FACULTY OF LAW

Lund University

Erik Källman

Blockchain v. Personal Data

– A Rising Conflict Between Technology and the Law?

LAGF03 Essay in Legal Science

Bachelor Thesis, Master of Laws programme

15 higher education credits

Supervisor: Emma Ahlm

Term: Spring 2019

Contents

Summary

This thesis examines the blockchain technology from a General Data Protection Regulation (GDPR) point of view. The focus area is the protection of personal data in blockchains.

Blockchain can be summarized as a shared, decentralized ledger where data can only be added, not removed. In essence, it works like a trust-creator that has the potential to remove the need for middlemen. From a legal perspective, the most relevant areas of use are things like transitions of ownership, derivatives market trades, storage of transaction history, and supply chain management.

The examination shows that most blockchains will process personal data in such a way that the GDPR becomes applicable. This is due to the wide definition of personal data together with the difficulties to successfully anonymize such data. The applicability of the GDPR evokes several responsibilities. This thesis focuses on the different obligations to erase personal data. It is shown that many blockchain configurations will directly violate the different obligations to erase. There may be methods to increase compliance, but none of them is without risk.

Sammanfattning

Denna uppsats undersöker blockkedjeteknologin utifrån dataskyddsförordningen (GDPR). Fokusområdet är skydd av personuppgifter på blockkedjor.

Blockkedjeteknologin kan sammanfattas som en delad, distribuerad liggare där information endast kan läggas till, inte tas bort. Tekniken är ett sätt att skapa tillit, och har i förlängningen potentialen att ta bort behovet av mellanhänder. Från ett juridiskt perspektiv är de främsta användningsområdena äganderättsövergångar, handel med derivat, lagrande av transaktionshistorik och hantering av logistik.

Undersökningen visar att de flesta blockkedjor kommer behandla personuppgifter på ett sådant sätt att dataskyddsförordningen blir tillämplig. Detta beror på den vida definitionen av personuppgifter samt svårigheterna att framgångsrikt anonymisera sådan data. Dataskyddsförordningens tillämplighet väcker flera skyldigheter. Denna uppsats fokuserar på de olika skyldigheterna att radera personuppgifter. Det visas att många blockkedjor inte kommer kunna uppfylla dessa skyldigheter. Det finns visserligen möjligheter att öka kompabiliteten mellan blockchain och GDPR, men alla dessa metoder innehåller risker som inte kan förbises.

1. Introduction

1.1. Background and Research Overview

The introduction of blockchain technology is often compared with the introduction of the limited liability for corporations. It is said that blockchains will fundamentally change the way we do business with each other. Securing decentralized trust is the most important aspect of the technology, this is done through the creation of an immutable ledger that can only be added. In essence, it has the potential to remove the need for middlemen. Instead of using a bank as an intermediary for things like transactions, loans, derivatives market trades etc., peers can interact directly with each other through a blockchain solution. This interaction can take place with complete strangers, with the technique as the trust-creator instead of the middlemen.

The immutability of the blockchain is both the biggest advantage and the greatest threat to the technology. This is due to the General Data Protection Regulation[1] (GDPR), and its different obligations to erase personal data. A violation of the GDPR could result in fines of up to 20 000 000 EUR or up to 4 % of the total worldwide annual turnover.[2] In other words, it is of great significance to ensure compliance.

Since both the GDPR and the breakthrough of blockchain technology are relatively recent, not much research has been done. Existing research often focuses on the issues of identifying controllers and processors in a blockchain environment, while the issue of personal data is often given less attention.[3] Other works only present specific design concepts to increase compliance.[4] This thesis will examine when the GDPR is applicable, what responsibilities applicability evoke and if there are any methods to increase compliance.

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

[2] GDPR, Art. 8.5.

[3] Among others, see Shmelz et al. Towards Using Public Blockchain In Information-Centric Networks: Challenges Imposed by the European Union's General Data Protection Regulation (pp. 223–228). *Proceedings of 2018 1st IEEE Conference on Hot Information-Centric Networking*, 2018; Lyons, Tom et al. *Blockchain & the GDPR*. EU Blockchain Observatory and Forum, 2018.

[4] Among others, see Bayle, Aurelie et al. When Blockchain Meets the Right to be Forgotten: Technology Versus Law in the Healthcare Industry. *IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, 2018; Binh Truong, Nguyen et al. GDPR-Compliant Personal Data Management: A Blockchain-based Solution. *IEEE Transaction on Information Forensics and Security,* 2019.

1.2. Aim and Research Questions

The aim of this essay is to examine if the GDPR risk to counteract the development of blockchain technology. The focus area is the protection of personal data on blockchains. For this purpose, the following questions will be answered:

- Under what circumstances will information in a blockchain be considered personal data under the GDPR?
- Is it possible to comply with the obligations to erase if personal data is stored on a blockchain?

1.3. Delimitation

This thesis will focus only on the issues related to the storage of personal data on a blockchain. Therefore, a number of aspects of the GDPR as well as the blockchain technology will be left out. What will not be discussed includes the territorial scope of the GDPR, organizational requirements under the GDPR, some technical aspects of blockchains, and enforcement of the GDPR. Furthermore, there will be no particular division between public and private blockchains. The GDPR applies to both and the issues of personal data on the blockchain will essentially be the same.

1.4. Method and Material

This thesis is written from an interdisciplinary perspective. Regarding the application of legal norms within IT, the term *Legal Informatics* is often mentioned. It pertains to unite the development of IT with the law, where the law is often not directly adapted to the technology.[5] The aim of using an interdisciplinary perspective is to avoid that the GDPR is observed without regard to the society where it is applicable. This is done to better understand the practical consequences of the GDPR.

The interpretation of the GDPR will be done with a legal dogmatic method. The purpose of the legal dogmatic method is often described as the reconstruction of a legal norm.[6] An often

[5] Cecilia Magnusson Sjöberg (ed.). *Rättsinformatik: Juridiken i det digitala informationssamhället*. 3rd edn. Lund: Studentlitteratur, 2018, p. 27.
[6] Kleinman, Jan in: Nääv, Maria. Zamboni, Mauro (eds.). *Juridisk metodlära*. 2nd ed. Lund: Studentlitteratur AB, 2018, p. 21.

put forward criticism of the method is that it is only interested in the norms themselves, not the application of them.[7]

Since the GDPR is an EU Regulation, interpretations of it must be done in accordance with an EU legal method. One should note that EU Law has its own rules of interpretation and legal principles that all Member States and national courts must follow.[8] Among other things, this means that the European Court of Justice (ECJ) is considered to have exclusive jurisdiction over the interpretation of EU Law.[9] The main methods of interpretation used by the ECJ are grammatical, systematic and purposive.[10] Furthermore, preambles are not legally binding, however, they provide context and purpose to the Articles.[11]

Regarding the material, the following can be said. A variety of sources are used to answer the research questions. Regarding the section about the GDPR, mainly literature and articles from practicing attorneys exists. It is treated with caution. However, significant parts of the GDPR are identical to the Data Protection Directive from 1995.[12] Therefore, much of the case-law and other documents will still have relevance for the interpretation of the GDPR. Regarding the parts that are identical, opinions from the Article 29 Working Party are especially relevant. The Working Party was set up under Article 29 of the former directive, as an independent advisory body.[13] Even though it is not an established legal source, the opinions enjoy a significant amount of respect in the field.

As regards the blockchain technology, mostly non-legal sources are used to describe it. This is to give a basic understanding of the technique before discussing it from a legal perspective.

A report from the EU Blockchain Observatory and Forum is used regarding the discussion about the GDPR and Blockchain. It is an initiative of the European Commission. It is however independent, and the views do not reflect the views of the European Commission.[14]

---

[7] Ibid, p. 24.
[8] Hettne, Jörgen. Otkens, Ida (ed.). *EU-rättslig metod*. 2nd edn. Stockholm: Norstedts juridik, 2011, p. 158–170.
[9] See Consolidated version of the Treaty on the Functioning of the European Union [2012], Art. 19.
[10] Rösler, Hannes in: Basedow, Jurgen. Zimmermann, Reinhard (eds.). *The Max Planck Encyclopedia of European Private Law*. Oxford: Oxford University Press, 2012, p. 979.
[11] Baratta, Roberto. Complexity of EU Law in the Domestic Implementing Process (pp. 293–308). *The Theory of Practice and Legislation*, Vol 2 issue 3. 2014, p. 302.
[12] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and the free movement of such data.
[13] Directive 95/46/EC, Article 29.1.
[14] Ibid.

2. The World of Blocks

2.1. Introduction to Blockchain Technology

There is no established legal definition of blockchain technology. However, some official and semi-official authorities have released various writings concerning blockchain. The EU Blockchain Observatory and Forum defines the technology as follows:

> "*At its core, Blockchain is a decentralized database technology. It allows large numbers of actors, including strangers or even adversaries, to store synchronized copies of the same data. The data is typically organized in the form of an append-only ledger, meaning that data can only be added, not taken out*".[15]

The blockchain technology was introduced with the cryptocurrency Bitcoin.[16] They should, however, not be mixed together. The blockchain technology is the foundation of Bitcoin, but the possible areas of use for the blockchain technology are much greater than just Bitcoin or other cryptocurrencies.[17] The blockchain technology was first developed to enable exchanges in a low-trust environment, to solve the issue with double-spending in ledgers and to create a distributed ledger that is practically impossible to tamper with.[18]

2.2. How the Blockchain Technology Functions

2.1.1. Sharing Information Without a Central Authority

Instead of having a centralized server, blockchain uses a peer-to-peer network. It consists of *nodes* that are *non-hierarchal*. In the illustration of the peer-to-peer ledgers in figure 1 below, each ledger also constitutes a node. Being a node means that it, on equal terms with the other nodes, send and receive information.[19] A centralized ledger is the sender, receiver and keeper of information in the network. A practical example of a centralized ledger is a bank, where all transactions are centrally registered, and all transactions move through the bank. In a peer-to-

[15] Lyons, Tom et al. *Blockchain & the GDPR*. EU Blockchain Observatory and Forum, 2018, p 14.
[16] Beck, Roman. Horst, Treiblmaier (eds.). *Business Transformation Through Blockchain.* Vol 2. Cham: Springer Nature Switzerland, 2019, p. 340.
[17] Ibid.
[18] Nakamoto, Satoshi. *Bitcoin: A Peer-to-Peer Electronic Cash System*, 2008, p. 1.
[19] Drescher, Daniel. *Blockchain basics: a non-technical introduction in 25 steps*. New York City: Apress LLC, 2017, p. 22.

peer network, however, the transactions are made and registered directly with the other peers, without the need for a middleman.

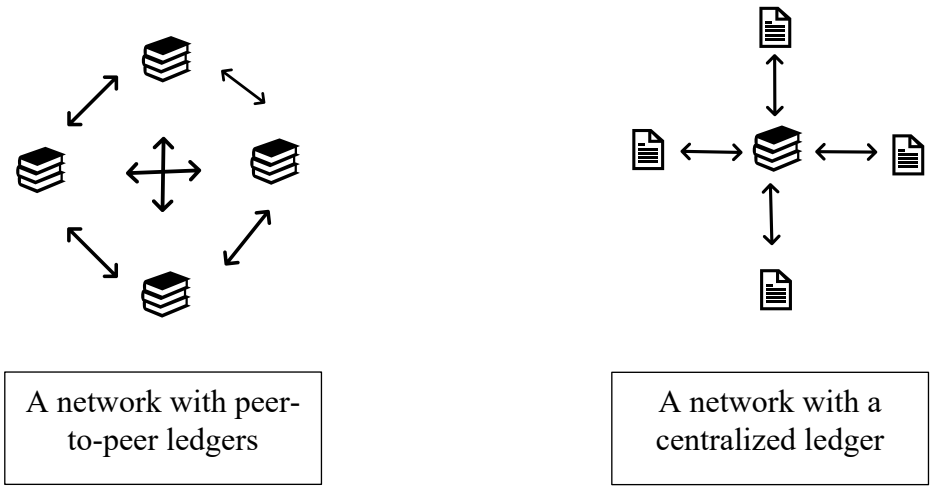

Figure 1. Illustration of peer-to-peer ledgers (left) and a centralized ledger (right).

2.2.2. Making Small Differences Stand Out

One of the most fundamental aspects of blockchain technology is the use of *hash functions*. They can be described as the digital fingerprint of the block.[20] It is the hash that makes the blockchain practically impossible to tamper with.[21] A hash function is a mathematical method that takes any information of any size and produces a fixed length output, known as a hash output or simply hash.[22] Normally, the hash output is much smaller than the inserted information. Every time the same information is put through the same hash function, it produces the same output. On the other hand, if any change is done to the information, the hash function produces a completely different hash output. In other words, a change would not be unnoticed.

[20] Appelbaum, Deniz. Stein Smith, Sean. Blockchain Basics and Hands-on-Guidance (pp. 28–37). *The CPA Journal*. June 2018, p. 30.
[21] Drescher, Daniel. *Blockchain basics: a non-technical introduction in 25 steps*. New York City: Apress LLC, 2017, pp. 84–85.
[22] Appelbaum, Deniz. Stein Smith, Sean. Blockchain Basics and Hands-on-Guidance (pp. 28–37). *The CPA Journal*. June 2018, p. 30.

The cryptographic hash used in blockchains is a one-way function, meaning that the same hash function cannot be used to make the hash output readable again.[23]

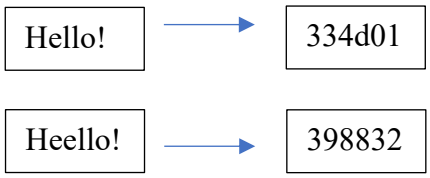

Figure 2. Visualization of a hash function, showing the significant difference in the output with only a small difference in the input. The hash function used is a shortened version of the SHA256function, showing only the first 6 values out of 256.[24]

2.2.3. Creating the Chain

Blockchain utilizes *hash references* to create a virtual chain to connect the blocks. The way it works is that every block contains a hash reference to the hash output of the previous block.[25] The consequence is that if a block would be removed or tampered with, every following block until the latest one would become void.[26]

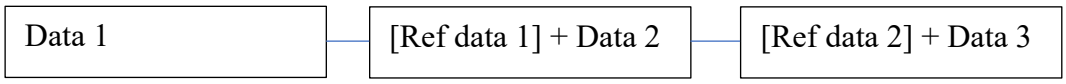

Figure 3. Simplified illustration of hash references.

2.2.4. Identifying the Users

Most blockchains are accessed through platforms. The users can be everything from natural persons to companies to official authorities. Users log in to the platform and the platform communicates with the blockchain. On the blockchain, users are identified through asymmetric cryptography.[27]

---

[23] Ibid.
[24] The SHA256 is a hash function developed by the US National Security Agency (NSA), patented in US patent 6829355.
[25] Drescher, Daniel. *Blockchain basics: a non-technical introduction in 25 steps*. New York City: Apress LLC, 2017, p. 87.
[26] Ibid., p. 88.
[27] Ibid., p. 93.

Asymmetric cryptography uses two linked keys for encryption and decryption.[28] The information encrypted with one key can only be decrypted with the other key and vice versa.[29]

The way asymmetric cryptography is used in a blockchain is that two corresponding asymmetric keys are created for every new user.[30] One is called the *public key* and one is called the *private key*.[31] The public key is announced to everyone on the network, while the private key is known only by the platform from which the natural person accesses the blockchain. The practical consequence is that everyone on the network can verify that information added by public key 'X' is created by the holder of the corresponding private key.

2.2.5. Verifying the information

Information is verified in a blockchain through the use of digital signatures. It can be verified both that the uploaded information is the intended one as well as which user uploaded it.[32] Digital signatures utilize a combination of asymmetric cryptography and hashing to create verifiable digital signatures that make it easy to identify frauds.[33]

The way it works is that if I want to send an authorized message saying 'Hello', I would start by hashing the message. The hash output with the shortened hash function used in figure 1 would be '185f8d'. This would subsequently be encrypted with my private key, giving me a random set of characters, let's say '123456'. My authorized message that is uploaded to the blockchain would contain both my message and the hash output from my message. Now, users in the network can verify that I was the sender of the message by using my public key to decrypt the message that would read 'Hello'. They can also verify the message itself by hashing 'hello' and compare it with the hash output. If the values match, my message is authorized.

2.2.6. Adding New Blocks

When a user wants to add a new block to the chain, the request is sent to the nodes in the peer-to-peer network. Each new block must be approved by consensus or by a majority of the nodes.[34]

---

[28] Ibid., p. 96.
[29] Ibid.
[30] Ibid., p. 94.
[31] Ibid.
[32] Ibid., pp. 104–106.
[33] Katz, Jonathan. *Digital Signatures*. New York City: Springer, 2010, pp. 4–6.
[34] Appelbaum, Deniz. Stein Smith, Sean. Blockchain Basics and Hands-on-Guidance (pp. 28–37). *The CPA Journal*. June 2018, p. 30.

The verification process can be customized depending on variables such as intended use and the number of nodes. One of the more common approaches is the proof-of-work verification.[35] It involves the nodes competing to be the first to solve a mathematical problem. The solution of the mathematical problem is deliberately time-consuming and require significant computational resources. The reason behind is to make it unattractive to tamper with the information in the blockchain, since rewriting of a block would entail rewriting and solving the mathematical problem of that block and every sequent block until the most recent one. In the Bitcoin blockchain, for example, the mathematical problem is set to a difficulty so that it takes around 10 minutes for a computer to solve it.[36]

2.2.7. The Big Picture: Building the Blockchain

Preserving integrity in an open system is a technical challenge. The method used in blockchain is to secure immutability, thus creating an add-only database. The blockchain technology contains three major elements that provide the immutability.

Firstly, the peer-to-peer approach provides a way to store information without the need for a central authority. Having the blockchain stored in several places at the same time entails significant difficulties to tamper with it, in a way that will be accepted by all the nodes.

Secondly, the use of hashes and hash references makes even the smallest manipulation stand out. The result is that one cannot manipulate or delete information, without causing invalidity of every block until the most recent one. Hence, the blockchain utilizes an all-or-nothing approach, where one either manipulate every block until the most recent one or leaves everything unchanged.

Thirdly, if someone is not afraid to manipulate the whole blockchain, this process is significantly time-consuming and resource demanding, thanks to the different verification protocols. Even if someone succeeds, they still need the manipulated version of the blockchain to be approved by all the nodes. In effect, this means hacking a majority of the nodes. With sufficient protection, that is close to impossible.

The information stored in a blockchain can represent many different things. From a legal perspective, the most relevant ones are things like transitions of ownership, transaction history, verification of documents, supply chain management, and derivatives market trades. There are

[35] Ibid.
[36] Ibid.

even examples of blockchains being used to document war crimes.[37] To better understand how blockchain functions, a visualized blockchain will be built, see figure 4 below. The visualized blockchain contains recordings of business events between Alice and Beatrice.

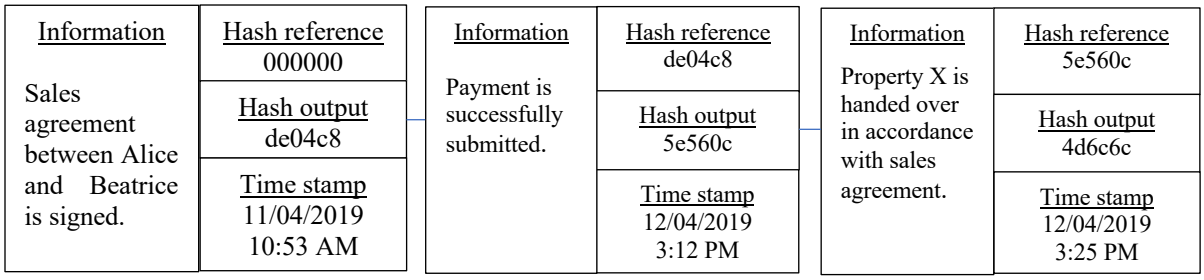

Figure 4. Simplified blockchain, recording events regarding the sales of property between Alice and Beatrice. The hash reference on the first block contains only zeros, this block is referred to as the genesis block. In reality, several events may be inserted in the same block.

If Alice wants to add a new block, she sends a request to the network through the platform that has her private key. The network knows that the request is made from Alice's account since they can decrypt it only with her corresponding public key. The request is digitally signed, meaning that the information is verified by Alice. When the nodes receive the request, the block is constructed and verified in accordance with the verification protocol. The same process is done to every following block. Both Alice and Beatrice can make requests and verify the information in the blockchain. Due to the immutability of the blockchain, paper originals should not be necessary. In other words, something like a digital original that can be trusted has been created. Furthermore, we can trace ownership securely, and thus prevent double selling of the same property.

[37] Beer, Nathan. Holding War Criminals Accountable with the Ethereum Blockchain. *Consensys*. 2018-09-18. https://media.consensys.net/holding-war-criminals-accountable-with-the-ethereum-blockchain-6b12471a7cdd (Accessed 2019-05-13).

3. The General Data Protection Regulation (GDPR)

3.1. Introduction to the GDPR

The protection of personal data is considered to be a fundamental right, laid down in Article 7 and 8 of the Charter of Fundamental Rights of the European Union (The EU-Charter) and Article 16 of the Treaty on the Functioning of the European Union (TFEU). Through the adoption of the GDPR, the EU wanted to raise the data protection level and further harmonize the data protection within the EU.[38]

The GDPR can be seen as a clarification of the right to data protection in Article 8 of the EU-Charter. Art. 8 provides everyone within the scope the right to data protection concerning him or her. It further provides that such data must be processed fairly for specified purposes and on the basis of consent of the person concerned or some other legitimate basis laid down by law. In preamble (4) of the GDPR, it is pointed out that the processing of personal data should be designed to serve mankind, and that the right of protection of personal data is not an absolute right. Instead, it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality.

The GDPR entered into force the 25th of May 2018. Unlike the former Data Protection Directive 95/46/EG, the GDPR is a regulation and thus immediately applicable and legally binding in all Member States, without the need for national legislation.[39] Even though the protection of personal data is now protected in a regulation, significant parts of the protection is unchanged compared to the former Data Protection Directive. Among other things, the definition of personal data is the same.[40]

3.2. Material Scope of the GDPR

3.2.1. General

The GDPR applies to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which forms part of a filing system or are intended to form part of a filing system.[41] The material scope of the GDPR

[38] GDPR, preamble (6).
[39] TFEU, Art. 288.
[40] Compare Directive 95/46/EG, Art. 2(a) to Article 4(1) of the GDPR.
[41] GDPR, Art. 2.1.

is thus very wide, it applies for almost every kind of treatment of personal data. It is more difficult to apply the GDPR to strictly manual processing, why the exception of personal data processed other than by automated means exists.[42]

3.2.2. Processing

Processing is defined as "any operation or set of operations which is performed on personal data, whether or not by automated means".[43] Practically every treatment of personal data will be considered processing.[44] This includes collection, structuring, storage, use, erasure etc.[45]

3.2.3. Personal data

The definition of personal data is the core of the applicability of the GDPR. Data equals to stored information, indication or signs.[46] However, data needs to be personal to fall within the scope of the GDPR. Data is considered personal if it *relates* to an *identified* or *identifiable* natural person.[47]

Often it is not a problem to determine if information *relates* to a natural person. For example, the information in a medical record will always relate to a natural person.[48] It can, of course, be less clear. The Article 29 Working Party lists a couple of situations where information may be considered as related, but it depends on the circumstances of the particular case. These are the value of a house, a car service record, a call log for a telephone, information regarding a meeting etc.[49]

A natural person is *identified* when he or she is distinguished from other natural persons. That is normally information with a particularly close relationship to the person, such as the name or detailed signs of appearance.[50]

---

[42] Frydlinger, David et al. *GDPR: Juridik, organisation och säkerhet enligt dataskyddsförordningen*. Stockholm: Norstedts Juridik, 2018, p. 64.
[43] GDPR, Art. 4(2).
[44] Voigt, Paul. Von dem Bussche, Axel. *The EU General Data Protection Regulation (GDPR)*. Cham: Springer Nature, 2017, p. 9.
[45] GDPR, Art. 4(2).
[46] Voigt, Paul. Von dem Bussche, Axel. *The EU General Data Protection Regulation (GDPR)*. Cham: Springer Nature, 2017, p. 11.
[47] GDPR, Art. 4(1).
[48] Article 29 Working Party, Opinion 4/2007 on the concept of personal data, p. 9.
[49] Ibid., pp. 9–12.
[50] Ibid., p. 12.

Accordingly, a natural person is i*dentifiable* when he or she is not yet identified, but it is possible to do it. For example, an encrypted name may not identify the person directly, but the natural person will be considered identifiable if a decryption key exists.[51] A person will also be considered identifiable if there is no decryption key, but the encryption is weak.[52]

It is an ongoing discussion in the community whether a theoretical identifiability is sufficient for the person to be considered identifiable. In 2016, the ECJ ruled that the reasonably likeliness of identifiability should be considered, taking into account the efforts needed in terms of manpower, time, cost and technological developments.[53] If the risk of identification appears insignificant based on the efforts needed, then the person is not considered identifiable.[54] The case regarded the Data Protection Directive, but there are strong indications that the criteria will be used in the GDPR as well.[55]

3.2.4. Anonymization and Pseudonymization

Anonymization is a way of removing the connection between data and the natural person. Preamble (26) of the GDPR provides that the principles of data protection should not apply to anonymous information, since it does not relate to an identified or identifiable natural person. However, the bar for data to be successfully anonymized is set high. It needs to be an irreversible de-identification. All means likely to be used to identify the person should be taken into account, considering the available technology and technological developments.[56] The Article 29 Working Party identifies risks that may allow identifiability in every commonly used anonymization technique.[57]

Pseudonymization, on the other hand, consists of processing personal data in a way that the data can no longer be attributed to a specific natural person without the use of additional information.[58] This is provided that the additional information is kept separately and is subject

[51] Ibid., p. 13.
[52] Ibid.
[53] Case C-582/14 (Breyer), ECLI:EU:C:2016:779, point 46; Opinion of the Advocate General, Case C-582/14 (Breyer), ECLI:EU:C:2016:339, points 42–49.
[54] Ibid.
[55] See GDPR, preamble (26); Schreiber, Lutz in: Plath, Kai Uwe (ed.). *BDSG/DSGVO*. 2nd edn. Cologne: Verlag Otto Schmidt, 2018; Voigt, Paul. Von dem Bussche, Axel. *The EU General Data Protection Regulation (GDPR)*. Cham: Springer Nature, 2017, pp. 12–14; For arguments against, see Buchner, Benedict. Grundsatze und Rechtmäßigkeit der Datenverarbeitung unter der DS-GVO (pp. 155–161). *Datenschutz und Datensicherheit – DuD.* Vol 40 issue 3, 2016.
[56] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques, p. 23; GDPR, preamble (26).
[57] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques, p. 23.
[58] GDPR, Art. 4(4).

to technical and organizational measures to ensure that the data is not attributed to an identified or identifiable person. The GDPR applies to pseudonymous data.

3.3. Duties and responsibilities evoked by the applicability

3.3.1. General

Applicability of the GDPR entails several duties and responsibilities. Special considerations for the use of a blockchain have to be done. The starting point is the core principles in Article 5. They specify the balance between natural persons and data controllers and should be understood as suitable compromises.[59] The principles are lawfulness; fairness and transparency; purpose limitation; data minimization; accuracy; storage limitation; integrity; and confidentiality.

The storage limitation principle is important for the examination in section 4. It requires personal data to be kept in a form which permits identification of natural persons for no longer than is necessary for the purposes for which they were collected. Storage for longer periods is allowed if the data will be processed solely for archiving purposes in the public interest, scientific or historical research or statistical purposes.[60]

3.3.2. The Lawfulness of processing

The next step is to ensure that a legal ground for processing applies. The main rule is that processing of personal data is forbidden unless justified by a legal basis.[61] The legal bases are

- consent for the specific purpose;
- contractual necessity;
- legal obligation necessity;
- vital interest of the data subject or another natural person;
- public interest or official authority necessity;
- the legitimate interest pursued by the controller or by a third party unless the interest is overridden by fundamental rights or freedoms of the data subject.

Several bases for processing can be considered for a blockchain solution. Consent will, however, always be a less appropriate ground. This is due to the fact that consents always can be withdrawn, but the information on the blockchain cannot be deleted accordingly.

---

[59] Frydlinger, David et al. *GDPR: Juridik, organisation och säkerhet enligt dataskydds-förordningen*. Stockholm: Norstedts Juridik, 2018, p. 35.
[60] GDPR, Article 5.1 (e).
[61] GDPR, Article 6.1.

3.3.3. The Obligations to Erase

The different obligations to erase personal data is the core of the tension between blockchain and the GDPR. There are a number of situations where personal data should be erased in the GDPR. Obligations to erase can be found mainly in the storage limitation principle and the right to erasure. These will be referred to as *the obligations to erase*. Moreover, the privacy by design and default contain requirements that are closely linked to the obligations to erase.

The storage limitation principle means that personal data should only be retained for as long as it is needed for the purpose of the processing.[62] Accordingly, when no legal basis that allows further processing exists, the data should be erased.

The right to erasure, often called *the right to be forgotten*, contains a right for the data subject to have personal data concerning him or her erased. It was brought to attention by the ECJ in *Google Spain & Google* and subsequently strengthened with the GDPR.[63] Article 17 of the GDPR is a codification and to some extent an extension of the right.[64] The obligation to erase applies on the following grounds:

- Lack of necessity in relation to the purpose for which the data were collected;
- Withdrawal of consent;
- Objection to the processing in accordance with Article 21;
- The data has been unlawfully processed;
- Legal obligation to erase;
- The data have been collected based on a child's consent in relation to the offer of information society services referred to in Article 8(1).

There are exceptions to the right to be forgotten. These are:

- Exercise of the freedom of expression and information;
- For compliance with a legal obligation;
- For public interest regarding public health;
- For archiving purposes in the public interest, scientific or historical purposes or statistical purposes;
- For the establishment, exercise or defense of legal claims. [65]

[62] GDPR, Article 5.1(e).
[63] Case C-131/12 (Google Spain & Google), ECLI:EU:C:2014:317.
[64] Frydlinger, David et al. *GDPR: Juridik, organisation och säkerhet enligt dataskydds-förordningen*. Stockholm: Norstedts Juridik, 2018, p. 284.
[65] GDPR, Article 17(3).

The concept of privacy by design is based on the idea that the conditions for data processing are being set by the hard- and software used for the task. Thus, when creating new technologies, developers are obliged to do that with data protection in mind. For example, the new technologies should have instruments for data minimization, pseudonymization and time limits for storage of personal data.[66] The concept of privacy by default, on the other hand, is based on the idea that only necessary personal data to the specific purpose should be processed. That applies to the amount of collected personal data, the extent of the processing, the period of storage, as well as their accessibility.[67]

[66] GDPR, Article 25.1; Voigt, Paul. Von dem Bussche, Axel. *The EU General Data Protection Regulation (GDPR)*. Cham: Springer Nature, 2017, p. 62.
[67] GDPR, Article 25.2.

4. Applying the GDPR in a Blockchain Context

4.1. Identifying Processing of Personal Data in a Blockchain

For the GDPR to be applicable, the information in a blockchain needs to be deemed processed and personal. That data is processed in a blockchain should not cause much debate, the examples in Article 4(2) of the GDPR covers practically every use of personal data in a blockchain.[68]

It is more difficult to say when data in a blockchain will be considered personal. The information in the blocks varies depending on the blockchain configuration. All data that can be used to, directly or indirectly, identify a natural person will entail applicability of the GDPR. For example, the information in the visualized blockchain in section 2.2.6 will be considered personal, since it contains Alice's and Beatrice's names. Even if the names were removed, there may be other identifiers. If the users are natural persons, one such identifier can be the public key of the asymmetric cryptography. This is due to two reasons.

Firstly, most blockchains are accessed through platforms. Many of those blockchain platforms will be subject to Anti Money Laundering laws that require identification of the users. Even without such requirements, they may hold information that allows identification of the users. As the ECJ held in C-582/14 *Breyer*, it is sufficient for identifiability that a third party has access to additional information that makes the person identifiable.[69] In my view, this applies to the public keys of the blockchain as well if the platforms have additional data that allows identification of a natural person. Secondly, a pattern may emerge if the same key is used by the same natural person in several blocks. That pattern may be used to identify a natural person and thus entail applicability of the GDPR.[70]

Identifiability may also emerge concerning other information in the block, such as addresses, phone numbers, license plate information or other similar information.

It should also be discussed whether personal data on a blockchain can be successfully anonymized, with the consequence that the GDPR does not apply. This question has not been settled by law or by clarifications from the ECJ. As previously said, the bar for anonymization

[68] See section 3.3.2.

[69] Case C-582/14 (Breyer), ECLI:EU:C:2016:779, point 49.

[70] This view is held by Schmelz, Dominik; Fischer, Gerald; Niemeier, Phillip; Zhu, Lei; and Grechenig, Thomas in: Shmelz et al. Towards Using Public Blockchain In Information-Centric Networks: Challenges Imposed by the European Union's General Data Protection Regulation (pp. 223–228). *Proceedings of 2018 1st IEEE Conference on Hot Information-Centric Networking*, 2018, p. 20.

is set high. The Article 29 Working Party acknowledge the risk that it will be difficult to fully anonymize whilst retaining as much of the underlying data as required for the task.[71] This is, as we will see, troubling for the possibilities to tweak the blockchain in order to increase compliance.

In summary, the material scope of the GDPR is very wide. The bar for what constitutes personal data is set low and the bar for successfully anonymized data is set high. Furthermore, other identifiers may emerge and make the GDPR applicable. Most blockchains will consequently be subject to the GDPR.

4.2. Possibilities to Comply with the Obligations to Erase

4.2.1. General

The immutability of the blockchain is the key property of the technology and necessary to secure decentralized trust. As we have seen, it is practically impossible to delete or change the information in a block without destroying the whole chain. The consequence is that personal data on the blockchain will be stored for an undefined amount of time, without any possibility to erase it. Storage of personal data directly on the blockchain will thus in many cases directly violate the different obligations to erase personal data, since erasing it is technically impossible. There are, however, methods to increase compliance with the GDPR. These will be examined below.

4.2.2. Making Use of the Exceptions

One method, used in a project by Lantmäteriet and Kairos Future, is to only store data that will fall under the exceptions to the obligations to erase.[72] The GDPR contains a number of such exceptions.

Consider, for example, if a group of Swedish Universities wants to create a safe way to store proofs of graduation for their former students. They want to allow simple verification of the proofs and it is vital that they are not tampered with or destroyed. Any attempts to manipulate the proofs must be noticed. A blockchain solution would allow all of this. The problem is that

[71] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques, p 10–25
[72] See: Kairos Future. *Fastighetsköp och lagfart genom en blockkedja – governance och juridik*. The land registry in the blockchain – implementation test, 2018. pp. 20–22.

the proofs of graduation will be considered personal data, and thus make the GDPR applicable. The processing may nevertheless be lawful due to the exceptions to the obligations to erase.

To begin with, the storage limitation principle requires data to be kept in a form which permits the identification of data subjects for no longer than is necessary for the purposes for which they are processed.[73] However, data may be stored for longer insofar as the personal data will be processed for archiving purposes in the public interest. Under Swedish law, proofs of graduation from Swedish universities are considered to be official documents.[74] Such documents are by law public and must be archived.[75] Regarding the legal bases for processing, there are several appropriate. Legal obligation necessity is probably the most suitable one since the proofs have to be archived under Swedish law.[76] If a natural person subsequently requests erasure, the Swedish Universities can deny the request with reference to that the processing is necessary for compliance with Swedish law.[77] The request can also be denied with reference to that the processing is necessary for archiving purposes in the public interest.[78]

One should keep in mind that there are many uses of a blockchain which will be less likely to fall under the exceptions. Imagine, for example, a sales contract between 'X' and 'Y' regarding a horse. The sales contract contains information that will make X and Y identifiable. It will be more difficult to justify storage for an undefined amount of time in this case. The horse may have been sold again, all potential claims may be time-barred, or the horse may have died. The storage limitation principle permits storage for longer periods only for certain purposes, that many blockchain will not fall under. Even if there are appropriate grounds to process the personal data for a limited time, it will be difficult to justify storage for an undefined amount of time. If 'X' or 'Y' subsequently request the erasure of the contract from the blockchain, there are no risk-free exceptions to rely on.[79] It will consequently pose a significant risk to upload the personal data to the blockchain, since storage of personal data that is deemed unlawful cannot be erased.

[73] GDPR, Article 5.1(e).
[74] Tryckfrihetsförordningen [The Freedom of the Press Act] chapter 2. Art. 1, 3 and 4.
[75] Tryckfrihetsförordningen [The Freedom of the Press Act] chapter 2; Arkivlag (1990:782) [The Archives Act] Art. 1 and 3.
[76] Compare GDPR, Article 6.1 (c). Refers to a legal obligation in the EU or in Member State Law, see Art. 6.3.
[77] GDPR, Article 17.3(b).
[78] GDPR, Article 17.3(d); see also GDPR, preamble (154).
[79] See section 3.4.3.

4.2.3. Tweaking the Blockchain to Enhance Compliance

There are a number of ways to tweak a blockchain in order to enhance compliance with the GDPR. Two methods will be briefly discussed here.

A possible method is to only upload hash outputs of the information to the blockchain. The pieces of information themselves would be stored in a normal database. This way, information can be verified by hashing the information and compare the hash output to the one on the blockchain. If any change is done, the values will not match. The argument is that this will potentially make it possible to comply with the obligations to erase data. The information in the normal database can be erased as normal, the decisive question is whether the hash output on the blockchain would be considered personal data or not. There are indications that it will, especially if the particular hash function and the range of input values are known. Among other things, the Article 29 Working Party considers hash outputs to be pseudonymous, not anonymous.[80] This is due to the theoretical risk of success with a so-called *Brute-Force Attack*, where all possible input values are hashed and compared to the information on the blockchain.[81] This risk would apply even if the corresponding information is erased. Furthermore, technical developments should be taken into account when evaluating if the hashes are anonymous. The most secure hash function at present may be considered too weak in a couple of years, ex. through developments in quantum computing.

Another method is to encrypt the personal data on the chain and store the decryption keys in a normal database. To "erase" data, the decryption key is destroyed. Again, the decisive question is whether the encrypted data on the blockchain will be considered personal or not, once the decryption key is destroyed. If the data would be considered personal, for example due to technological developments that enable easy decipher of the encryption, there would be no possibility to erase it from the blockchain.

4.3. The Big Picture: A Rising Conflict Between Technology and the Law?

As we have seen, all blockchains are not *per se* incompatible with the GDPR. Official authorities with a legal obligation to archive data can, in many cases, do so with a blockchain solution.

[80] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques, p 10–25
[81] Ibid., p. 20.

Many blockchain configurations will, however, directly violate the GDPR. The wide scope of the GDPR means that it is difficult to escape the applicability of the GDPR. The definition of personal data is wide and will hit information in most blockchains. When information in a blockchain is deemed personal, the GDPR is applicable. That means that the obligations to erase are evoked. The storage limitation principle and the right to erasure will not be possible to comply with if none of the exceptions are actualized. The concepts of privacy by design and default increases the tension further. Considering the privacy by design, it will be difficult to justify the use of a technique that does not allow erasure. One could even argue that the blockchain technique, in its nature, violates the privacy by design by not allowing erasure.

As shown, there are methods to tweak the blockchain to enhance compliance. However, all of them rely on techniques that does not result in the complete erasure of the data. Traces of the data will be left in the blockchain. These traces pose a risk to be considered personal data, especially since technological developments should be taken into account. This risk cannot be removed.

For this reason, it can be concluded that the GDPR will hinder many blockchain configurations.

The use of blockchains are easy and cost-efficient ways to secure decentralized trust. It is often said that blockchains will fundamentally change the way we do business with each other. Innovation such as Ethereum Smart Contracts, Ethereum Smart Apps, decentralized finance, tokenization of ownership etc. all relies on the blockchain technology. There are thus considerable arguments for lawmakers not to counteract the development of the technology.

In the GDPR today, there are no appropriate grounds that allow further processing of data that do not fall under the exceptions. It is an open question whether there should be any such grounds, or if the tensions should be solved by computer scientists through the development of blockchains with time-limited processing.

References

EU-Law

Consolidated version of the Treaty on the Functioning of the European Union [2012].

Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection if individuals with regard to the processing of personal data and the free movement of such data.

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

Swedish Law

Tryckfrihetsförordningen [The Freedom of the Press Act].
Arkivlag (1990:782) [The Archives Act].

Cases & Advocate General Opinions

Case C-582/14 (Breyer), ECLI:EU:C:2016:779.
Case C-131/12 (Google Spain & Google), ECLI:EU:C:2014:317.
Opinion of the Advocate General, Case C-582/14 (Breyer), ECLI:EU:C:2016:339.

Article 29 Working Party Opinions

Article 29 Working Party, Opinion 4/2007 on the concept of personal data.
Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques.

Literature

Cecilia Magnusson Sjöberg (ed.). *Rättsinformatik: Juridiken i det digitala informationssamhället*. 3rd edn. Lund: Studentlitteratur, 2018.

Drescher, Daniel. *Blockchain basics: a non-technical introduction in 25 steps*. New York City: Apress LLC, 2017.

Frydlinger, David et al. *GDPR: Juridik, organisation och säkerhet enligt dataskyddsförordningen*. Stockholm: Norstedts Juridik, 2018.

Hettne, Jörgen. Otkens, Ida (ed.). *EU-rättslig metod*. 2nd edn. Stockholm: Norstedts juridik, 2011.
Katz, Jonathan. *Digital Signatures*. New York City: Springer, 2010.

Nääv, Maria. Zamboni, Mauro (eds.). *Juridisk metodlära*. 2nd ed. Lund: Studentlitteratur AB, 2018.

Plath, Kai Uwe (ed.). *BDSG/DSGVO*. 2nd edn. Cologne: Verlag Otto Schmidt.

Voigt, Paul. Von dem Bussche, Axel. *The EU General Data Protection Regulation (GDPR)*. Cham: Springer Nature, 2017,

Articles, Reports and White Papers

Appelbaum, Deniz. Stein Smith, Sean. Blockchain Basics and Hands-on-Guidance (p. 28-37). *The CPA Journal*. June 2018.

Baratta, Roberto. Complexity of EU Law in the Domestic Implementing Process (pp. 293-308). *The Theory of Practice and Legislation*, Vol 2 issue 3. 2014.

Basedow, Jurgen. Zimmermann, Reinhard (eds.). *The Max Planck Encyclopedia of European Private Law* (pp. 979-892). Oxford: Oxford University Press, 2012.

Bayle, Aurelie et al. When Blockchain Meets the Right to be Forgotten: Technology Versus Law in the Healthcare Industry. *IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, 2018.

Binh Truong, Nguyen et al. GDPR-Compliant Personal Data Management: A Blockchain-based Solution. *IEEE Transaction on Information Forensics and Security,* 2019.

Buchner, Benedict. Grundsatze und Rechtmäßigkeit der Datenverarbeitung unter der DS-GVO (pp. 155-161). *Datenschutz und Datensicherheit – DuD.* Vol 40 issue 3, 2016.

Kairos Future. Fastighetsköp och lagfart genom en blockkedja – governance och juridik. The land registry in the blockchain – implementation test, 2018.

Lyons, Tom et al. *Blockchain & the GDPR*. EU Blockchain Observatory and Forum, 2018.

Nakamoto, Satoshi. *Bitcoin: A Peer-to-Peer Electronic Cash System*, 2008.

Shmelz et al. Towards Using Public Blockchain In Information-Centric Networks: Challenges Imposed by the European Union's General Data Protection Regulation (pp. 223-228). *Proceedings of 2018 1st IEEE Conference on Hot Information-Centric Networking*, 2018.

Internet

EU Blockchain Observatory and Forum. https://www.eublockchainforum.eu/about (Accessed 2019-05-13).

Beer, Nathan. Holding War Criminals Accountable with the Ethereum Blockchain. *Consensys*. 2018-09-18. https://media.consensys.net/holding-war-criminals-accountable-with-the-ethereum-blockchain-6b12471a7cdd (Accessed 2019-05-13).