud. Isikuandmete kaitse üldmäärus arendati Euroopa Liidus (EL) välja eesmärgiga standardiseerida privaatsusregulatsioone kogu Euroopas ning tutvustada muudatusi, kuidas isikuandmeid tuleks käsitleda. Isikuandmete kaitse üldmäärus koosneb reast artiklitest ja peatükkidest, mis muuhulgas nõuavad nõusoleku hankimist andmete kogumiseks ja salvestamiseks, andmete vastutava töötleja identifitseerimist, andmete krüpteerimist ning anonüümseks muutmist. Käesolev lõputöö uurib seost plokiahela ja isikuandmete kaitse üldmääruse vahel. Täpsemalt esitab töö otsustusraamistiku, mis aitab arendajaid ja otsustajaid isikuandmete kaitse üldmäärusega vastavuses olevate plokiahela projektide loomisel. Magistritöös esitatud raamistik on välja töötatud olemasolevate, tehnoloogia ning regulatsiooni koosmõju uurivate, tööde süstemaatilise kirjandusülevaatena. Esitatud tulemused aitavad plokiahela arendustiimidel arvesse võtta isikuandmete kaitse üldmäärusega seonduvaid tegureid plokiahela projekte disainides ning ellu rakendades.

Võtmesõnad: Plokiahel, Isikuandmete kaitse üldmäärus, Vastavus, Süstemaatiline ülevaade

CERCS: P170

Table of contents

1. Introduction

In recent years, there has been an increasing interest towards the development of Blockchain technologies and solutions. Blockchain can be described as one of the major headways' technologies of the past decade, which gives a possibility to large groups of people and organizations to reach agreement on and permanently record information without a central authority. It has been acknowledged as a significant tool for building a fair, inclusive, secure, and democratic digital economy [1].

Blockchain is represented as a decentralized trustless system using a public ledger with peer-to-peer file sharing and public key cryptography, which can be named as a key innovation [2]. While there are several types of blockchains in existence (permissioned blockchains and permissionless or public blockchains [3]), they share certain functional characteristics - means for nodes on the network for communication, a mechanism for nodes to bring the addition of information to the database, usually in the form of some transaction, and a consensus mechanism for the network to validate the accepted version of the database. The data is stored in groups known as blocks, each validated block is cryptographically sealed to the previous block, forming an ever-growing chain of data. Once a transaction has been added to the chain, it generally cannot be altered or removed [3].

Blockchain technology makes it clear to become an extremely disruptive technology with exponential development in various applications and domains that could have the capacity to reconfiguring the aspects of society and its operations. Along with the rapidly developing technological environment, the critical question regarding personal data protection has arisen and became one of the most important aspects to consider while working on the development and implementation of technological innovations, solutions, and processes.

With the aim to protect and empower all EU citizens' data privacy and to reshape the way organizations across the region approach data privacy, the European Union (EU) in May 2018 has raised into effect the General Data Protection Regulation (GDPR) which replaced the 1995 Data Protection Directive (Directive 95/46/EC). The emphasis of the regulation is to reduce the capability asymmetry between organizations that use and manage personal data, and the individual to whom the data belongs [4]. The GDPR pursues to adjust all Data Privacy Laws across Europe in order to exploit an area of freedom, security, and justice of an economic union, as well as progress in the strengthening of the economies within the internal market. The GDPR imposes it by making adjustments and clarifications of existing rules and establishing new rules on how personal data is processed with the goal to increase the level of integrity.[1] Non-compliance with the GDPR may lead under some circumstances to severe fines of up to 4% of the worldwide annual turnover of the organization.

The GDPR is applicable in cases when personal data is processed. According to the document, information is considered to be personal data when throughout it the natural person is identified or can be identifiable. If personal data is, however, provided in the form of anonymized information that cannot be relatable to an identified or identifiable natural person, the GDPR is not applicable.[2] The main roles introduced in the context of the GDPR are a data subject

[1] GDPR, Recital 2
[2] GDPR, Recital 26

(owner of the data) who turns to exercise the rights to a data controller with the access to his/her personal data, with the possibility of forwarding it to a data processor (organization) that will be responsible for processing the data on behalf of the controller. It is the obligation of the data controller to define how personal data can be processed in accordance with the GDPR constraints [6]. At this step, it is already possible to underline that according to the GDPR it is assumed that it is possible to identify a data controller, but within the Blockchain solutions, based on decentralized systems, it is not always easy to make. The GDPR was formulated and designed following the assumption that data will be stored within centralized entities. However, blockchain technologies are facilitating a move towards a decentralized model of data management. In this regard it becomes a debate question what parts of the GDPR can be complied with when using Blockchain technology, and how the technology will be influenced and affected by the introduction of the Regulation.

It is pointed out that the GDPR expends the principles, rights, and obligations in terms of which data processing can take place. This brings a concern to the Blockchain solutions' developers and decision-makers as now they are required to consider multiple GDPR-related factors before designing the product and they face the situation when there is a lack of familiarization on how and at what stage the requirements of the Regulation are applied, and how to address the issues raised. Despite this question to be actual for practitioners and companies which are working to build blockchain solutions, in the research area there are still a lot of gaps and under-researched issues toward the interconnection of the Blockchain to the GDPR. Even though the GDPR's introduction is an actual and important topic, researchers have not presented the framework to support the design decisions of Blockchain solutions with consideration of the GDPR-related factors. Taking into consideration the described above, it is possible to conclude that the need for developing the respective "decision framework" is determined.

That's why the present thesis aims to address the main research objective while proposing a *decision framework* for decision-makers, analysts, and Blockchain developers so they can consider GDPR within the design process, as long as make their solutions compliant with the GDPR. The main research question is formulated as follows: "How can Blockchain-based solutions be compliant with the GDPR?"

To answer the main research question, this study aimed to address the following sub-questions:

1. What are the fundamental aspects of the GDPR that affect Blockchain solutions?
2. How Blockchain solutions are affected the GDPR?
3. How can Blockchain solutions become GDPR-compliant?

The research will be conducted using the systematic literature review (SLR), which is a methodologically precise review of research results. SLR aims not only just to accumulate the existing documentation on a research question, but it is also designated to support the development of documentation-based guidelines for practitioners [11]. The goal of the research is to analyze, basing on the available literature and resources, the state of art in scientific discussion in part of Blockchain compliance to the GDPR and their interconnection; based on the results received from the analysis of studies, to deduce the decision framework, as well as respective recommendations to the developers and managers of Blockchain solutions in part of the GDPR compliance.

A review of technological aspects of the Blockchain will be assessed and presented through the relevant scientific articles, books, and material, to provide the required understanding of the technological aspects to be able to discuss them. Taking into account the fact that the blockchain being a moderately new technology, and that there is the absence of large-scale research by scholars, the certain non-traditional sources, like internet sources, are required and will be used with the aim to collect and present the enforced knowledge and data.

This study makes an important contribution to the advancement of the conceptual model of interconnection between the Blockchain and the GDPR, by exploring the technology's compliance with the Regulation. The project provides to Blockchain solutions' developers and managers the decision framework to develop a design and implement the GDPR-compliant Blockchain solutions.

This work consists of six chapters and is organized as follows. Following this Introduction, Chapter 2 reviews the background information about Blockchain technology, the key elements, and obligations of the GDPR. Chapter 3 presents the SLR, state of art in scientific discussion in part of the interrelation of GDPR and Blockchain. In Chapter 4 results of the analysis of the implications of the GDPR towards Blockchain solutions are provided and discussed. Chapter 5 presents the decision framework for GDPR-compliant Blockchain solutions. Chapter 6 summarizes the writer's conclusions.

2. Background

With the aim to overview and analyze the respective interconnection between Blockchain and the GDPR, it is required to have an understanding of the Blockchain technology, as well as of general elements and aspects of the GDPR which are relevant to the technology. This chapter will explain the main concepts and elements of Blockchain and followed by the corresponding appraisal of the GDPR.

2.1. Blockchain Technology

2.1.1. Blockchain architecture

There are various ways to develop and implement the software system, and one of the key determinations to settle on is the architecture of the system, in which way the components of the system are structured and associated one to another. The two main architectural approaches are centralized and decentralized (distributed). In the case of centralized systems, the components are positioned around and associated with one central component. Opposed to the mentioned system, the components of decentralized (distributed) systems establish a network of associated components without having any central element of control or coordination [12].

Blockchain is a decentralized, shared, encrypted database that works as a non-convertible and incorruptible storage of information. The technology performs the next step in the peer-to-peer economy, where people are provided with the technical possibility to agree on a certain state of affairs and note the agreement in a protected and verifiable manner with the use of decentralized consensus mechanisms [13]. Peer-to-peer systems can be described as systems that consist of *nodes* (individual computers), which provide their computational resources (e.g., processing power, storage capacity, diffusion of information) to another. Users convert their

computers to nodes of the system after joining a peer-to-peer system and act equally concerning their rights and roles. Within the system, all the nodes have the same working potential and responsibility despite the difference in resources that are contributed [12].

Blockchain is described as a range of data sets which are consisted of a chain of blocks (data sets) where a block consists of multiple transactions. Adding the blocks expands blockchain and introduces a complete ledger of the transaction history. Blocks are legitimized by the network using the means of cryptography. A part of information about a transaction, every block comprises a timestamp, the hash value of the previous (parent) block, and a nonce, random number to certify the hash. The first block of a blockchain, genesis block, has no parent block [15]. The generated *hash values* are unique, and fraudulence can be forfended as in case of changes of the block in the chain the immediate change of hash value will take place. The new block can be added to the chain in the event when the majority of nodes are consent by consensus mechanism to affirm the transactions and the block itself.

A block itself is composed of the block header and the block body. The following parameters are included in the block header: a version of the block, Merkle tree root hash (hash value of all block's transactions), timestamp, nBits (target threshold of a valid block hash), nonce (4-byte field), parent block hash (256-bit hash value pointing to the previous block). The body of the block consists of the transaction counter and transactions related to the block. In order to assert the authentication of transactions, the asymmetric cryptography mechanism is applied [16].

2.1.1.1. Data cryptography

The blockchain uses the technology of the mathematical cryptographic model, named the asymmetric cryptography, with the aim to recognize users in the blockchain and preserve their property respectively [12]. Cryptography can be described as a set of techniques associated with the features of information security representing confidentiality, data integrity, entity authentication, and data origin authentication [18]. The main goal of cryptography is the provision of respective data protection from being acquired by unauthorized people. The cryptographic tools to ensure information security include encryption schemes, hashing, and digital signatures.

Cryptography utilizes the keys to preserve the data, and the digital equivalents of those keys' usage for locking and unlocking the protected data are *encryption* and decryption, respectively [12]. Encrypted data are called cypher text and represent a set of letters and figures which require possession of the relevant key to decrypt it. Decrypted cypher text is equivalent to the original data that have been encrypted. For the number of years, the symmetric cryptography method was used, where the identical key was used to do both the data encrypting and decrypting. This means that each person, who was able to encrypt the data with the use of the key, could automatically decrypt the created cypher text with the same key. However, it became apparent that having one key for encryption and decryption is not expedient, and the *asymmetric* method of cryptography was introduced [12]. Asymmetric cryptography can be utilized in untrusted networks to send messages and can also be used to entirely recognize users and validate transactions [19].

The encryption method used for blockchain technology is said to be a *public-key encryption* scheme where for each associated encryption/decryption pair, one key is made publicly

available (the public key), while the other (the private key) is kept secret [18]. Cypher text generated with one of these keys can only be decrypted with the other key and vice versa. The key obtainable by the person who encrypts the message doesn't need to be secret. The essential part is that the receiver can only decrypt it using a secret key. To realize such a system, the receiver discloses a public encryption key which is known to everyone. At the same time receiver also has a matching secret key, which is utilized for decryption [20].

One of the most principal concepts in ensuring the information security and blockchains integrity, and consequently protecting the ownership from unauthorized access, is 'hash functions', also known as 'hashing'. *Hashing* is a process of text transformation from an arbitrary length to the text of some fixed length, where the converted text is called *hash-value* [18]. The hash reproduces the explicit content of the original text/file, and wherever the content should be re-approved, the same hash algorithm is applied to the file, and the hash signature tends to be the same if the file has not been modified [2].

Hashing serves the goal to compare data (e.g., files or transaction data) without observing their content piece by piece and to detect if the data (e.g., a file or transaction data) that is supposed to stay unchanged was changed at a certain date/moment, or after sending it to someone, or after it was stored in a database [12]. The blockchain is highly dependent on *hash references*, which can be described as a slightly more progressive application case of cryptographic hash values' power. The aim of hash reference is to refer to data (e.g., transaction data) that are placed somewhere (e.g., on a hard disk or in a database) and assure that the data have remained *unchanged*. In order to serve such assurance, the combination of the cryptographic hash value of the data being stored with the respective information about the place where the data are located is provided. In case the data were altered, both parts of information would no longer be sequential and therefore the hash reference would become invalid. References to data represent the digital equivalent to the physical location of it, they are performed as pieces of data that refer to other data. Computer programs use references to mark the respective site *where the data have been placed* and to retrieve them later [12]. Within the blockchain the mechanism presents the possibility to establish the *chain* of hash values and hash references, in which the part of data also contains the hash reference to the previous part of data, generating a chain of linked data [21].

Apart from placing data on-chain, some businesses consider also the *off-chain* database storage to be used for locating the respective part of the data. When using the off-chain database systems, hashes of the corresponding data are stored on the blockchain itself and at the same time in a separately located off-chain database storage, next to the original personal data that was put through the hash function [57]. Although this method sacrifices many of the benefits of using a blockchain in the first place, it is described and recognized as a possibility for storing sensitive data in a more secure and efficient way [58].

2.1.2. Blockchain characteristics

The following two main *types* of blockchain can be distinguished [3]:

- *Permissionless or public blockchains*. These systems represent the open-source network, which anyone can access and use the shared ledger, as well as participate in the consensus process.

- *Permissioned or private blockchains.* These networks are the networks centralized to one organization/entity that specific individuals or entities use to conduct transactions.

Besides the mentioned types of blockchain, in several cases the third type is also presented [14]:

- *Consortium blockchains*. In this type of network, the consensus process is controlled by a pre-selected set of nodes and in this regard, these blockchains may be observed as "partially decentralized".

2.1.2.1. Key components of the Blockchain

With the aim to investigate the substantial parameters of the technology, the following key characteristics of blockchain can be underlined [17]:

- *Decentralization*. In traditional, centralized transaction systems, each transaction needs to be legitimized through the central trusted authority, which leads to the occurred costs and the performance bottlenecks at the central servers. Blockchain doesn't require a third party for the validation of transactions. Data consistency is maintained using the respective consensus algorithms.
  When describing the decentralization, Buterin [22] suggests distinguishing three components for the classification: architectural, political, and logical decentralization. The architectural part defines the number of computers the system consists of, and the amount it tolerates to be simultaneously out of order. Political refers to the individuals or organizations which control the computers in the network. Logical decentralization is responsible for interface and data structure presentation as a monolithic object. Blockchains are represented as politically decentralized (as there is no single point of control) and architecturally decentralized (no infrastructural central point of failure), but they are logically centralized (the one common state is agreed, and the system *behaves* like a single computer). These characteristics support the following arguments for the usability of Blockchain: *tolerance to fail*, *resistance to* probable *attacks*, and *resistance to* possible *collusions* among participants.

- *Persistency and immutability*. Verification of transactions takes very little time and invalid transactions are not conceded. Blocks that are holding improper and false transactions could be instantly revealed. After validation of the transaction and block, it is almost impossible to roll back or remove the transaction. Immutability *ensures* that all data and transaction history in the blockchain is in one trusted source. The immutability of a blockchain adds another layer of security and complexity for those who think to attack the network – even when gaining access rights to a blockchain, the data on the blockchain itself is immutable, thus *preventing* anyone from ever completely erasing or altering records.

- *Anonymity*. The real identity of the users is not revealed as they interact with the blockchain system with the use of the generated address. Anonymity refers to the non-identifiability of the sender and the receiver in one transaction. Blockchain is believed to be very safe in terms of anonymity provision, as users only make transactions with

generated addresses rather than with real identity. In situations where there are leakages, users can create several addresses, which makes the probable traceability to become hard.

- *Auditability*. Blockchain is characterized by immutable transactions to be a part trusted mechanism used by the network. All the history and data about transactions and respective changes are stored within the system. As all the information and data regarding the transaction are staying within the system and not deleted, the corresponding details about the transaction could be tracked and justified upon the need or request.

Another aspect of Blockchain that worth to be mentioned is the provision of so-called smart contracts. Initially, the concept of smart contract was proposed by Nick Szabo, a cryptographer, back in 1994 [30]. Nevertheless, this idea was not popular until the rise of blockchain as a distributed ledger technology. A smart contract is presented as a computer program that is self-verifying, self-executing, and tamper-resistant and it runs on the blockchain platform. Smart contracts are characterized as the protocols that digitally facilitate, verify, and enforce contracts made between two or more participants on the blockchain [31]. Smart contracts are naturally redistributed on the blockchain platform and necessarily exist on the network. They enforce participants to accept the rules written in the contract without any centralized, third-party authorities controlling it in order to reach agreements and solve common problems with minimal confidence [32]. With the aim to provide the security of the blockchain network, smart contracts are mainly performed in closed environments, avoiding the access and import external data [32]. Smart contracts deliver Blockchain solutions both flexibility and power, and due to that reason, it is desired to use them.

2.2. GDPR

The processing of personal data intervenes in diverse areas in economic and social activities, and the respective development of information technologies affects the processing and exchange of such data considerably easier. Contemporaneously, data protection standards are becoming increasingly high with regards to the mentioned reasons. In this context, the European Union (EU) adopted the General Data Protection Regulation (GDPR) to consort the rules for data protection within the EU member states and to enhance the level of privacy for the affected individuals [23]. With the aim to observe the implications of the GDPR it is important to evaluate two key components: the cases where the regulation applies and to who does the regulation employ? In a part of the material scope, it applies to any processing of personal data. The GDPR is intended to be employed to anyone, regardless of a legal entity, that processes or controls the processing of personal data[3]. To the extent of comprehending the GDPR, the relevant definitions are required to be recognized.

2.2.1. Main definitions of the GDPR

Personal data is a core area and definition of the GDPR and it appeals to any information that is related to an identified or identifiable natural person, corresponding ‘data subject’[4]. The data

[3] GDPR, Article 2, 3
[4] GDPR, Article 4

can be considered as personal in case if the identification of a person is possible based on the accessible data, which indicates that person can be detected directly, or indirectly, by the reference to an identifier [23]. The natural person can be identified, for example, by name, an identification number, location data, an online identifier (such as internet protocol addresses, cookie identifiers or radio frequency identification tags) [5] or to one or more particular characteristics to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person[6]. However, the regulation does not apply to the personal data of deceased persons[7]. Identification of the natural person/data subject may occur in case of combining various information, which themselves would not have retraced to the person but does so in combination [23].

Processing under the GDPR represents any operation or set of operations that are performed on personal data or sets of personal data, whether or not by automated means[8]. Substantially, any handling of data can be considered as processing. The means of processing are identified in wide possible range, such as: collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction[9].

The GDPR develops the term of *pseudonymization* as a method used to reinforce privacy by dissociating information that allows data to be attributed to a specific person. The complementary information that might refer data to a specific person is supposed to be kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person[10]. Therefore, every individual pseudonymizing process should be assessed to be able to infer if a pseudonymized dataset is to be marked as personal data or not. In contradistinction to pseudonymized data, anonymized data is entirely excluded from the scope of the GDPR[11].

Controller indicates the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data[12]. This signifies that the controller is responsible for the processing of personal data, which inflicts several legal responsibilities for the controller[13]. In Article 26 the GDPR presents the notion of 'joint controllership' for the cases where two or more controllers jointly determine the purposes and means of processing[14]. In case if aim and processing tools are determined by numerous organizations together, those organizations will divide data protection obligations under the GDPR and have to arrange a clear allocation of responsibilities [23].

Processor is presented as a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller[15]. Processors can, like the controllers,

[5] GDPR, Recital 30
[6] GDPR, Article 4
[7] GDPR, Recital 27
[8] GDPR, Article 4
[9] Ibid.
[10] GDPR, Article 4
[11] GDPR, Recital 26
[12] GDPR, Article 4
[13] GDPR, Article 5
[14] GDPR, Article 26
[15] GDPR, Article 4

take a broad diversity of legal forms and several processors can be directed to act at the same time by the controller [23].

2.2.2. General principles of the GDPR

Lawfulness, fairness, and transparency. The GDPR establishes very particular criteria regarding how the controller and processor process personal data to assure integrity and enhance *transparency*. The aim is to amplify processing that is *lawful and fair*[16]. In order to do it in such a way, only the demanded data is collected, the methods and objectives are communicated with the data subject while awareness is raised regarding how and why the personal data is collected[17].

Purpose limitation principle underlines the importance of processing and collecting personal data in an explicit manner and for legitimate purposes, with the data subject being aware of the forthcoming data processing. Moreover, the collected data should not be further processed in a manner that is not consistent with the objectives because of which it was originally processed. A notable exception from this is the processing for archiving purposes in the public interest, historical research, statistical purpose[18].

Data minimization principle is enforced with the aim to regulate that data is to be kept relevant, adequate, and limited to what is necessary [16]. This relates to pursuing the goal for which data are processed[19].

Accuracy. Personal data are supposed to be processed and stored in accordance with the GDPR in an accurate and up to date manner. All reasonable steps are required to be taken in order to ensure that all inaccurate personal data are erased or rectified without any possible delay[20].

The principle of *storage limitation* imposes that the period for which the personal data can be stored is limited to a very strict minimum and in direct correspondence with the fulfillment of the processing's goal[21].

The '*integrity and confidentiality*' principle guarantees that the data subjects' personal data are processed and stored in a format that provides safety measures to preclude unauthorized access and unlawful processing, and further avoid damage and entire loss of the data[22].

The controller is considered to be *accountable* and should be able to demonstrate compliance with the data processing related principles under the GDPR[23].

2.2.3. Data subject rights

With the implication of the GDPR, the data subjects receive an extensive amount of rights concerning the processing of their personal data that is represented in Chapter 3 of the Regulation.

[16] GDPR, Recital 39
[17] GDPR, Article 5
[18] Ibid.
[19] GDPR, Article 5
[20] Ibid.
[21] Ibid.
[22] GDPR, Article 5
[23] Ibid.

Right to information and access to personal data. The controllers are obliged to grant the data subject with comprehensive *information* regarding the processing and storage of their personal data according to Article 12 in the GDPR. The communication of such information must be made in a "concise, transparent, intelligible and easily accessible form, using clear and plain language"[24]. Furthermore, the controller must, when collecting personal data, notify the data subject not only about the identity of the data controller but also who is the corresponding data protection officer, and besides the aim of the data collection and information regarding its processing and storage[25]. In the case of processing the data subjects' personal information, they also have a right to request *access* for more information on the recipients of such data, categories of data being processed, purposes of the processing, existence of automated decision-making, to lodge a complaint with the supervisory authority[26].

Right to Erasure and Rectification. The GDPR has introduced the empowerment of the data subject and the right to their data by imposing the right to *erasure* (referred to as 'the *right to be forgotten*') and broadening the scope to inflict to the controller, who has made the personal data public, with the duty of informing the controllers processing the data to erase all links to, or copies or replications of such personal data[27]. The right to *rectification* is also based on the concept of empowering data subjects to be in control of their own personal data, enforcing the controller with the responsibility of correcting, on-demand, and without undue delay, inaccurate personal data concerning the data subject[28].

2.2.4. Transfer of personal data to third countries

Further obligations arise on the condition of data processing and transfer outside of the European Union (EU) can take place, also known as 'transfers of personal data to third countries'[29]. The GDPR specifies that personal data can generally only be transferred to third countries if they are expected 'adequate'[30] — that is, if the country implements sufficient data protection regulation that is fundamentally equivalent to that in the EU — or in the case if the data controller can suggest allocated safeguards that the data will be processed in a manner consistent with that law. In any manner, transfers to third countries may only be executed in full compliance with the GDPR [42].

The *aim of the GDPR* is to enhance transparency and empower the data subject to be in control of their own personal data by imposing certain rights on the data subject that transfers into obligations for the controller and processor.

2.3. Summary

As it is described in the previous section, the GDPR applies in all the cases when the personal data is processed in the EU territory, or when the data subject itself is located in the EU. On

[24] GDPR, Article 12
[25] GDPR, Article 13,14
[26] GDPR, Article 15
[27] GDPR, Article 17 and Recital 66
[28] GDPR, Article 16 and Recital 65
[29] GDPR, Article 44
[30] GDPR, Article 45

the ground of this, if the data subject's personal data is processed and stored in the Blockchain, it is affected by the introduced Regulation.

The main principles of the GDPR – Lawfulness, Fairness, and Data minimization, can be assessed in Blockchain on a case-by-case basis, as it's needed to evaluate the legal obligations and relationships imposed by contracts. But for the execution of those principles the corresponding point of contact form the technology's side should be located, which is not always possible, especially with the public blockchains. The principle of Storage limitation directly conflicts with the Blockchain's characteristic of immutability, as the information cannot be removed from the network.

Apart from the introduced principles of processing the personal data of data subjects, the definition of data subject rights is highly relevant for the Blockchain. Referring to the immutability aspect, the data once inserted and stored in the Blockchain cannot be removed. This implies that the GDPR's right to erasure (right to be forgotten) cannot be realized with blockchain applications and products, as it will require the deletion or alteration of data inside the system, which contradicts the basic idea of Blockchain.

Blockchain may be specified as the distributed network where the data is located on the computer of every participant. The concept of data transferring is one of the main for the technology. At the same time the GDPR focuses on the territorial scope with the aim to ensure that the data is processed according to approved regulations, the data controller is supposed to be familiar with all the locations of Blockchain participants. This is very difficult to implement, and almost impossible with the public blockchains.

The data distribution within the network and the impossibility to make constant changes are basic concepts of the Blockchain technology. As was mentioned before the processing of personal data must comply with certain principles according to the GDPR[31], which apparently leads to the number of tensions to be further discussed.

3. Systematic Literature Review

The present study aims to investigate the scope to which the GDPR and the obligations it creates can be applied to the Blockchain solutions. The work is conducted in the form of a systematic literature review (SLR). The systematic literature review seeks to obtain data that will help to investigate the Blockchain's compliance with the GDPR and their interconnection.

The methodology of systematic literature review proceeds the guidelines proposed by Kitchenham [24], which is composed for software engineering researches. Systematic literature review contains three phases: planning the review, conducting the review, and reporting the review. The first phase is associated with the identification of the need for the review, and development of the review protocol. The second phase is related to the identification of research, selection of primary studies, data extraction, and monitoring, and data synthesis. And the third stage is recognized for the formatting and evaluation of the report. The following section particularizes the first phase of the SLR.

3.1. Systematic Literature Review Protocol

[31] GDPR, Article 5

The main goal of the SLR is to recognize studies where the interconnection of the GDPR and Blockchain solutions is analyzed and discussed. The systematic and methodical summary of all existing information about the Blockchain compliance to the GDPR is required in a thorough and unbiased manner. The SLR uses a trustworthy, rigorous, and auditable methodology of the research topic's evaluation when comparing to the provisional reviews.

3.1.1. Research Questions

The current SLR aims to identify the published studies in regard to the GDPR and Blockchain solutions interconnection. In order to recognize the primary studies where the GDPR concerns are addressed with the relation to the blockchain solutions, it is needed to decompound the research objective to the number of research questions.

As the scope of the study is to analyze the implications of the GDPR's introduction to the Blockchain solutions, the following research question (RQ) is formulated:

RQ: How can Blockchain-based solutions be compliant with the GDPR?

To answer this research question, the following sub-questions (SRQ) are to be addressed:

SRQ1. What are the fundamental aspects of the GDPR that affect Blockchain solutions?

SRQ2. How Blockchain solutions are affected the GDPR?

SRQ3. How can Blockchain solutions become GDPR-compliant?

3.1.2. Search strategy

The search strategy aims to detect as much of the relevant scientific studies and literature as possible. Published studies to be identified using a search strategy for the primary and secondary search. For the primary search, the respective keyword strings on several electronic databases are planned to be used. Pursuing the primary exploration, the secondary search is planned to be handled with the use of backward and forward lookouts [25].

For the comprehensive primary search to identify the primordial set of papers to be observed, the following keyword string (ST) is specified:

ST: ((“Blockchain” OR “Distributed Ledger Technology”) AND “GDPR”)).

The electronic databases, being a predominant source of literature collection [26], are selected based on the inclusion of journal papers, conference reports, workshops' documents in the domains of computer science, cryptography and the legal regulation:

- Google Scholar
- ScienceDirect
- Web of Science
- Scopus (including SpringerLink)
- ACM Digital Library.

After the identification of the extensive list of potentially pertinent papers with the use of primary search, the secondary search is to be conducted. In order to recognize the respective

relevant papers, the backward and forward search approaches are planned to be used. The final list of papers to be arranged for further analysis basing on the application of inclusion and exclusion criteria

3.1.3. Study selection criteria

Formulation of the selection criteria is required to determine the studies where the relevant information to discuss the research questions is provided. The following inclusion (IC) and exclusion criteria (EC) are supposed to be used for the selection of papers:

IC1: Paper discusses aspects of Blockchain within the context of the GDPR

IC2: Paper specifically discusses the GDPR

Taking into consideration the aim of the study, it is required that papers consider the Blockchain solutions and its aspects in the context of addressing the GDPR, thus the first criterion refers to a study where the technology is presented and reviewed along with taking into account the legal concerns. It is essential that papers cover the fundamental facets addressed in the Regulation and present the studies which are sufficient for further analysis. In this respect, the second criterion aims to reveal the studies that cover the respective aspects of the GDPR.

EC1: Study published before 2015

EC2: Study is not written in the English language

EC3: Full-text version of the study is not digitally accessible

EC4: Paper is a duplicate

The first exclusion criterion is designated to assure the up-to-date researches and reporting to be analyzed, as Blockchain is relatively new, developing and upgrading technology, as well as the GDPR, which has raised into effect only in 2018. The second and third criteria are prescribed to provide the access and apprehensibility of the observed studies. The studies accessible in digital databases and available on the internet are acknowledged to be accessible. Duplicates are the papers with the same or approximately the same title from the same author that occur in various digital libraries. In case of the exact duplicate, only one is included, and in case of duplication of version, the most recent one is included.

3.1.4. Screening procedure

The papers' screening procedure is planned to be organized on the base of primary research studies that were identified, by reviewing the title and the abstract of the proposed study. The reviewing principles are the determined inclusion and exclusion criteria. In case if the observed paper fails to meet the criteria, it is excluded from observation and the other criteria are not analyzed in its respect. Every study from the initial research list is to be audited against the inclusion criteria by following the identical procedure. After the accurate evaluation of the studies available, the final list of papers to be identified for further data extraction [27].

3.1.5. Data extraction strategy and synthesis of the extracted data

Following the recognition of the ultimate list of papers, the relevant data is planned to be extracted. Each of the papers supposed to be read one by one and the extraction sheet to be

updated accordingly. The data extraction sheet includes columns to retrieve all the required information needed for further review and reporting. The results of the summarized and reviewed data to be used for the creation of a decision framework for designing and implementing the GDPR-compliant Blockchain solutions.

3.2. Selection and review of studies

In this section, the author presents the steps and intermediate results received that lead to selecting the final set of primary studies. The section also contains a description of the data extraction strategy and information that was extracted from the papers. Table 1 contains a summary of the number of papers retrieved from each electronic database.

In the first step, I have collected the list of query results from each of the defined sources. All the sources indicated in Table 1 allowed exporting the results, except Google Scholar, and for it, the browser extension was used with the aim to scrap data from the search results[32]. Google Scholar currently has a search limitation showing only the first 1000 results. This means that with a "Results per page settings" of 20, only the first 50 pages were available which amounted to 1000 papers. After retrieving results from the databases, a total of 1891 papers were found from all sources. The results from these searches were filtered through the inclusion/exclusion criteria. I started the reviewing process by using the EC4, and 103 papers were removed, resulting in 1788 papers. Considering the difference in export format across various sources, I used only the following data headers: - title (paper title), - authors (list of paper's authors), - year (year of publication).

Database	***Abbreviation***	***Number of papers***
Google Scholar	GSC	1000
Science Direct	SDI	178
Web of Science	WOS	629
Scopus	SCP	44
ACM Digital Library	ACM	40
Total		***1891***

Table 1: Number of results per database

In the next step, using exclusion criteria EC1 and EC2 13 pre-2015 and 49 non-English papers were removed. As a result of this step total of 1726 papers remained. This was reduced to 1555 after checking the EC3 and removing not accessible studies. In the next step, I have filtered the paper titles using the inclusion criteria identified in Section 3.1.3. At this phase, there were 1246 irrelevant paper titles, resulting in 309 final papers. As a result of the next step, 73 papers

[32] Data Scraper - Easy Web Scraping Chrome extension bit.ly/2IEVRiP

were chosen based on their abstract and potential content relevance. In the last step to ensure that these papers do contain the information that is needed for the SLR and using inclusion criteria used in Section 3.1.3, I further examined each of the 73 papers to assess their relevance. This resulted in a total of 37 papers that were appropriate for the final inclusion. The final list of papers included whitepapers and academic papers. I then reviewed these papers and a total of 5 additional papers were found as a result of backward and forward tracing. In total 42 papers were used for analysis in this thesis.

A flowchart of the strategy implemented is presented in Figure 1. When the abstract of a study was not available, the full article was retrieved and assessed for relevance. All potentially relevant articles were retrieved in full text.

3.2.1. Data extraction strategy

Succeeding the identification of the final list of papers, the respective relevant data was extracted. To ensure an unbiased data extraction strategy, it has been recommended [33] to develop a data extraction form and strategy. The corresponding data extraction form was generated after conducting the screening process, which provided an opportunity to utilize the visions revealed during the screening phase. Three types of data were specified and extracted. The first type of data relates to data about the paper. The second data is related to the context and outcomes of the study. The third type is related to the discussion of the interrelation between GDPR and Blockchain and properties of both that are affected, as well as possible concepts that are described which enable the compliance. Table 2 summarizes the information extracted.

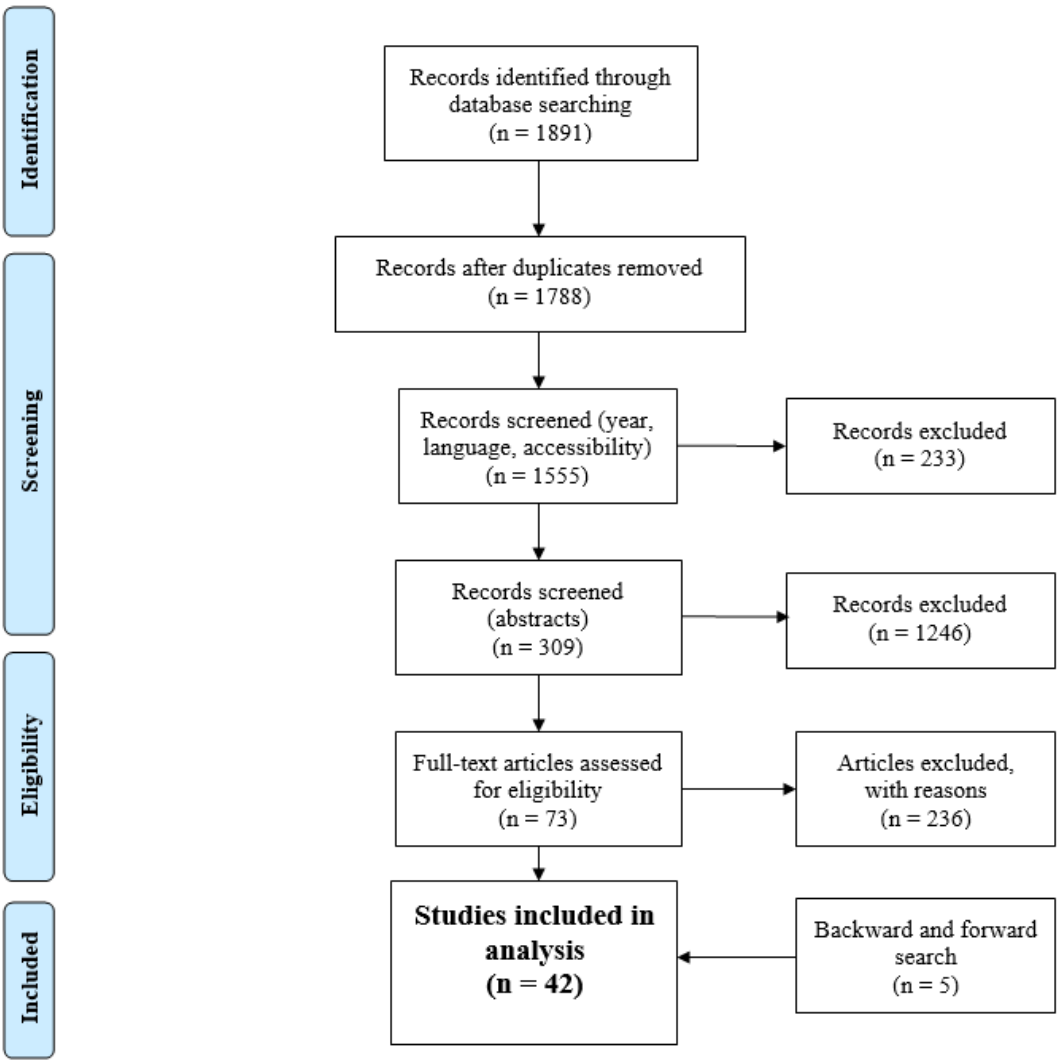

Figure 1. Flowchart of the search strategy

The data from the studies were extracted in a recurrent manner by the author, who extracted the data and filled the corresponding fields of the form.

Data Extraction Form

Data about paper	
Authors	Author(s) of the paper
Title	Title of the paper
Publication year	Paper's year of publication
Type	Type of the paper
Data about the study's context	
Author's specialization	Author's main specialization(s)

Methodology	Methodology used in the study
Main findings and outcomes	Main findings and outcomes of the paper
Data about compliance	
Type of Blockchain	Type of Blockchain discussed/proposed for review
Blockchain aspects	Aspects of Blockchain that affect the GDPR
GDPR aspects	Aspects of the GDPR that affect blockchain solutions
Scenarios	Scenarios/prototypes/designs discussed in terms of interconnection

Table 2: Data Extraction Form

3.3. Overview of studies (report)

This section provides an overview of the reviewed studies. Blockchain in general and its interconnection to the GDPR are relatively new topics in academia. Figure 2 below provides and evidence of the number of publications increased within the last three years, along with the indication of the source of publication/research. The GDPR concept was introduced in 2016 with implication into force in 2018, this also explains the expansion of the interest and clear trend.

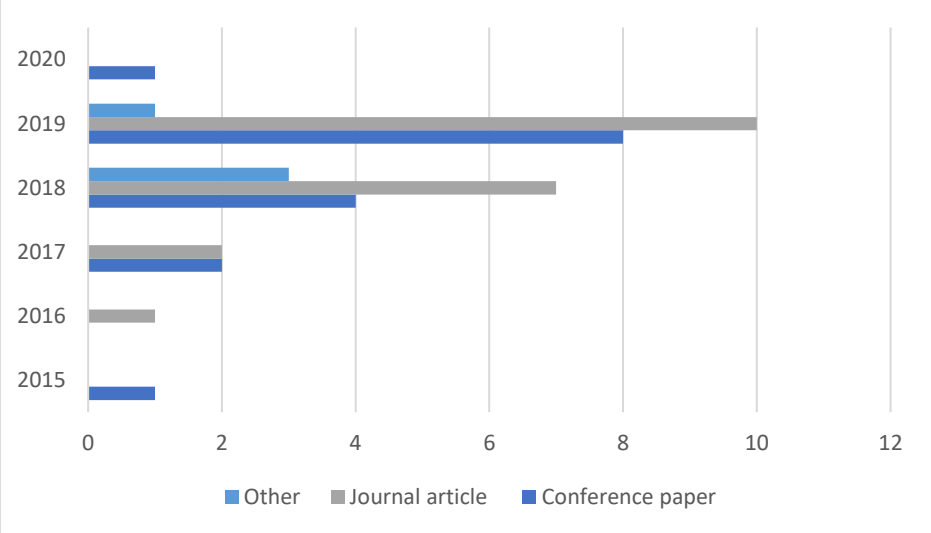

Figure 2. Distribution of reviewed studies by type and by year

Figure 3 represents the distribution of papers among the years with special indication of the authors' specialization – this allows us to observe whether legal specialists or specialists with the industry and technology knowledge are putting more focus towards analyzing the interplay occurred. Results show the almost equal attention from both sides, it might be explained that both, legal and technology specialists, are facing challenges in the interpretation of the

Regulation after its introduction in relation to the technical and technological terms. The topic seems to be very relevant for the understanding of the compliance options.

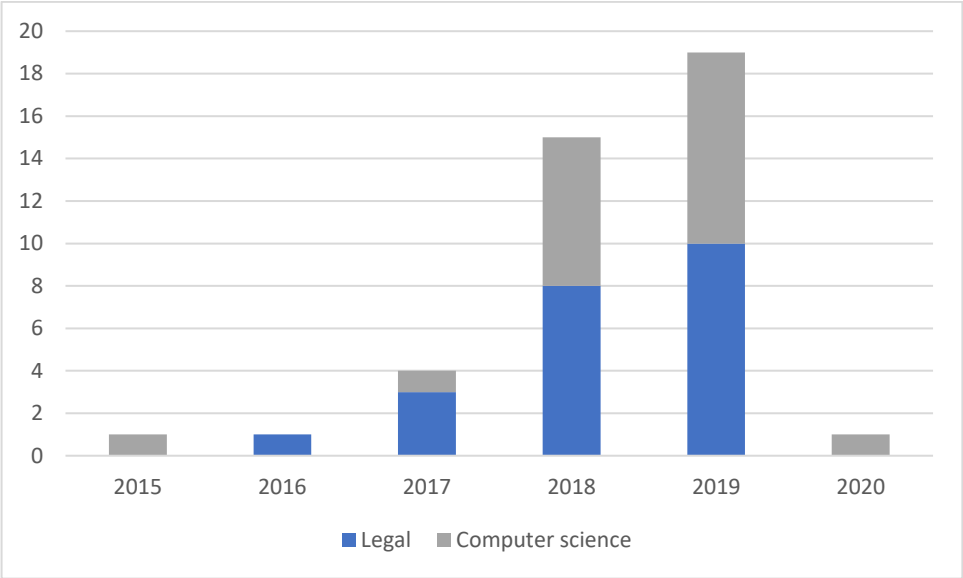

Figure 3. Distribution of reviewed studies by year and author(s) specialization

4. Results

In this chapter of the thesis, the author presents the results obtained from the SLR study with regard to each sub-research question, answering the research question using the information from the state-of-art studies.

4.1. GDPR Aspects that Impact Blockchain Solutions

In this section, the thesis aims to answer the RQ1: What aspects of the GDPR affect the Blockchain solutions?

In this thesis, papers describe the aspects of the GDPR and Blockchain solutions and present the respective interplay and tension that is raised due to the data protection obligations that were introduced by the GDPR. The review identified several aspects of the GDPR that affect Blockchain technology mentioned or discussed in the papers. Almost all the observed papers underlined the aspect that Regulation required the roles of data controllers and data processors to be clearly defined [38], [41], [42], [57], as long as discuss the presented rights of data subjects [40], [50]. Few papers reviewed the so-called "territorial scope" which is referred to the transfer of personal data to third countries [47], [63].

4.1.1. Roles of controller and processor

The GDPR's material scope applies to any processing of personal data[33]. Regulation affects anyone, regardless of a legal entity, that processes or controls the processing of personal data [23]. The factual term utilized by the GDPR is the *controller* of personal data, who is the one, a natural or legal person, public authority, agency or other body which alone or jointly with

[33] GDPR, Article 2

others determines the purposes and means of the processing of personal data[34]. The *processor* of personal data is also subject to the GDPR and is the person who processes personal data on behalf of the controller[35]. In case when the data controller uses a processor, the GDPR oblige of making a contract or other legal act between the two of them guiding the subject-matter [53], [56]. On account of that, it is essential to discuss the relationship between the controller and a potential processor in each case with the aim to define who will be responsible for complying with the GDPR, what kind of data protection authority will audit the respective compliance. Notwithstanding, the roles of the involved entities can be complex as usually there are many parties processing the same personal data simultaneously or jointly [44, [55]. The GDPR recognizes the situation where controllers also being processors, as well as the growth in outsourcing processing activities [48].

The data subject is empowered by the GDPR to have access to a defined minimum amount of information regarding the use and management of personal data. The respective minimum includes the right to know *who the relevant controllers are* along with their location, the reasons for processing the data in question, and other types of information required to ensure the lawful and transparent processing of the data[36]. Apart of the data controllers and data processors, a Data Protection Officer (DPO) is supposed to be designated as an officer of the firm with corresponding "expert knowledge of data protection law and practices" [74], [75]. According to the GDPR, there are also changes in the requested forms of demonstrating compliance. Companies are supposed to indicate the presence and implementation of a profound data compliance program in case of an audit, and not just making the notifications with the data protection authorities. Companies will need to build up, introduce, and maintain policies, documentation and reports, rules, and contracts that prove the compliance with the GDPR requirements. These obligations are expected to be extended to both data controllers and processors.

In terms of the processing actions, there are a number of liabilities to follow by controllers: record information in regard to the goals of processing, groups of data and data subjects, the types of recipients to whom the data will be disclosed, and the possible time limit for the erasure of the different groups of data. During processing definite types of personal data there might occur a high risk to the data subject, and the controller is entitled to conduct a Data Protection Impact Assessment (DPIA) delineating and evaluating the base for preventing feasible blanks and breaches in the system. Despite that, companies are also supposed to maintain a record of recipients and the documentation of the appropriate safeguards while transferring the respective data. Further responsibilities include retaining recorded descriptions of the technical and organizational measures taken for the protection of personal data security, keeping testimonies for the process of selecting data processors, and carrying the copies of contracts with data processors. Controllers may acquire data mapping techniques with the aim to show the aforementioned information. Data mapping can be used by the companies not only for recording the respective information requirements but also in evaluating the risks of data processing activities, tracking the data flows appeared. Along with that, other applicable technical and organizational measures should be introduced to review and defend from any probable risks [48].

[34] GDPR, Article 4
[35] Ibid.
[36] GDPR, Recital 57

4.1.2. Data subject's rights

With the introduction of the GDPR new rights are granted for the data subject, as well as some rights that currently exist are strengthened. Hereof, in an effort to show compliance with the Regulation, data controllers will need to acknowledge the rights of the data subject in case he decides to exercise his rights. The rights provided to the data subject are still open to some interpretation, and the scope of the constraints and prosecution is yet to be defined [49].

With the right to access the data subject empowers the possibility to verify that the controller is acting licitly, hence it is enforcing the data subject's rights[37] [52]. The right to access represents the responsibility of providing information under Article 13 and 14 and provides the data subject with the right to receive in-depth information from the controller about his personal data processing. In its turn, the controllers are compelled to provide the data subject with comprehensive information about the processing and storage of their personal data according to Article 12 in the GDPR[38]. The communication of such information must be made in a "concise, transparent, intelligible, and easily accessible form, using clear and plain language" [50]. This large-scaled obligation follows the aim to improve trust and clarity while processing personal data[39].

The right to erasure, commonly introduced as '*the right to be forgotten*', authorizes data subjects to be in control of their personal data in the conditions and environment that are not adapted towards forgetting [58]. The review found that the vast majority of the studies underline the fact that this fundamental right is in direct conflict with the Blockchain. Following the knowledge that the key security feature of the blockchain-data-structure is determined by the previous transaction data to be stored, taking out a piece of data from the structure would depict the rest of the data to become useless. This introduced right to be forgotten is a subject matter for the extended discussion within the past few years, taking into consideration fast technological development, as well as cases when big corporations use personal data for monetary benefits. In order to have more strict control for the usage of data subject's personal information, the GDPR is prosecuting the right to erasure and enlarging the scope to impose the controller, who is supposed to be responsible for making the personal data public, to erase all copies, replications, or links to, such personal data[40].

Another fundamental right introduced by the GDPR, the right to rectification, is also based on the view of entitling data subjects to be in control of their own personal data, imposing the controller with the obligation to correct, on demand and without undue delay, inaccurate personal data concerning the data subject[41]. Rectification implies that data should be updated to be factual. As was described above, a new transaction can always amend incorrectness in a previous transaction, but the previous transaction will stay and cannot be removed. Basing on Article 16 of the GDPR, it is not generally clear if the improper data must be deleted for compliance [67]. Nevertheless, the primary studies reviewed [68] emphasize the fact that the blockchain does not operate like a traditional register where new data takes the place of the old data.

[37] GDPR, Recital 63
[38] GDPR, Article 12
[39] GDPR, Recital 58
[40] GDPR, Recital 66
[41] GDPR, Recital 65

The described rights demonstrate not only the empowerment of data subject's control for their personal data but also a measure to eliminate infringements of the law, performed by a controller with the personal data [23], [59].

4.1.3. Territorial scope

Along with the above-mentioned aspects, several studies [65], [66] review and discuss Articles 44 through 50 of the Regulation, which directly deal with the governance of cross-border transfers of personal data. Cross-border data transfers may occur in cases when the transfer is processed to a country or international organization with adequate levels of protection or the exporter has a lawful transfer mechanism[42]. The adequacy of the protection level is determined by the European Commission, which means that the corresponding level of protection is determinately equivalent to the EU law's level of protection[43]. Those decisions concerning adequacy are subject to periodic reviews where the Commission may appeal, amend, or suspend a nation's adequacy decision. For the date of the performed research, the Commission has issued 12 adequacy recognitions. While the GDPR provides a defined number of countries with the adequacy of protection designation, it does not prevent cross-border transfers to the countries without an adequacy designation considering that the controller or processor uses appropriate safeguards[44] [67]. The named appropriate safeguards include Binding Corporate Rules (BCRs), standard data protection contractual specifications, an approved code of conduct, an approved mechanism of certification, or legitimately binding channels between public authorities[45].

Taking into account the described requirement, the fact that the cross-border transfers can be problematic if personal data is stored on a permissionless blockchain appear, and it may also be an issue for permissioned blockchain networks if their scope is global, as it is often the case [42], [53], [60].

4.1.4. Summary

The presented section (4.1) introduced an overview of the important aspects of the GDPR that were introduced by the Regulation and affect the data processing in Blockchain-based solutions. The specific sub-aspects discussed in the literature were declared in conjunction with the characterizations and criticism when applied. The revision of literature introduced three main aspects for the analyses and consideration when understanding how the Regulation impacts the development of Blockchain solutions. The following Table 3 provides an overview of the main aspects of the GDPR that impact Blockchain solutions. Each of the recognized aspects is taken into consideration when analyzing the possible concepts' development for Blockchain-GDPR compliance.

The scope of the GDPR applies to any processing of personal data. Regulation affects anyone, regardless of a legal entity, that processes or controls the processing of personal data. The data subject is legitimized to have access to a defined minimum amount of information about the use and management of the personal data, including the right to know *who the relevant controllers and processors are* along with their location, the reasons for processing the data,

[42] GDPR, Article 44
[43] GDPR, Article 45
[44] GDPR, Article 49
[45] GDPR, Article 46

and other types of information with the aim to ensure the lawful and transparent processing of the data. With the decentralized, distributed architecture of a Blockchain, this can be considered as an issue, especially for the public networks. The data subject *rights* implemented in the Regulation not only gave the data subject control for their personal data but also eliminated possible violations of the law from the side of a controller. Taking into consideration the limitations for *cross-border* data transfers, it may become problematic if personal data is stored on a permissionless blockchain, and it may also be an issue for permissioned blockchain networks if their scope is global, as it is often the case.

Aspect	Sub-aspect	Description	Papers
Roles of controller and processor	Defining controller	Controller - determines the purposes and means of personal data processing. The data subject has a right to know who is the controller and his location, in order to contact, provide/withdraw consent for processing personal data, request data's change/removal. Might be an issue with distributed systems, especially with public blockchains.	[36] – [38], [40] – [48], [51] – [66], [70] – [72], [75]
	Defining processor	Processor - processes personal data on behalf of the controller.	
	Compliance	GDPR: contract or other legal act between controller and processor guiding the subject-matter. Companies to develop and implement a profound data compliance program (policies, documentation, reports), description of the technical and organizational measures taken for the protection of personal data security.	
Data subject's rights	Right to access	The data subject has a possibility to verify that the controller is acting justly, right to receive in-depth information from the controller about his data subject's personal data processing.	[47], [48], [52], [55], [56], [61], [67], [71]
	Right to erasure (Right to be forgotten)	More strict control for the usage of the data subject's personal information. The data subject may request from the controller to erase all copies, replications, or links to	[37], [39], [47], [51], [53], [55], [56], [59], [61], [64],

		the personal data. In direct *conflict* with the Blockchain.	[67], [71], [73], [74]
	Right to rectification	Data subject to control use of his personal data. The controller has an obligation to correct on demand inaccurate personal data concerning the data subject. Blockchain allows to amend the incorrect data, but the previous transaction cannot be removed, the new data doesn't take the place of the old data (like in traditional register).	[37], [48], [58], [59], [64], [73], [74]
Territorial scope	Cross-border transfer of personal data	Cross-border data transfers (outside EU) may occur when the transfer is processed to a country/organization with adequate levels of protection or exporter has a lawful transfer mechanism. Can be problematic, as the scope of Blockchain is global.	[36], [38], [40] – [42], [45], [47] – [49], [52], [63], [65] – [67]

Table 3. Aspects of the GDPR to impact Blockchain solutions

4.2. Aspects of Blockchain solutions that are affected by GDPR

Understanding the Blockchain characteristics and properties is crucial when analyzing the interconnection with the GDPR. In this section, the thesis aims to answer the RQ2: What aspects of the Blockchain technology are affected by the GDPR?

4.2.1. Decentralization

The GDPR was drafted in a world and system with a clear division of responsibilities between controllers and processors. The company acts as a data controller and is responsible for handling its users' data. In the defined cases they can pass on these responsibilities to any third-party data processors with conducting a specific data processing agreement [45], [55]. As several reviewed studies point out [38], [45], the Blockchain model distincts from the initial scheme contemplated by the data protection laws due to it is a *decentralized* based technology in which a various number of parties/individuals are directly engaged in the processing of personal data.

With such key characteristic of the technology, Blockchain networks involve an extensive amount of actors, varying from 'protocol developers who create and maintain open-source blockchain technology' to 'actors who run the blockchain protocol on their computers in order to act as validating nodes or participating nodes' or 'network users who sign and submit transactions to the blockchain network via a node' or 'publishers of smart contracts' [38], [45]. In Blockchain's technological reality, *the role of each actor is not precisely determined* and can range depending on the private or public nature of the Blockchain network that is in consideration. Consequently, it appears challenging to define who is nominated as the

controller within the meaning of the GDPR, namely the person who determines the objectives and means of the processing. Following the issue, several studies, including the EU Blockchain Observatory and Forum report, consider that 'protocol developers' and 'actors who run the blockchain protocol' are doubtful to be qualified as data controllers [42], [46], [53]. The users of the particular network may be considered as controller 'if they submit personal data to the Blockchain ledger as part of business activity', but in case of providing the personal data with the intention of the personal use, then the household exemption will be applied and the users cannot be treated as data controllers [42], [46], [53]. Analyzing the publishers of smart contracts, it is indicated that the debate exists regarding how the software should be recognized: managed by its publisher, by the network user calling it or by both [42], [46], [53]. Difficulty in defining who is the controller in terms of a Blockchain network is questionable as the controller delivers the main obligation for complying with the requirements specified by the GDPR, according to the principle of accountability. The principle of accountability is one of a core principle since it accredits data subjects to determine the person responsible for handling their requests and enforcing their rights [42], [53], [60]. Apart from that, the decentralized structure of Blockchain networks also discords with the principle of the lawfulness of processing, specifically, if the processing of personal data is based on consent by reason of the data subject does not know to whom he is giving the consent.

The shortcoming in identifying the entity or entities in the position of 'control' on the permissionless blockchain applications, being a feature of this technology, has generated the question of whether this technology is subject to the GDPR at all. Berberich and Steiner [47] suggest that in a case where the notion of data controller indicates any actual control over the information, two consequences may be feasible: either no node is characterized as a data controller within the meaning of the GDPR or every node where the copies of the distributed ledger occur. At the same time authors highlight that none of those consequences are meaningful and due to the fundamental uncertainty of the regulators' approach to this the entities that use the blockchain as the infrastructure run the risk of walking in the not fully researched and described area. Another opinion is provided by De Filippi [8] who suggests that "the responsibility of keeping data private merely shifts from the operator to the individual user" in multi-node systems. In such a situation, nominating the accountability for one's own data to oneself is, by definition, consistent with the EU's consumer empowerment trend within the data protection and privacy[46]. Notwithstanding, that in practice this concept necessarily counts on well-informed users with advanced knowledge of modern technologies.

When analyzing and discussing the point, the CNIL's (Commission Nationale de l'informatique et des libertés[47]) guidance [45], [76] on Blockchain acknowledges those who have some sort of administration capacities over the chain (which capability to give access, send data for validation and add data) as data controllers. At the same time they are not considered as data controllers if the Blockchain is used only by individuals who are not using the technology for business purposes (e.g. a group of family members using it to build a database of their family tree –for purely household purposes only). The CNIL also proposes to establish legal entities if a group of natural persons aims to use Blockchain technology to achieve a common goal (so

[46] GDPR, Recital 7

[47] National Commission on Informatics and Liberty, independent French administrative regulatory body whose mission is to ensure that data privacy law is applied to the collection, storage, and use of personal data

they will jointly conclude the reasons why and the way how personal data will be processed) [45], [76]. In this context, the responsibilities of each party must be determined and a solution for this is putting in place a contract between all parties. The contract concluded should regulate the responsibilities of every participant when it comes to data protection obligations.

As described above, the current state-of-art is likely to generate an accountability gap which is not ideal for the regulators or citizens. As a consequence, at least for some period of time, it is more likely for the regulators to designate at least some degree of accountability to the entities who are using decentralized architectures with respect to the protection of data and privacy initially [64].

4.2.2. Immutability

One of the key characteristics of Blockchain technologies is that once data is stored, it cannot be altered without leaving a clear trace behind, which signifies that once stored, the data is *immutable*. Regarding that the EU Blockchain Observatory and Forum report indicates that 'the whole point of such a Blockchain is to ensure that transaction including the parties involved, are never forgotten in order to enable decentralized trust' [42], [53], [60]. Such a feature of immutability clashes with some of the GDPR principles, such as the principle of data *minimization*, which dictates the personal data in process should be limited to what is necessary, as well as the principle of *storage limitation* which obliges to keep personal data no longer than needed. In addition to this, the property of blockchain technology also conflicts with the right of the data subject to request the *erasure* of their personal data [42], [48], [53].

As it was presented before, the GDPR implemented the right to erasure which gives the data subject the right to request the deletion of their personal data. While analyzing the key features of Blockchain, it is pointed out that this fundamental right stands in direct conflict with Blockchain technology. The key security feature of the blockchain-data-structure relies on previous transaction data to be stored, that is why removing a piece of data from the structure would interpret the rest of the data useless [68]. Moreover, being mindful of the mathematically implemented structure, the processing power that is required to delete something from the ledger is nearly technologically unattainable, making it close to unfeasible for any controller or processor to do so. With the Blockchain solution, you can conjoin a new transaction that overrides the old transaction data, but at the same time, the old transaction data will always be a part of the ledger [61], [68]. Regardless of the grounds for which the data subject requires the controller or processor to erase the personal data, it appears to be not possible according to architecture, which in turn means that the obligations of the GDPR cannot be respectively fulfilled.

The identical settings employ when the data subject requires that his personal data be rectified. Basing on Article 16 of the GDPR, it is not made clear if the inaccurate data have to be deleted for compliance. Nevertheless, as the Blockchain does not function as a regular register, the only way to rectify information is by adding a layer that displays the old data to be outdated. In such a situation by Blockchain standards, the data has been rectified and corrected and the data subject has the power to be able to request the rectification which consecutively means that a literal interpretation could probably result in the provision being satisfied. However, a strict interpretation would presumably oblige that the outdated data to be removed, which is technically inexecutable on the Blockchain.

4.2.3. Anonymization and pseudonymization

According to the GDPR requirements, the personal data that is under processing should be fully *anonymized* with the purpose of the company/use-cases and application to be out of the scope of the Regulation. The techniques that are used in the Blockchain are classified as the provision of *pseudonymization*, the encryption of personal data is not helping to deceive the problem [49], [58]. The GDPR provides that pseudonymity "means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject *without the use of additional information*, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person"[48] [69]. Encryption, hash-function, keyed-hash function with the stored key, deterministic encryption, or keyed-hash function with deletion of the key and tokenization are listed as pseudonymization techniques [69].

While the GDPR does not apply to anonymized personal data, the criteria of anonymized introduced by the GDPR are hard to fulfill [42], [55]: the anonymization technique should be exceptional enough to make it unfeasible to identify a natural person through any of the means possible to be used, but the process must also be irreversible. It should not be achievable to reconstitute the original data from the anonymized form. The studies reviewed [4], [61], [64] the different techniques available to anonymize data (obfuscation, encryption, aggregation), but it does not identify any technique in the current state of the art which by itself could attain the level of anonymization required by the GDPR. The underlined problem with encryption is that it always leaves the metadata accessible [4], [61], [64]. Nevertheless, the GDPR compliance obliges the implementation of state-of-art security measures and privacy-preserving methods to reinforce privacy and data integrity[49].

4.2.4. Summary

This section (4.2) introduced the overview of the various aspects of the Blockchain Technology to be affected by the implementation of the GDPR. The particular sub-aspects identified in the literature were presented along with the descriptions and critical reflections. While reviewing the papers, three main aspects were deliberated with the aim to identify the colliding areas of compliance between the Blockchain technology and the GDPR. The following Table 4 provides an overview of the main aspects of Blockchain Solutions that are affected by the GDPR. Each of the recognized aspects is taken into consideration when analyzing the possible solution for Blockchain-GDPR compliance.

The GDPR was developed and drafted for the centralized entities and system with a clear division of responsibilities between controllers and processors, where the controller is responsible for handling the data of the users. As the model of Blockchain is completely different from the initial scheme and introduces the *decentralized* technology, in which a various number of parties are involved in the processing of personal data, *the role of each actor is not precisely determined* and can range depending on the private or public nature of the Blockchain network. *Immutability*, being one of the key characteristics of the Blockchain, is giving a tension in terms of compliance with GDPR's principles of data minimization and

[48] GDPR, Article 4
[49] GDPR, Article 32, Recitals 76–80

storage limitation, as it represents the impossibility to alter or delete data from the network. The currently recognized techniques to anonymize data (obfuscation, encryption, aggregation), could not attain the level of anonymization required by the GDPR, and encryption algorithms used in Blockchain solutions are classified as *pseudonymization*.

Aspect	Sub-aspect	Description	Papers
Decentralization	Contact for the data subject (controller)	Blockchain model is not the traditional (centralized) scheme with a clear division of responsibilities, due to its *decentralized* based technology where a various number of parties/individuals are directly engaged in the processing of personal data. The role of each actor is *not precisely determined*, challenging to define who is the controller within the meaning of the GDPR.	[36] – [38], [40] – [46], [48], [52], [53], [55], [56], [58] – [61], [70] – [73]
	Discussion on nominating the controller	Opinions of legal and industry specialists vary (individual nodes/those with administrative rights/developers to be data controllers) – further clarification by regulators needed.	
Immutability	Data storage	Blockchain - once data is stored, it cannot be altered without leaving a clear trace behind, which shows that was stored; all the history is kept within the network. It clashes with the GDPR principles: data *minimization* (personal data in the process should be limited to what is necessary), and the principle of *storage limitation* (keep personal data no longer than needed).	[7], [35] – [37], [39], [41], [43] – [45], [48], [50], [53], [61], [63], [66] – [70]
	Impossibility to remove data	Blockchain's architecture is built in a way that even if a new transaction is added it overrides the old transaction data, but at	[7], [35] – [37], [43] – [45], [48], [53], [56] –

		the same time, the old transaction data will always be a part of the ledger. GDPR implemented rights to erasure and to rectification (deletion or amendment of the data), which cannot be correspondently fulfilled.	[59], [63], [66] – [69]
Anonymization	Pseudonymization	The techniques to be used in the Blockchain are classified as the provision of *pseudonymization*, with implemented encryption it always leaves the metadata accessible. According to the GDPR, the personal data that is under processing should be fully *anonymized* to be out of Regulation's scope, but the current techniques could not accomplish the level of anonymization required by the GDPR.	[4], [36], [38], [40] – [42], [45], [47] – [49], [53], [66] – [69]

Table 4. Aspects of Blockchain solutions to be affected by the GDPR

4.3. Compliance of Blockchain solutions with the GDPR

In this section the review tackles following RQ3: How can Blockchain solutions become GDPR-compliant?

During the review of selected studies, it was identified that researchers and specialists from the industry discuss and propose the following approaches to implement Blockchain solutions in a GDPR-compliant way: the Off-Chain storage of personal data, the implementation of editable Blockchain. Along with the mentioned concepts, some authors suggested various scenarios, mainly including the idea of encryption and afterward key destruction. With the intention of answering the research question, every approach is characterized according to the reviewed literature and investigated against the obligations of the GDPR, whilst the corresponding pros and cons are summarized.

The respective review of the papers and reports outlined two main approaches discussed: the Off-Chain storage and implementation of the Redactable Blockchain. Moreover, added scenarios were also identified which are summarized under the type "Other". Each of the approaches is presented in detail below. The following Table 5 indicates the list of the studies analyzed along with the evidence of approaches which are discussed in those.

Solution	Description	Example	Papers
Storage Off-Chain	Corresponding personal data is not stored inside the Blockchain network but is kept "off-chain". In the Blockchain network itself, there is only the reference/hash value preserved to locate the outside storage with the actual data. As all personal data is stored off-chain can be altered/deleted.	For example, the respective data may be stored in some traditional database, or in cloud storage system.	[36] – [38], [42] – [43], [45], [48] – [50], [52] – [56], [58], [60] – [61], [63] – [66], [69] – [73]
Editable / Redactable Blockchain	The scheme proposes to make changes in one or several blocks of the network, add one or several blocks to the actually available chain. Despite being the characteristic of immutability, the opinion is provided that immutability might not be used for all the newly developed Blockchain solutions. The alterations are possible to introduce with the use of so-called "chameleon hash functions". Those alterations can make changes in data without making a change in the hash value in the block.	Proof of implementation the concept is not available, the option is under theoretical evaluation.	[35], [59], [60]
Other	-Additional legal guidance is requested to be provided by the regulatory to avoid misinterpretation of the regulation, as specialists underline the unclarity of formulations and definitions.	French administrative authority (CNIL) provided the Blockchain/GDPR discussion with proposed interpretations and proposals for further research.	[4], [38] – [42], [44] – [47], [49] – [51], [55], [61], [65] – [69], [71] – [75]

	-Use of more advanced encryption technique / Encrypting the data and destructing the encryption key	Relevant critics: Advanced encryption techniques are required to be developed, not available at the moment. Current encryption algorithms might become insecure with the development of technology.	

Table 5. List of identified proposed solutions

4.3.1. Identified approaches

Storage of personal data Off-Chain

In the studies reviewed the approach that is considered in the vast amount of cases is the idea of the Off-Chain storage of personal data with the object of providing the GDPR-compliant Blockchain solutions in terms of processing the data subject's personal data. The interpretation of this approach follows the logic that the corresponding personal data is not stored inside the Blockchain network but is kept "off-chain" (for example, in some traditional database) [4], [54], [56]. At the same time in the Blockchain network itself, there is only the reference/hash value preserved and locating the outside storage with the actual data saved there [51], [58],[68], [69]. Having regard to the requirements of the GDPR imposed, the concept of storing the sensitive information and personal data outside the Blockchain framework is actively advised, as well as locating there the other type of data and big datasets with the aim to allow better performance of the network [70], [71]. The data supposed to be handled in the Off-Chain storage is located outside the Blockchain network which is used for the data subjects' primary transactions. With such design the introduction of a centralized party to control the data's completeness and confidentiality is underlined and discussed, as it appears to become a violation of one of the main characteristics of the Blockchain – decentralization [4], [44], [71]. At the same time as a part of a discourse in regard to the Off-Chain storage, the opinion presented when the Blockchain is used in alliance with a distributed file system. In this case, only the corresponding reference is located On-Chain and it leads to the prevention of using centralized storage [69].

Acknowledging the GDPR's obligations, the concept where the personal and sensitive data is stored outside the network is supposed to deliver the advantages and probable resolution of the tension that occurred between the core characteristics of the Blockchain and the GDPR requirements. The investigated opinions confer that using the above-described procedure could bring the match with the introduced obligations [4], [44], [54]. Nevertheless, the point is that with such architecture the hash values to the personal data will be still kept inside the Blockchain network and will work as a reference to it. Researchers and industry specialists with the current formulation of the Regulation, and lack of explanations in terms of clash points, cannot provide the clear conclusion whether hash values of personal data will be categorized and treated as anonymous data or will be evaluated as pseudonymous data [36]. The question is currently over the consideration with the participation of regulatory [42], [53],

but several authors are already categorizing hashing as pseudonymization, mentioning that important point is the presence of chance to discover the respective link between the personal data and the one in the hashes [4], [63], [65]. The risk of probable recalculation and options to resolve it with the use of a secret key are currently in the deliberation [42], [53].

With the addition of ongoing discussion about the hash properties, the recommendation of personal data storing as an advanced cryptographic commitment is introduced [45], [76]. In the condition when it is not feasible, the keyed has values and encryption principles should be employed. As per the core rights introduced by the GDPR – right to be forgotten and right to rectification, it is indicated that with aforementioned options the erasure of personal data can be executed by deleting the data Off-Chain along with the corresponding key, that was used to generate the hash value and is stored On-Chain, which gives no proof of which data has been hashed and the data locating On-chain could be expressed as being anonymous. In respect to the right of rectification, the same mechanism might be used when the old data is deleted, and corrected data is loaded to the network. In summary, it can be said, that at the moment there is no legal guarantee that working with cryptographic references based on personal data On-chain is a GDPR-compliant concept for personal data processing using Blockchain. Even referenced data may still be perceived as pseudonymized personal data from a legal perspective and erasure or rectification is technically not possible. Of course, these circumstances can change over time with changing legal perspectives on this procedure.

Editable / Redactable Blockchain

A relatively small number of studies present the idea of developing the Editable/Redactable Blockchain [4], [35], [63]. In the article from 2017 [35] by "redactable", the authors propose to make changes in one or several blocks of the network, as well as add one or several blocks to the actually available chain. Such a scheme is presumed to act against the characteristic of immutability, but authors provide an opinion [35] that immutability might not be used for all the newly developed Blockchain solutions and introduced applications.

The immutability of the Blockchain technology is introduced by the hash values and their connectivity to the previous value and the block in the chain. As the change in this scheme, authors propose the usage of the so-called "chameleon hash function" [35]. The features of chameleon hash functions are entitled to be the same as with usual hashes, but the novelty discussed is in regard to making it possible to introduce the alterations. Those alterations can relate to making changes in data without making a corresponding change in the hash value in the block, this will give the preservation of established connection with the previous block in the chain. Researchers illustrate the process as introducing the connection between two blocks which is supposed to be opened with the right key [35].

Within the list of studies under the review, only three of them mention such an approach with reference to the raised tension between the Blockchain technology and introduced Regulation [4], [35], [63]. Proof of implementation the concept is not available, so the conclusion might be done that for the moment this option is under theoretical evaluation. In the meanwhile, despite possible advantages to bringing in terms of removal of data provision, the introduced approach also has the shortcomings. First of all, this might not be applied to the already existent solutions and applications, which provides with the need to establish the concept at the design stage [4]. Also, the existing available old copies of the network will still contain the

respectively redacted data during the same time a compliant Blockchain node accepts the redacted data and delete the old copies in that version [4], [35]. Notwithstanding, that the possibility to alter data is a huge benefit and responsibility and the corresponding measures should be conducted under the GDPR compliance and by trustworthy parties involved.

Despite the general contradiction of fact of data altering to the basic concept of the Blockchain, authors insist [4], [35] this concept might be evaluated for the further research and studies, as the Blockchain technology is also under the constant expansion and improvement.

Other

During the literature review and evaluation, the other proposals for solutions were revealed. Some of the studies [4], [49] underline the fact that in some definitions and explanations of the Regulation provided there is a lack of clarity, as well as not precise formulations. This creates the misinterpretations of the parts of the Regulation, as well as leaves questions towards the possible interconnection with the Blockchain technology. Legal specialists, as well as colleagues from the technological industry, require the regulatory bodies to observe the formulations in question and provide additional guidance [49], [65].
Besides the precision of the formulations by the legal authorities, several studies mention the usage either of more advanced encryption techniques (which is required to be developed), either the method of encrypting the data on the Blockchain which is followed by the destruction of the encryption key. This method presumes that with the usage of available encryption techniques data gets inaccessible when the encryption key is no longer present [4], [7], [53]. Following such an argumentation, this procedure approaches the understanding of the erasure of data. In respect to the GDPR's requirements, such an approach might be criticized, as the encryption algorithms that are currently in use can become insecure in the future with the development of new techniques, and the encrypted data could possibly be decrypted [4], [7], [53]. Controverting that, the anonymization of data should last forever according to the Regulation. Being mindful of the aforementioned, the use of Encryption & Key destruction should not be considered as the main technique for a GDPR-compliant concept.

4.3.2. Summary

This section (4.3) presented an overview of the different concepts identified in the literature along with corresponding descriptions and critical reflections. The systematic literature review disclosed the two approaches being discussed with the purpose to process personal data on Blockchain according to the compliance with the GDPR. Apart from these, additional solutions that are less feasible to be implemented were also described. The revealed probable solutions could be considered more or less applicable to attain the obligations of the GDPR. Table 5 provided an overview of the discussed solutions with the respective basic description of those. Each of the considered concepts its own advantages and disadvantages, which should be taken into consideration when following the decision framework and making the design determination.

At the moment Redactable Blockchain appears to become a concept that may fulfill the GDPR's obligations, as it empowers the potential changes of data to be done directly on a Blockchain. However, this approach contradicts the immutability characteristic of Blockchain and could be used by the network's participants not in a legitimate way. Presuming the GDPR's privacy regulations, Off-Chain Storage can be used together with a Blockchain in order to store

the corresponding personal data of data subjects. Nevertheless, this may require the establishment of the third-party authority, which also negates the initial idea of the Blockchain. The outlined "Other" concepts that were determined during the research provide the questionable perceptions, and it appears rational to wait for additional explanations to be issued by the regulators.

5. Decision Framework

Presented in this section Decision framework aims to support decision-makers, analysts, and/or Blockchain developers in making an informed decision on the design of the Blockchain solutions taking into account the GDPR-related factors in order to make their solutions compliant with the GDPR obligations. The goal of the framework is to summarize the results of the SLR analysis by creation of a coherent and clear model for Blockchain-GDPR compliant decision-making on the basis of current research in this area. The main goal of it to support Blockchain teams in the process of making weighted decisions at the stage of project initiation within the projects which require the processing of data subject's personal data. It should be noted that the author discusses the cases where the Blockchain architecture is used as a possible solution and does not describes whether it should or should not be used at all, thus, the proposed framework sets the *Blockchain architecture as a main and only option*.

After identification of questions which need to be carefully discussed in order to align Blockchain and the GDPR, aspects of both, technology and regulation, that are affected were covered. This indicated the areas which require the further elaboration and highlighted the two main concepts of how Blockchain-based applications and products can comply with the GDPR in part of personal data processing. Each of the proposed approaches could be considered as an option to fulfill the GDPR's requirements. Other than that, the author discusses the other findings revealed during the research process which can be useful for decision-making process. It should be taken into account that one of the main prerequisites for introduction of Blockchain architecture in areas related to personal data' processing is the way how the personal data is handled. The data processing in the EU-based companies should be aligned to European Union GDPR regulations. Even if data is not processed by this company but send to another one based outside EU – the actor processing the data must follow the European Union data protections laws. Otherwise, this process of data circulation would not be GDPR-compliant and can use the proposed framework only after implementing the respective measures for adequate data protection, for decision-making process in order to identify the best possible architectural solution for implementing Blockchain.

One of the two concepts that are widely presented and discussed, is Off-Chain storage of personal and sensitive data which could require the launching of Third Trusted Party (TTP), and this aspect might be considered as a contradiction towards the core idea of the Blockchain. At the same time the alternatives of how to bypass the introduction of a centralized location for data storage and TTP (for example, to reallocate data among the participants of Blockchain (mainly the private one), but this might not be applied to all the cases). As per the point of privacy-preserving, this concept can fulfill the GDPR obligations with introduction of processing data inside the Blockchain together with addition of personal data storage Off-Chain.

Relative to the idea of Editable/Redactable Blockchain, despite having options to become compliant with the requirements of the GDPR by altering the data in the Blockchain itself, it confronts the characteristic of Blockchain's immutability. Other concepts and opinions presented the perceptions of the GDPR that could be acknowledged invalid by the court of law. The additional judgments from regulatory bodies are needed for the further precision of the issue.

The decision framework is designed in the form of the decision tree (Figure 4), which should be used by the Blockchain development team during the decision-making process in the area of defining what type of Blockchain is needed to be introduced from design point of view and in order to comply with the GDPR regulations. The reader should interpret the framework from top to bottom and starts with presenting the steps and questions to be discussed and answered by the team when choosing the type of Blockchain to be used for the business purpose, following the natural flow of questions to be analyzed when developing a GDPR-compliant Blockchain-based solution. The structure of the framework is based on the factors which need to be evaluated both from external sources that are related to the GDPR specifications, and specific particularities of the blockchain project to be developed. In that order framework first offers to decide whether the Blockchain solutions will need personal data to be processed. In case if it is not the issue, then GDPR will not apply. Nevertheless, in most of the cases, the next point should be discussed – if the products fall into EU territorial scope or the measures for the adequate data protection should be taken. Next is the indication of the Trusted Third Party (TTP) to be used. This is followed by defining the probable number of writers and participants of the network. The consecutive stages are related to the recognition and trustworthiness of the writers and participants. Moving forward, the number of possible parties/organizations that are planned to be involved should be evaluated. At the end, when all the relevant questions are considered, investigated, and discussed, the author proposes specific types of blockchain solutions design that can be used in order to answer the specific questions of each situation. This approach allows to make a fully constituted decision about the Blockchain's type being more relevant for designing a solution. After deciding about the type of the Blockchain to be implemented, the possible decisions for the teams to proceed are provided in the form of schemes of the GDPR-Blockchain compliant architectures (Figure 5 and Figure 6). In the following Table 6 the description of steps from decision process is presented.

Step of decision framework	**Description**	**Paths to follow**	
		Yes	**No**
Are all writes known?	When following the decision framework, it is assumed that online TTP is used and multiple writers are presented. The step in question requires the precision	In case if all writers are known, the decision-making team should proceed to next question in the framework for	If not all writers are known, it imposes the fact that *Permissionless Private Blockchain* should have been chosen for the

	whether all writers are known or not. This will define what steps should follow, or indication of blockchain's type to use.	respective consideration.	solution. [7], [41], [45], [53], [61], [68] After defining the type of Blockchain the team can start observing the available options for making the solution GDPR-compliant.
Is public verifiability required?	Following the consideration that all the writers are trusted, current step observes whether it is needed to have public verifiability.	If verifiability is needed, it brings to the fact that *Permissioned Public Blockchain* should be considered by decision team as an architecture. [7], [41], [45], [53], [61], [68] When decided about the architecture's type to use, team will choose from possible solutions for GDPR-compliance (Figure 4, 5, 6).	In case if public verifiability is not required, the decision-making team should proceed to next question in the framework for respective consideration.
Are multiple organizations involved?	After defining that verifiability is not a point for taking into account, step analyzes whether various organizations are involved in transactions and data processing.	In case of answering all respective former questions and coming to understanding that multiple number of organizations are planned to be participants of the information and data processing, then team should consider using the *Permissioned Consortium Blockchain*. [41], [42], [48], [61]	If after all previously analyzed points, it is underlined that there are no multiple organizations to take part in the process and only the defined number of actors, it provides the team with possibility to imply the *Permissioned Private Blockchain* as the architecture for the solution. [4], [7], [41], [42], [45], [53], [61]

		After defining the type of Blockchain to be used, team can start evaluating the possible options for making the solution, that is planned to be developed, GDPR-compliant (Figure 4, 5, 6).	When the type of architecture to be used is defined, team will discuss offered solutions for GDPR-compliance.

Table 6. Steps of the decision tree

As an example, think of a project decision-maker or a blockchain developer exploring the possibilities for introducing an application based on the Blockchain technology and compliant with the GDPR. In the application, the personal data of data subjects are supposed to be handled, which requires the special concept to be developed on order to comply with the Regulation. When deciding about sufficient architecture to be implemented, decision-maker takes into account the following aspects: the appliance of territorial scope – in what countries/regions the application is planned to be launched; whether the TTP is required – are there any entities available; how many participants are expected and whether they are known and trusted; how many organizations are assumed to be involved in the data processing. For defining answers to those subject areas, the framework would support the specialists in their decision making and provide a guideline to choose the type of blockchain to work with. And as the next step, the decision-maker observes the proposed architecture schemes, deliberates with the developer on the options and chooses the specific design to follow and acquire the GDPR-compliant Blockchain solution. The illustration of the decision model (Figures 4, 5, 6) provides visual support for the above-mentioned route.

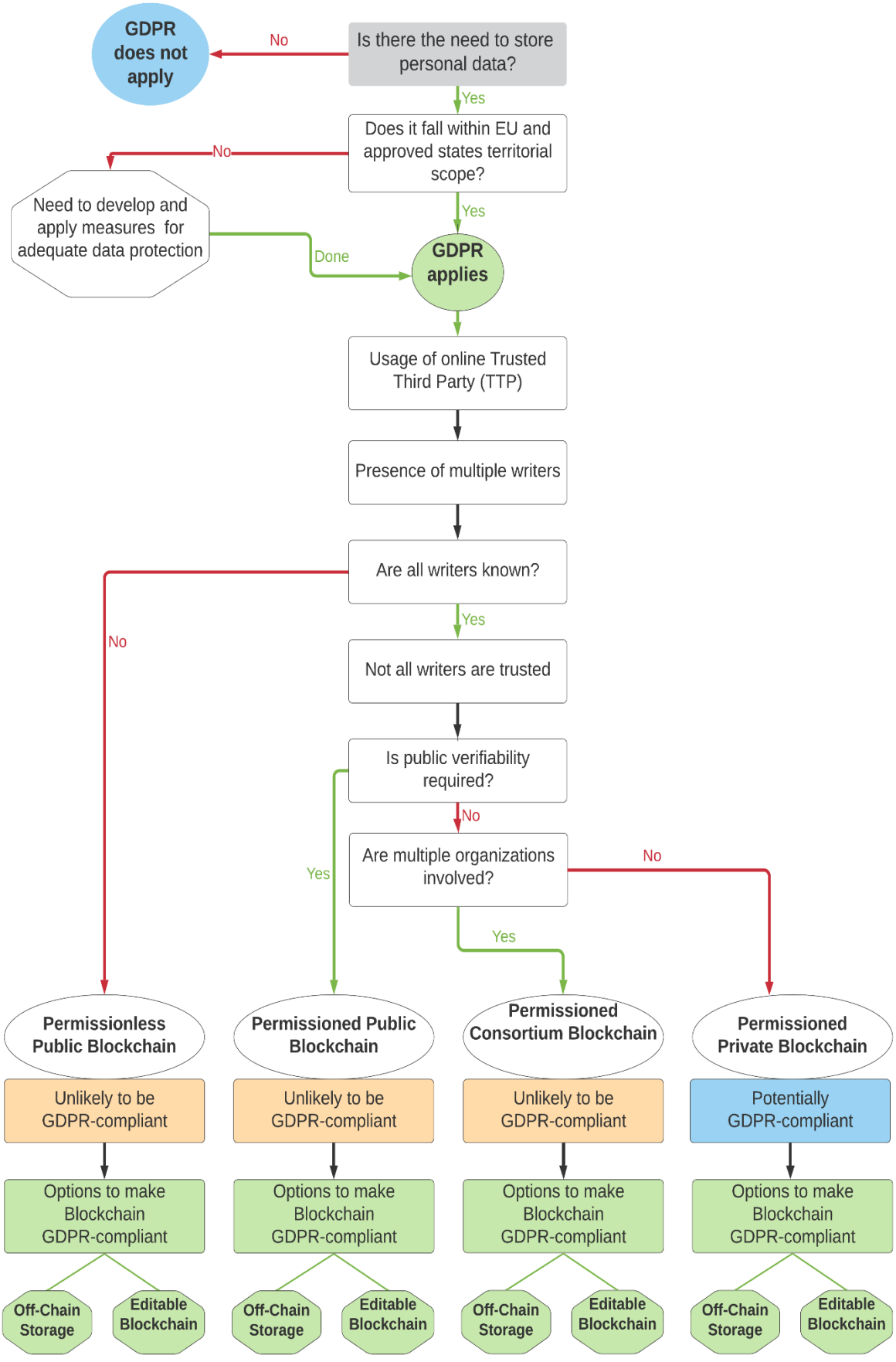

Figure 4. Model for the Blockchain decision

Table 7 provides the respective references towards the proposed solutions for various types of blockchains.

Type of Blockchain	Proposed solutions for GDPR-compliance	References
Permissionless Public Blockchain	Off-Chain Storage	[36]-[38], [41]-[43], [49]-[50], [52]-[53], [55], [60], [61], [68], [71]
	Editable Blockchain	[35], [59], [60]
Permissioned Public Blockchain	Off-Chain Storage	[36]-[38], [41], [49]-[50], [52]-[53], [55], [60], [61], [68], [71], [74]
	Editable Blockchain	[35], [59], [60]
Permissioned Consortium Blockchain	Off-Chain Storage	[35], [37], [38], [41], [48], [60], [61]
	Editable Blockchain	[35], [59], [60]
Permissioned Private Blockchain	Off-Chain Storage	[36]-[38], [41]-[43], [45], [48]-[50], [53]-[56], [61]
	Editable Blockchain	[35], [59], [60]

Table 7. References for decision tree

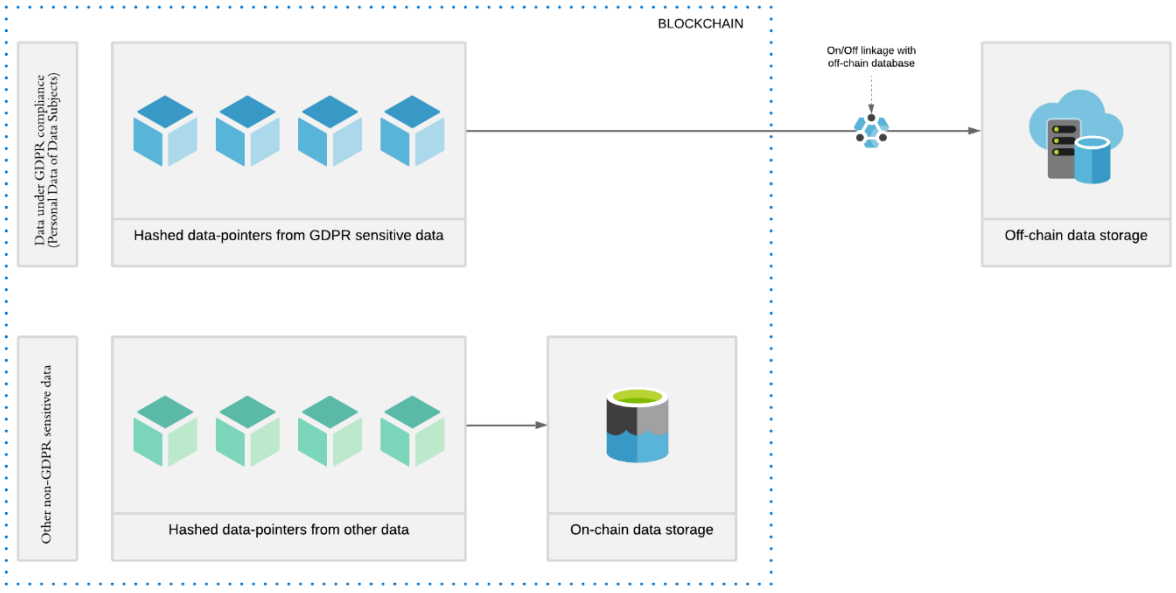

Figure 5. GDPR-Blockchain compliant architecture with the use of Off-Chain storage

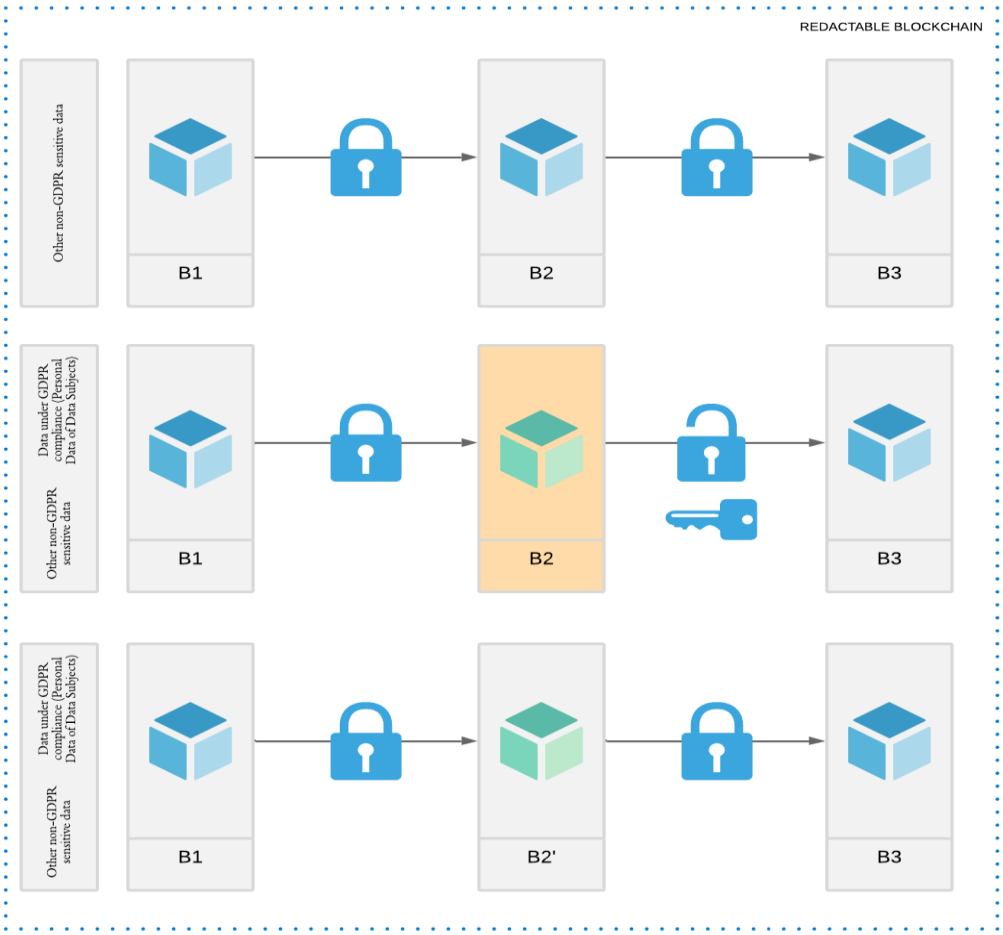

Figure 6. GDPR-Blockchain compliant architecture in form of Redactable Blockchain

5.1. Threats to validity and limitations

In this section, the thesis discusses possible threats to validity (TTV) based on the mapping of possible threats [34]. Threats to validity that apply to this SLR are restricted time span, the bias in study selection and bias in data extraction. Restricted period threat stands for the inability of the author to identify and take into consideration other relevant studies outside the time for which the studies were collected. Blockchain is a constantly developing topic and technology with more information and techniques to appear regularly. In this matter, the author of the thesis could not analyze other relevant studies due to the reason they might appear later on and could not have been included in the list of primary papers. Bias in the study selection threat stands for the subjective assumption of reviewer during the process of identification, review and analysis, and the possibility that not all inclusion and exclusion criteria were used correctly. This bias could have appeared in the present work as the reason of the author's personal and professional knowledge and experience. Since the field is not set and stone in terms of definitions and categories, the author could have introduced bias in the selection of studies specifically concerning papers where the introduction of the GDPR was discussed not in relation to the Blockchain technology. To reduce this type of bias, the author reviewed and assessed the full article for relevance when the abstract of a study was not available, as well as

all potentially relevant articles, were retrieved in full text. Another probable bias to be mentioned is the bias in data extraction, it could be possible that specific invalid or biased outcomes were provided. In this case, to reduce this threat the author used the backward and forward tracing of the studies with the aim to overview all the available opinions and suggestions in the industry.

6. Conclusion

Blockchain has become a highly relevant technological enabler that can be implemented in a various number of use case scenarios. The technology provides several possible options, like the immutability of the information that is stored on the network, as well as the absence of the need to trust all the participants due to the availability of respective consensus mechanisms. Along with the development of technology, the GDPR was introduced in 2018 with the aim to ensure the standardization of data protection policies in the EU. The Regulation established the key principles in regard to processing the personal data, the introduction of data subjects' rights, and transferring personal data to third countries. Blockchain is affected by the GDPR because of material and territorial scope as the processing of personal data is taking place. Compliance with the data protection laws is a crucial element for Blockchain technology both from a regulatory point of view and a business perspective in order to gain the trust of participants and encourage its use. The scientific research explores the points of interconnection between technology and Regulation with the aim to define the possible occurred tension and propose concepts on how the Blockchain-based solutions can be compliant with the GDPR. The conducted review enabled author of the thesis to present a decision framework to support Blockchain project decision-makers and developers in their efforts. In this thesis, a systematic literature review, following guidelines proposed by Kitchenham [24], [27] was used to achieve the objective of the research. As a result, 42 papers were collected and examined in detail.

The review reviled conflict situations may arise when personal data is processed on Blockchain. The tensions between the GDPR and Blockchain rotate mainly around the following issues: 1) The identification and responsibilities of data controllers and data processors; 2) The exercise of a number of data subjects' rights introduced by the Regulation; 3) The anonymization of personal data. Identification of data controller and data processor is required by the GDPR with the purpose to enhance control and traceability on the processing of personal data. While there are some cases where data controllers and data processors can be identified and comply with their obligations, there are also situations when it is difficult, and even impossible, to identify a data controller, especially with a public blockchain. The provided data subjects' rights allow to request erasure or amendment of data, however, if personal data registered and stored in a Blockchain network, it may be problematic to rectify or remove it taking into consideration the concept of the immutability introduced within the Blockchain technology. It is worth mentioning, that precision of what can be considered erasure in the context of blockchains is currently under revision and discussion. Regarding the data anonymization, the intense debates are taking place, without achieved consensus for the moment, on what it takes to anonymize personal data to the point where the resulting output can potentially be stored in a Blockchain network, and what kind of technology developments are required for that. *Two* types of *concepts* are being actively discussed to achieve compliance of Blockchain solutions with the GDPR: Off-chain storage of personal data and designing the Redactable blockchain. Despite both of them logically contradict the basic ideas of the Blockchain, they are considered as the options of possible adjustments to ensure compliance with the Regulation.

Based on the findings of the review, a Blockchain decision framework was developed. The framework consists of seven important aspects. First, it is offered to decide whether Blockchain

solutions will require the processing of personal data. Then the next point should be discussed – if the products fall into EU territorial scope or the measures for adequate data protection should be taken. It follows with the next step of indication the usage of the Trusted Third Party (TTP). The number of writers and participants to use the network should be analyzed. The consecutive stages are related to the recognition and trustworthiness of the writers and participants and defining the parties/organizations to be involved. This approach allows to make a fully integrated decision about the Blockchain's type being more relevant for designing a solution. After determining the type of the Blockchain to be implemented, the two schemes of the GDPR-Blockchain compliant architecture are introduced.

This framework intends to serve as a guide and support for Blockchain developers, project managers and teams in their efforts to make an informed decision in terms of the GDPR-compliant Blockchain architecture. It would enable Blockchain project teams to better understand the available set of options for compliance with the Regulation, but also be able to design new Blockchain solutions that would fit their specialized needs.

Some of the findings in the review have the potential for further examination and future research. In the findings, it was noted that most of the studies were discussing the option of Off-chain storage despite contradiction to the basic concept of Blockchain technology. Given that fact the other observations, including technological changes and developments, might be relevant for design plans and further investigation. Therefore, this topic along with the development of technology stays actual for analysis and can be a valuable route for further research.

References:

[1] Lyons T., Courcelas L., Timsit K. (2018). Blockchain and the GDPR. Report prepared on behalf of the European Union Blockchain Observatory and Forum.

[2] Swan, M. (2015). Blockchain: Blueprint for new economy. O'Reilly.

[3] Rennock M., Cohn A., & Butcher J. (2018, February/March). Blockchain Technology and Regulatory Regulations. Thomas Reuters.

[4] Ibáñez L.D., O'Hara K., Simperl E. (2018), On Blockchains and the General Data Protection Regulation. University of Southampton. https://eprints.soton.ac.uk/422879/1/BLockchains_GDPR_4.pdf

[5] General Data Protection Regulation (GDPR)

[6] Neisse R., Steri G., Nai-Fovino I. (2017). A Blockchain-based Approach for Data Accountability and Provenance Tracking. European Commission Joint Research Centre (JRC)

[7] Van Geelkerken F.W.J., Konings K. (2017). Using Blockchain to strengthen the rights granted through the GDPR

[8] P. De Filippi. (2016). The Interplay between Decentralization and Privacy: The Case of Blockchain Technologies

[9] "How new EU privacy laws will impact blockchain: Expert take." https://cointelegraph.com/news/how-new-eu-privacy-laws-will-impact-blockchain-expert-take, Mar.2018

[10] "Private Blockchains Could Be Compatible with EU Privacy Rules, Research Shows" https://cointelegraph.com/news/private-blockchains-could-be-compatible-with-eu-privacy-rules-research-shows, Nov.2018

[11] Kitchenham B., Brereton O.P., Budgen D., Turner M., Bailey J., Linkman S. (2009). Systematic literature reviews in software engineering – A systematic literature review. Information and Software Technology

[12] Drescher D. (2017). Blockchain Basics: A Non-Technical Introduction in 25 Steps

[13] Width A., P. de Filippi. (2015). DECENTRALIZED BLOCKCHAIN TECHNOLOGY AND THE RISE OF LEX CRYPTOGRAPHIA

[14] Buterin V. (2015). On Public and Private Blockchains, available at: https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/

[15] Nofer M. et al.: Blockchain, Bus Inf Syst Eng 59(3):183–187 (2017)

[16] Zheng Z. et al.: An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends, 6th International Congress on Big Data (2017)

[17] Zheng Z. et al.: Blockchain Challenges and Opportunities: A Survey, Int. J. Web and Grid Services (2017)

[18] Menezes, A., Van Oorschot, P. and Vanstone, S. (1997) Handbook of applied cryptography. Boca Raton: CRC Press

[19] Romano, D. and Schmid, G. (2017). Beyond Bitcoin: A Critical Look at Blockchain-Based Systems. Cryptography, 1

[20] Paar, C. and Pelzl, J. (2010). Understanding cryptography. Berlin: Springer

[21] K.-W. Wong. (2003). A combined chaotic cryptographic and hashing scheme. Physics Letters A 307

[22] Buterin, V., (2017). The Meaning of Decentralization. Available at: https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274

[23] Voigt, P. and Bussche, A. (2017). The EU general data protection regulation (GDPR). Cham: Springer International Publishing

[24] Kitchenham B., Charters S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. Engineering. 2, 1051

[25] Levy Y., Ellis T.J. (2006). A systems approach to conduct an effective literature review in support of information systems research. Informing Science Journal, 9, 181–211

[26] Okoli C. (2015). A guide to conducting a standalone systematic literature review. Communications of the Association for Information Systems, 37, 879–910

[27] Kitchenham B., Brereton O.P., Budgen D., Turner M., Bailey J., Linkman S. (2009). Systematic literature reviews in software engineering – A systematic literature review. Information and Software Technology

[28] Kitchenham B. (2004). Procedures for Performing Systematic Reviews. Keele, UK, Keele University

[29] Randolph J.J. (2009). A Guide to Writing the Dissertation Literature Review. Practical Assessment, Research and Evaluation, vol. 14.

[30] Raj, K. (2019), Foundations of Blockchain, Packt Publishing. E-book. Available online:https://learning.oreilly.com/library/view/foundations-ofblockchain/9781789139396/

[31] S. Wang, L. Ouyang, Y. Yuan, X. Ni, X. Han, and F.-Y. Wang, "Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends," IEEE Trans. Syst. Man, Cybern. Syst., vol. PP, pp. 1–12, 2019

[32] X. Xu et al., "The blockchain as a software connector," in Proceedings - 2016 13th Working IEEE/IFIP Conference on Software Architecture, WICSA 2016, 2016

[33] A. Fink, "Conducting research literature reviews: From the Internet to paper (3rd ed.).," Conducting research literature reviews: From the Internet to paper (3rd ed.). 2010.

[34] X. Zhou, Y. Jin, H. Zhang, S. Li, and X. Huang, "A map of threats to validity of systematic literature reviews in software engineering," Proc. - Asia-Pacific Softw. Eng. Conf. APSEC, pp. 153–160, 2017

[35] G. Ateniese, B. Magri, D. Venturi and E. Andrade, "Redactable Blockchain – or – Rewriting History in Bitcoin and Friends," 2017 IEEE European Symposium on Security and Privacy (EuroS&P), Paris, 2017, pp. 111-126, doi: 10.1109/EuroSP.2017.37

[36] Cruz, Rocio de la, Privacy Laws in the Blockchain Environment (December 15, 2019). Annals of Emerging Technologies in Computing (AETiC), Print ISSN: 2516-0281, Online ISSN: 2516-029X, pp. 34-44, Vol. 3, No. 5, 15th December 2019, Published by International Association of Educators and Researchers (IAER), DOI: 10.33166/AETiC.2019.05.005, Available at SSRN: https://ssrn.com/abstract=3543901

[37] Guggenmos, F., Lockl, J., Rieger, A., Wenninger, A., & Fridgen, G. (2020, January). How to Develop a GDPR-Compliant Blockchain Solution for Cross-Organizational Workflow Management: Evidence from the German Asylum Procedure. In Proceedings of the 53rd Hawaii International Conference on System Sciences.

[38] Duarte, Diogo, An Introduction to Blockchain Technology From a Legal Perspective and Its Tensions With the GDPR (September 30, 2019). Cyberlaw Journal of the Cyberlaw Research Centre of the University of Lisbon School of Law -CIJIC, 2019, Available at SSRN: https://ssrn.com/abstract=3545331 or http://dx.doi.org/10.2139/ssrn.3545331

[39] E. Politou, F. Casino, E. Alepis and C. Patsakis, "Blockchain Mutability: Challenges and Proposed Solutions," in IEEE Transactions on Emerging Topics in Computing, doi: 10.1109/TETC.2019.2949510.

[40] Jaccard, G., Tharin A., GDPR & Blockchain: The Swiss Take (December 4, 2018). Jusletter IT 4 December 2018 , Available at SSRN: https://ssrn.com/abstract=3575231

[41] F. Zemler and M. Westner, "Blockchain and GDPR: Application Scenarios and Compliance Requirements," 2019 Portland International Conference on Management of Engineering and Technology (PICMET), Portland, OR, USA, 2019, pp. 1-8, doi:10.23919/PICMET.2019.8893923

[42] The European Union Blockchain Observatory & Forum (2018): Blockchain and the GDPR. The European Union Blockchain Observatory & Forum. Available online at https://www.eublockchainforum.eu/sites/default/file s/reports/20181016_report_gdpr.pdf, checked on 21/07/2020

[43] Lima, Claudio (2018): Blockchain-GDPR privacy by design. How decentralized Blockchain Internet will comply with GDPR data privacy. Blockchain Engineering Council. Available online at https://blockchain.ieee.org/images/files/pdf/blockcha in-gdpr-privacy-by-design.pdf, checked on 20/07/2020

[44] Rampone, Francesco, Data Protection in the Blockchain Environment: GDPR is not a Hurdle to Permissionless DLT Solutions (May 1, 2018). Ciberspazio e diritto, vol. 19, n. 61 (3 - 2018), pp. 457-20, Available at SSRN: https://ssrn.com/abstract=3383619

[45] GDPR, blockchain and the French data protection authority: Many answers but some remaining questions. Stanford Journal of Blockchain Law & Policy, 2019, 2(2), 1-14

[46] Wirth, C., & Kolain, M. (2018). Privacy by blockchain design: a blockchain-enabled GDPR-compliant approach for handling personal data. In Proceedings of 1st ERCIM

Blockchain Workshop 2018. European Society for Socially Embedded Technologies (EUSSET). DOI: http://dx.doi.org/10.18420/blockchain2018_03

[47] Berberich, M., & Steiner, M. (2016). Blockchain Technology and the GDPR-How to Reconcile Privacy and Distributed Ledgers. European Data Protection Law Review, 2(3), 422-426.

[48] Sater, S. (2017). Blockchain and the European Union's General Data Protection Regulation: A Chance to Harmonize International Data Flows. Available at SSRN: https://ssrn.com/abstract=3080987

[49] Hofman, D., Lemieux, V. L., Joo, A., & Batista, D. A. (2019). "The margin between the edge of the world and infinite possibility" Blockchain, GDPR and information governance. Records Management Journal, 29(1/2), 240-257.

[50] Faber, B., Michelet, G. C., Weidmann, N., Mukkamala, R. R., & Vatrapu, R. (2019, January). BPDIMS: A blockchain-based personal data and identity management system. In Proceedings of the 52nd Hawaii International Conference on System Sciences.

[51] Politou, E., Alepis, E., & Patsakis, C. (2018). Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. Journal of Cybersecurity, 4(1), tyy001.

[52] Bacon, J., Michels, J. D., Millard, C., & Singh, J. (2018). Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralized Ledgers. Rich. JL & Tech., 25, 1.

[53] Van Eecke, P., & Haie, A. G. (2018). Blockchain and the GDPR: The EU Blockchain Observatory Report. European Data Protection Law Review, 4, 531.

[54] Camilo, J. (2019). Blockchain-based consent manager for GDPR compliance. Open Identity Summit 2019. https://dl.gi.de/bitstream/handle/20.500.12116/20985/proceedings-14.pdf?sequence=1&isAllowed=y

[55] Blockchain Bundesverband: Blockchain, Data Protection, and the GDPR, https://www.bundesblock.de/wpcontent/uploads/2018/05/GDPR_Position_Paper_v1.0.pdf, 2018

[56] Compert, C.; Luinetti, M.; Portier, B. (2018). Blockchain and GDPR How blockchain could address five areas associated with GDPR compliance, https://www01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=61014461USEN.

[57] Mirchandani, A. (2019). The GDPR-Blockchain Paradox: Exempting Permissioned Blockchains from the GDPR. Fordham Intellectual Property, Media and Entertainment Law Journal, 29(4), 1201-1241.

[58] Truong, N. B., Sun, K., Lee, G. M., & Guo, Y. (2019). GDPR-Compliant Personal Data Management: A Blockchain-based Solution. arXiv preprint arXiv:1904.03038.

[59] LABANCZ, A. THE CONFLICT OF BLOCKCHAIN AND THE EU GENERAL DATA PROTECTION REGULATION IN THE AREA OF BUSINESS LAW. Law 4.0–Challenges of the Digital Age, 76-81.

[60] Moerel, L. (2018). Blockchain & data protection… and why they are not on a collision course. European Review of Private Law, 26(6), 825-851.

[61] Munier, L., & Kemball-Cook, A. (2019). Blockchain and the General Data Protection Regulation: Reconciling protection and innovation. Journal of Securities Operations & Custody, 11(2), 145-157.

[62] Finck, M. (2019). Smart contracts as a form of solely automated processing under the GDPR. Max Planck Institute for Innovation & Competition Research Paper, (19-01).

[63] Finck, M. (2018). Blockchains and data protection in the European union. European Data Protection Law Review, 4, 17.

[64] Salmensuu, C. (2018). The General Data Protection Regulation and the Blockchains. Liikejuridiikka, 1.

[65] Fabiano, N. Internet of Things and the Legal Issues related to the Data Protection Law according to the new European General Data Protection Regulation. Athens Journal of Law - Volume 3, Issue 3 – Pages 201-214

[66] Buocz, T., Ehrke-Rabel, T., Hödl, E., Eisenberger, I. (2019), Bitcoin and the GDPR: Allocating responsibility in distributed networks, Computer Law and Security Review, 35 (2), pp. 182-198. https://reader.elsevier.com/reader/sd/pii/S0267364918303170?token=D9E304D6D93386F9A6557228E9837B8C40388FB06470C7A64EE55B4009DEDBCB9D828675CA79DBF04F9C772ABD1D75B4

[67] Al-Zaben, N., Onik, M.M.H., Yang, J., Lee, N.-Y., Kim, C.-S. (2019), General Data Protection Regulation Complied Blockchain Architecture for Personally Identifiable Information Management, Proceedings - 2018 International Conference on Computing, Electronics and Communications Engineering, iCCECE 2018, art. no. 8658586, pp. 77-82. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8658586

[68] Schmelz, D., Fischer, G., Niemeier, P., Zhu, L., Grechenig, T. (2019) , Towards Using Public Blockchain in Information-Centric Networks: Challenges Imposed by the European Union's General Data Protection Regulation, Proceedings of 2018 1st IEEE International Conference on Hot Information-Centric Networking, HotICN 2018, art. no. 8606000, pp. 223-228. https://ieeexplore.ieee.org/document/8606000

[69] Zyskind G., Nathan O., Decentralizing privacy: Using blockchain to protect personal data, In: Security and Privacy Workshops (SPW), IEEE, 2015, 180–184

[70] Onik, M.M.H., Kim, C.-S., Lee, N.-Y., Yang, J. (2019), Privacy-aware blockchain for personal data sharing and tracking, Open Computer Science, 9 (1), pp. 80-91. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85064987875&doi=10.1515%2fcomp-2019-0005&partnerID=40&md5=80ee30d5cb1c661a5d5f4510e8a6ed6a

[71] Jambert, A. (2019), Blockchain and the GDPR: A data protection authority point of view, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial

Intelligence and Lecture Notes in Bioinformatics), 11469 LNCS, pp. 3-6. https://link.springer.com/chapter/10.1007%2F978-3-030-20074-9_1

[72] Giannopoulou, A., Ferrari, V. (2019), Distributed data protection and liability on blockchains, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11551 LNCS, pp. 203-211. https://link.springer.com/chapter/10.1007%2F978-3-030-17705-8_17

[73] Sarkar, S., Banatre, J.-P., Rilling, L., Morin, C. (2018), Towards Enforcement of the EU GDPR: Enabling Data Erasure, Proceedings - IEEE 2018 International Congress on Cybermatics: 2018 IEEE Conferences on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing, Smart Data, Blockchain, Computer and Information Technology, iThings/GreenCom/CPSCom/SmartData/Blockchain/CIT 2018, art. no. 8726847, pp. 222-229. https://ieeexplore.ieee.org/document/8726847

[74] Pagallo, U., Bassi, E., Crepaldia, M., Durante, M. (2018), Chronicle of a clash foretold: Blockchains and the GDPR's right to erasure, Frontiers in Artificial Intelligence and Applications, 313, pp. 81-90. http://ebooks.iospress.nl/publication/50837

[75] Dewitte, P., Wuyts, K., Sion, L., Van Landuyt, D., Emanuilov, I., Valcke, P., & Joosen, W. (2019, April). A comparison of system description models for data protection by design. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing (pp. 1512-1515)

[76] CNIL: Solutions for a responsible use of the blockchain in the context of personal data, November 2018. https://www.cnil.fr/sites/default/files/atoms/files/blockchain.pdf

Appendix

I. The full list of selected primary papers

	Title of the paper	Author(s) of the paper	Year of publication
1	Redactable Blockchain – or – Rewriting History in Bitcoin and Friends	Ateniese,G., Magri, B., Venturi, D., Andrade E.(Ateniese et al.)	2017
2	Privacy Laws in the Blockchain Environment	Cruz, Rocio de la	2019
3	How to Develop a GDPR-Compliant Blockchain Solution for Cross-Organizational Workflow Management: Evidence from the German Asylum Procedure	Guggenmos, F., Lockl, J., Rieger, A., Wenninger, A., & Fridgen, G. (Guggenmos et al.)	2020
4	An Introduction to Blockchain Technology From a Legal Perspective and Its Tensions With the GDPR	Duarte D.	2019
5	Blockchain Mutability: Challenges and Proposed Solutions	Politou, E., Casino, F., E. Alepis, E., Patsakis, C. (Politou et al.)	2019
6	GDPR & Blockchain: The Swiss Take	Jaccard, G., Tharin A.	2018
7	Blockchain and GDPR: Application Scenarios and Compliance Requirements	Zemler, F., Westner, M.	2019
8	The European Union Blockchain Observatory & Forum (2018): Blockchain and the GDPR	The European Union Blockchain Observatory & Forum	2018
9	Blockchain-GDPR privacy by design. How decentralized Blockchain Internet will comply with GDPR data privacy	Lima, Claudio	2018
10	Data Protection in the Blockchain Environment: GDPR is not a Hurdle to Permissionless DLT Solutions	Rampone, Francesco	2018
11	GDPR, blockchain and the French data protection authority: Many answers but some remaining questions	Stanford Journal of Blockchain Law & Policy	2019

12	Privacy by blockchain design: a blockchain-enabled GDPR-compliant approach for handling personal data	Wirth, C., Kolain, M.	2018
13	Blockchain Technology and the GDPR-How to Reconcile Privacy and Distributed Ledgers	Berberich, M., Steiner, M.	2016
14	Blockchain and the European Union's General Data Protection Regulation: A Chance to Harmonize International Data Flows	Sater, S.	2017
15	The margin between the edge of the world and infinite possibility. Blockchain, GDPR and information governance	Hofman, D., Lemieux, V. L., Joo, A., & Batista, D. A. (Hofman et al.)	2019
16	BPDIMS: A blockchain-based personal data and identity management system	Faber, B., Michelet, G. C., Weidmann, N., Mukkamala, R. R., & Vatrapu, R. (Faber et al.)	2019
17	Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions	Politou, E., Alepis, E., & Patsakis, C. (Politou et al.)	2018
18	Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralized Ledgers	Bacon, J., Michels, J. D., Millard, C., & Singh, J. (Bacon et al.)	2018
19	Blockchain and the GDPR: The EU Blockchain Observatory Report	Van Eecke, P., & Haie, A. G.	2018
20	Blockchain-based consent manager for GDPR compliance	Camilo, J.	2019
21	Blockchain Bundesverband: Blockchain, Data Protection, and the GDPR	Blockchain Bundesverband	2018
22	Blockchain and GDPR How blockchain could address five areas associated with GDPR compliance	Compert, C.; Luinetti, M.; Portier, B. (Compert et al.)	2018
23	The GDPR-Blockchain Paradox: Exempting Permissioned Blockchains from the GDPR	Mirchandani, A.	2019

24	GDPR-Compliant Personal Data Management: A Blockchain-based Solution	Truong, N. B., Sun, K., Lee, G. M., & Guo, Y. (Truong et al.)	2019
25	The Conflict of Blockchain and The EU General Data Protection Regulation in The Area of Business Law	Labancz, A.	2019
26	Blockchain & data protection… and why they are not on a collision course	Moerel, L.	2018
27	Blockchain and the General Data Protection Regulation: Reconciling protection and innovation	Munier, L., & Kemball-Cook, A.	2019
28	Smart contracts as a form of solely automated processing under the GDPR	Finck, M.	2019
29	Blockchains and data protection in the European union	Finck, M.	2018
30	On Blockchains and the General Data Protection Regulation	Ibáñez L.D., O'Hara K., Simperl E. (Ibáñez et al.)	2018
31	The General Data Protection Regulation and the Blockchains	Salmensuu, C.	2018
32	Internet of Things and the Legal Issues related to the Data Protection Law according to the new European General Data Protection Regulation	Fabiano, N.	2017
33	Bitcoin and the GDPR: Allocating responsibility in distributed networks	Buocz, T., Ehrke-Rabel, T., Hödl, E., Eisenberger, I. (Buocz et al.)	2019
34	General Data Protection Regulation Complied Blockchain Architecture for Personally Identifiable Information Management	Al-Zaben, N., Onik, M.M.H., Yang, J., Lee, N.-Y., Kim, C.-S. (Al-Zaben et al.)	2019
35	Towards Using Public Blockchain in Information-Centric Networks: Challenges Imposed by the European Union's General Data Protection Regulation	Schmelz, D., Fischer, G., Niemeier, P., Zhu, L., Grechenig, T. (Schmelz et al.)	2019
36	Decentralizing privacy: Using blockchain to protect personal data	Zyskind G., Nathan O.	2015

37	Privacy-aware blockchain for personal data sharing and tracking	Onik, M.M.H., Kim, C.-S., Lee, N.-Y., Yang, J. (Onik et al.)	2019
38	Blockchain and the GDPR: A data protection authority point of view	Jambert, A.	2019
39	Distributed data protection and liability on blockchains	Giannopoulou, A., Ferrari, V.	2019
40	Towards Enforcement of the EU GDPR: Enabling Data Erasure	Sarkar, S., Banatre, J.-P., Rilling, L., Morin, C. (Sarkar et al.)	2018
41	Chronicle of a clash foretold: Blockchains and the GDPR's right to erasure	Pagallo, U., Bassi, E., Crepaldia, M., Durante, M. (Pagallo et al.)	2018
42	A comparison of system description models for data protection by design	Dewitte, P., Wuyts, K., Sion, L., Van Landuyt, D., Emanuilov, I., Valcke, P., & Joosen, W. (Dewitte et al.)	2019

II. Licence

Non-exclusive licence to reproduce thesis and make thesis public

I, _____Ilona Pavlenkova_____,

(*author's name*)

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

_____GDPR and Blockchain Solutions_____,

(*title of thesis*)

supervised by _____Fredrik Payman Milani_____.

(*supervisor's name*)

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Ilona Pavlenkova

10/08/2020