Juridiska Institutionen
Handelshögskolan vid Göteborgs Universitet
Examensarbete Juristprogrammet
HRO800 HT21, 30 hp.

GDPR, Blockchain & Personal data

The rights of the individual v. the immutability of Blockchain

David Matsson

Handledare: Kristoffer Schollin
Examinator: Jens Andreasson

GÖTEBORGS UNIVERSITET
HANDELSHÖGSKOLAN

Abstract

Blockchain was introduced by the pseudonym “Satoshi Nakamoto” in 2008 and took the world by storm. One of the main features of Blockchain technology is the immutability of its ledgers. A necessary component to be able to function in a low-trust environment as they are designed to do. The GDPR entered into force in 2018 and established further possibility of data subjects to protect their individual rights. Such rights are, by others, the Right of Rectification and the Right of Erasure. These rights demand that where personal data exists, the data must be mutable. Thus, making the base for the conflict at hand. The aim of this thesis is to investigate the compatibility between the Right of Rectification and the Right of Erasure, as granted to the individual by the GDPR, and the append-only ledgers of Blockchain.

To remedy the issue of an immutable, append-only ledgers with the rights of the individual, a risk-based approach regarding the concept anonymous data must be adopted. Thus, allowing different interpretations at different times. In the light of this, with technological solutions and a risk-based approach towards the concept of non-personal data, Blockchain can be compatible with the Right of Rectification and Right of Erasure.

Abbreviations

BTC	Bitcoin
E.g.	For example
Etc.	Et cetera
ECJ	European Court of Justice
EU	European Union
GDPR	General Data Protection Regulation
I.e.,	In other words
NFT	Non-Fungible Token
OECD	Organization for Economic Co-operation and Development
P2P	Peer-to-peer
RFID	Radio Frequency Identifier
ROE	Right of Erasure
ROR	Right of Rectification
TFEU	Treaty of the Functioning of the European Union

Table of contents

1 Introduction

The technology of Blockchain is something new and innovative. Aspects of the technology have been discussed and proposed in theoretical academical papers since the 1980s.[1] It was introduced by a group which goes under the pseudonym Satoshi Nakamoto through the publishing of the paper "*Bitcoin: A peer-to-peer electronic cash system"*. Blockchain is an innovative structure allowing transactions between two parties within a trustless environment, without the necessity of a trusted third party. This is done through one of Blockchains most fundamental features, transparent immutable transaction history. Blockchain constitutes a class of technologies and as such, allowing Blockchain to be utilized within many different markets and towards different purposes.

Alongside the technological evolution, the need for the protection of the natural person's personal data has increased. In order to meet this, the General Data Protection Regulation (GDPR) was adopted. The GDPR establishes the rights of each individual regarding their own personal data. Two of these rights are the Right of Rectification and the Right of Erasure. These two rights require the possibility of changing existing data or the deletion of said data.

As the European Commission has stated in its digital strategy that the EU shall be world leader within Blockchain technology, an interesting friction between two goals of the Union arise. Namely, the goal of protection of personal data according to the GDPR and the goal of the Commission to be world leader within Blockchain. On the one hand, Blockchains are by definition immutable, with their append-only ledgers, and on the other hand, the GDPR puts up requirements that require these same ledgers to be the opposite, namely mutable.

1.1 Aim and research questions

The broad aim of this thesis is to investigate the compatibility of GDPR and Blockchain technology and the possible future for coexistence within the EU.

The Blockchain technology is designed to achieve decentralization, resilience through replication and to be an append-only ledger. The GDPR is based in the 1995 Data Protection Directive, and it seeks to facilitate the free movement of personal data between the EU's member states as well as establishing a framework of fundamental rights protection. The

[1] A. Narayanan and J. Clark, *Bitcoin's Academic Pedigree*, (2017), 60 Communications of the ACM, p.36

framework contains such rights as the Right of Rectification and the Right of Erasure to name a few. These two examples from the GDPR highlights the conflict between the personal data regulation and the append-only ledgers of Blockchain. In order for the EU to be leaders in Blockchain technology, it needs to be examined if the GDPR hinders this goal from being fulfilled and if so, how such a situation could be rectified.

Developed from this aim, the research question of the thesis is:

- Are the append-only ledgers of Blockchain technology compatible with the Right of Rectification and the Right of Erasure of the GDPR?

Blockchains compatibility with the GDPR is dependent on the possible interpretation of the GDPR as well as different possible and already existing technological designs of Blockchain. The issue at hand is mostly focused on the general definition of personal data. Thus, one part of the Blockchain possible incompatibility with GDPR focuses on the legal definition of personal data. Another possible way for Blockchain's append-only ledgers to be considered compatible with the GDPR is different technological solutions which allow for the personal data to be stored outside of the Blockchain. Thus, the following sub-questions to the main research question emerges:

- Can an interpretation of the GDPR's definition of personal data make the append-only ledgers compatible with the GDPR?
- Can the technological design "off-chain storage" make Blockchain compatible with the GDPR?

1.2 Delimitations

The targeted audience for this thesis is lawyers with a basic understanding of software and software architecture. The area of Blockchain technology is relatively new and technologically complex in its structure. This thesis will therefore not consider technological structures of Blockchain in such a way that would require preexisting knowledge of the subject. It will provide a basic account of how the technology is structured in order to provide an understanding that is sufficient to comprehend the complexity of the relationship between the GDPR and Blockchain technology.

This thesis will furthermore focus on open permissionless Blockchains in a peer-to-peer structure, rather than private and centralized alternatives. Blockchains can be designed to do a vast number of things, this thesis will only regard Blockchains that are designed to perform transactions of different kinds. Blockchain is normally associated with cryptocurrencies. This thesis will not, however, manage cryptocurrencies unless it is imperative to highlight aspects that is of importance. This thesis will not manage the consensus protocol of Blockchain.

There are several different suggestions for technological solutions regarding the different requirements which the GDPR applies on Blockchain. This thesis will solely discuss the possibility regarding off-chain data storage and therefore not discuss alternatives such as "pruning", "chameleon hashes" etc.

Furthermore, a basic introduction to the GDPR will be made. This thesis will focus on the rights of the data subject, but only in regards of the Right of Rectification and the Right of Erasure. A further discussion regarding the concept of personal data will be made and thus a more in-depth introduction regarding this will follow. Many aspects of the GDPR provide interesting question and problems within the context of Blockchain technology and distributed ledgers, e.g., the legality of processing data indefinitely. This thesis will however not discuss any other aspects than the Right of Rectification, Right of Erasure and personal data to a certain extent. However, a quick introduction to the different areas of the GDPR that is necessary to give the reader a sound understanding of its structure.

1.3 Method

The overall theme and the research question are concerning a relatively new phenomenon, both regarding the legislation as well as the technology at hand. The research questions concern compatibility and possible remedies for achieving compatibility. To ascertain a stable ground for the analysis to be conducted, the method of legal dogmatics will be used. A clear description of legal dogmatics is not an easy task since the definitions of it is a much-discussed topic between scholars.[2] It is however based on utilizing traditional accepted sources of law such as in the legislation and case law to solve the question at hand. [3]

[2] C, Sandgren, *Är rättsdogmatiken dogmatisk?* Tidsskrift for rettsvitenskap, (2005) Tfr, 118(4-5), p. 648
[3] J, Kleineman. "Rättsdogmatisk metod", In M, Nääv and M, Zamboni (eds.), *Juridisk Metodlära*, 2nd edn., Lund, Studentlitteratur AB, (2018), p. 21

The wording dogmatic inclines a lack of flexibility and might be deemed to be unsuited for an academic discipline. Kleinman argues that the wording in legal dogmatics does not have to be correlated to the definition of dogmatic. That legal dogmatics does not have to be dogmatics in its core, instead it is about analyzing the different elements in jurisprudence to achieve a finished conclusion that will be assumed to reflect how the rule of law should be perceived in a concrete context.[4]

Olsen also argues that any legal dogmatic dictum should be considered to be normative. Thus, the use of sources must be derived from authoritative sources.[5] According to Olsen, the range of sources should be very narrow, and any use of sources outside from what is considered to be authoritative will have a sociological incuse on the discussion. Due to the source's axiomatic significance for legal dogmatics, the choices of sources should not be specifically discussed.[6]

This view on the sources proposed by Olsen is not something that has been unchallenged. Jareborg has stated that even though that legal dogmatics is closely associated with the reconstruction of applicable law, it should not hinder the possibility to go outside of applicable law when arguing within legal dogmatics. That it is legitimate for legal dogmatics to search for ideal solutions.[7] In the light of this, Kleineman also highlights the possibility of conducting critical legal dogmatic research.[8] That while conducting research through the legal dogmatic method, go a step further and argue through the use of purpose arguments, that the legal situation is not satisfying and that in one way or another it should change.[9] Kleineman has clarified that legal dogmatics is a tied argumentation[10] and thus you have to base any purpose arguments within applicable law, otherwise it will not have independent significance.[11] Kleineman has continued to state legal dogmatics constitute an important link between interpretations from authoritative sources and independent arguments derived from justice policy.[12]

As previous stated, this thesis aims to analyze technology and its compliance with data protection regulation. The applicable law in this case is the GDPR which entered into force

[4] J. Kleineman, "Rättsdogmatisk metod", In M, Nääv and M, Zamboni (eds.), *Juridisk Metodlära*, p. 26
[5] L. Olsen, *Rättsvetenskapliga perspektiv*, SvJT, (2004) p. 119
[6] Ibid.
[7] N. Jareborg, *Rättsdogmatik som vetenskap*, SvJT, (2004) p. 4
[8] J. Kleineman, "Rättsdogmatisk metod", In M, Nääv and M, Zamboni (eds.), *Juridisk Metodlära*, p. 44
[9] Ibid.
[10] Here the author means that the arguments in legal dogmatics must be grounded in the traditional law sources.
[11] J. Kleineman, "Rättsdogmatisk metod", In M, Nääv and M, Zamboni (eds.), *Juridisk Metodlära*, p. 44
[12] Ibid., p. 45

during 2018.[13] The technology at hand is also a novelty which popularity and spread has increased since its introduction in 2008. The novelty of both applicable law and the technology entails a scarcity regarding sources of law which is an issue to conclude a legal dogmatics analysis. If the analysis were to be conducted according to the Olsen's view on the method legal dogmatics there would be an issue of sources and the legitimacy of any conclusion would be highly questionable. Therefore, the applied method will be conducted through the view of legal dogmatics which has been argued by both Jareborg and Kleineman, allowing for a wide span of sources to be able to conduct a legitimate analysis.

Since the legal dogmatics method is, as aforementioned, in its very basic nature a method of reconstructing the applicable law in the relevant context, it is deemed to be the best suited method for this thesis. The research questions are aimed towards the compatibility between the GDPR and the structure of Blockchain. Thus, the chosen method will allow for the reconstruction of applicable law in this context. Furthermore, the use of Kleineman's critical legal dogmatics will be used in order to be able to answer the research question regarding a new possible interpretation of the GDPR in order to remedy to compatibility issue.

1.4 Material

As aforementioned, the field is novel in both the legislation and the subject. Thus, the supply of material is scarce. The material used in this thesis consists of material relevant for answering the research questions. The questions are not posed on a national level but instead on an international EU-level. The material will therefore, in short terms, include EU-legislation, doctrine and articles concerning given EU-legislation, as well as statements and strategies delivered from the European Commission.

The GDPR established the European Data Protection Board, which replaces the Article 29 Working party. As of now, European Data Protection Board guidance is still rather scarce regarding many questions of implementation. There is a consensus that the opinions of Article 29 Working Party, where the opinions have corresponding wording in the GDPR, are still

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

relevant regarding the implementation of the GDPR.[14] Regarding Article 29 Working Party, it is important to note that these are not legally binding documents but instead are used as a form of guidance regarding its implementation. These opinions will be used in this thesis as a way to show how the different elements of the GDPR should be interpreted.

The design and intricate parts of the Blockchain are very technical and complex by nature. To be able to provide a correct but simplified understanding of this, there is an issue of sources. This thesis will in this regard mostly be based on one book[15] which is designed to provide a basic, non-technical introduction and understanding of Blockchain. The author of the book is Daniel Drescher who holds a doctorate in econometrics and a Master of Science in software engineering. To ensure the correctness of the technical aspects of the book, a technical review of the content has been concluded by an expert on the matter. With that said, the book is not written for lawyers per se, but it is written in a manner in which no preexistent knowledge of the matter is necessary.

The use of national data protection authority's implementation of the GDPR and their recommendations will occur in the thesis. E.g., the French data protection authority will be used as an example of recommendation regarding compliance with the Right of Erasure established by the GDPR.

The basis for the analysis will mostly consist of articles and doctrine concerning the application of the GDPR to Blockchain. These sources are not in abundance and therefore the number of sources referenced in this thesis will be affected. Regarding these it is worth noting that in 2019, a report regarding the GDPR and Blockchain was conducted by the Panel for the Future of Science and Technology.[16] This report will be used throughout the thesis. However, the parts which will be discussed will not solely be based on the conclusion of Finck in the report as well as most parts are, in the report, left as open questions and request for further regulatory guidance.

[14] See M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* Panel for the Future of Science and Technology (STOA), European Parliament Research Service, 2019
[15] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, Main, APress, 2017,
[16] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?*

1.5 Disposition

This thesis will consist of four chapters.

After this introduction to the thesis, chapter two will consist of an introduction and investigation into the GDPR. This part is focused on giving the reader a basic understanding of the structure of the GDPR. This introduction will go as far as needed to give the reader a sound understanding which will aid in the comprehension of the analysis and the research questions. Following this, an introduction to the concept and technology of Blockchain will be provided. This introduction will be left at a very basic level in order to avoid unnecessarily complicating the matter, an in-depth understanding of the matter will not be necessary. With both these chapters, the goal is to give the reader a comprehensive understanding of both subjects that will allow the reader to comprehend the application of the GDPR to the Blockchain in the light of the research question that will be shown in the fourth chapter. Here the analysis of the compatibility between the Right of Rectification and erasure and the append-only ledger of the Blockchain will be carried out. This chapter will be divided, first it will show the basic compatibility between these rights and the append-only ledgers and then it will dive into specific circumstances for each right and possible remedies for incompatibility. This chapter ends with final remarks from the author.

2 General Data Protection Regulation

2.1 A Brief History of Data Protection Law

Within Europe, Sweden was the first state to establish a law concerning the processing of personal data which was founded in 1973.[17] However, Sweden was not unique in regard of data protection regulation. The Swedish establishment of these were tightly associated with the international development within the area. In 1980 the Organization for Economic Co-operation and Development (OECD) laid out a set of guidelines concerning the right of privacy and the cross-border transfer of personal data. In the OECD guidelines, the discrepancy between the national regulation concerning data protection was recognized as a hindrance for international transaction of personal data and therefore also a hindrance against growth.[18]

[17] D. Frydlinger et al., *GDPR: Juridik, organisation och säkerhet enligt dataskyddslagen,* Stockholm, Norstedts Juridik, 2018, p. 22
[18] Ibid., p. 23

Parallel with the guidelines set by the OECD. The Council of Europe's (the Council) convention for the protection of individuals about automatic processing of Personal Data CETS No. 108 was accepted. The aim was to create a greater unity within the member states, as well as to extend the protection for everybody's rights and fundamental freedom with special focus to the right of privacy.[19]

In 1993, the European Community was transformed through the Treaty of Maastricht to the European Union (EU). The European Commission (Commission) acknowledged in accordance with both the Council and OECD that the discrepancy between the data protection regulation of the member states as a threat towards the internal market.[20] In order to remedy this, the Commission adopted Directive 95/46/EG on *Protection of Individuals with regard to the Processing of Personal Data and on the Free Movement of such Data* (Data Protection Directive).[21] The Data Protection Directive was implemented as a way to remedy the inconsistency of data protection law between the member states.[22] It acknowledged the double edged sword of data protection, namely the right of privacy and the free mobility of personal data between the member states.[23] The Data Protection Directive established the independent advisory board Article 29 Working Party (the Working Party).[24] The Working Party had the role to examine any question regarding the national implementation of the Data Protection Directive. The Working Party also gave opinions regarding different codes of conduct, proposed amendments and additional measures to safeguard the protection of personal data. [25]

The EU Treaty of Lisbon resulted in the amendment of two treaties which form the constitutional basis of the EU. The *Treaty of the Functioning of the European Union* (TFEU) and the *Charter of Fundamental Rights of the European Union* (the Charter). Both constitutional treaties established different rights of the data subject.[26] The TFEU established the right of the protection of personal data that is relating to them.[27] The Charter established the

[19] Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data No. 108, Preamble
[20] D. Frydlinger et al., *GDPR: Juridik, organisation och säkerhet enligt dataskyddslagen*, p. 24
[21] 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 on the free movement of such data,
[22] P. Voigt and Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, Cham, Springer International Publishing, 2017, p. 2
[23] D. Frydlinger et al., *GDPR: Juridik, organisation och säkerhet enligt dataskyddslagen*, p. 24
[24] Art. 29 the Data Protection Directive
[25] Art. 30 the Data Protection Directive
[26] D. Frydlinger et al., *GDPR: Juridik, organisation och säkerhet enligt dataskyddslagen*, p. 26
[27] Art. 16, The treaty on the functioning of the European Union

right of privacy and life[28] as well as the protection of personal data.[29] In contrast to the TFEU, the Charter established more thorough rules regarding how the personal data have to be processed and the requirements of transparency.[30]

The Data Protection Directive was not directly applicable in the member states. Instead, it had to be implemented into the national law of each state. This led for different states to have different interpretation and implementation of the directive, which resulted in fragmentation across the EU regarding the implementation for the Data Protection Directive. This resulted in that data processing activities could be lawful in one member state and unlawful in another. Due to this, the Data Protection Directive did not live up to one of its fundamental goals namely harmonizing the level of data protection and the different regulations within the Union.[31]

2.2 General Data Protection Regulation

Regulation (EU) 2016/679 on *the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of such Data, General Data Protection Regulation* was entered into force on May 25, 2018, to replace the Data Protection Directive.

The GDPR is, contrary to the Data Protection Directive, directly applicable in each member state without further implementation.[32] Since the Data Protection Directive, technological advancements have changed both the economic and the social life. Together with the rapid technological developments and globalization, it establishes new challenges on data protection.[33] While the principles of the Data Protection Directive remain sound, it did not prevent the fragmentation in the implementation where harmonization constitutes a necessity to provide protection of personal data and for the further economic development of the Union.[34]

The GDPR implemented the European Data Protection Board (the Board).[35] The board was constructed to promote a consistent application of the GDPR, thus replacing the Working

[28] Art. 7, The Charter of Fundamental Rights of the European Union
[29] Art. 8, The Charter
[30] Ibid.
[31] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 2
[32] Ibid.
[33] Rec. 6 and 7 GDPR
[34] Rec. 9 GDPR
[35] Art. 68 GDPR

Party.[36] The Board provides guidelines, recommendations and best practices for the interpretation and application of the GDPR.[37]

The fundamental purpose of the GDPR is to contribute to an area of freedom, security, and justice as well as the convergence of the economies within the internal market.[38] The recitals of the GDPR establish that the processing of personal data should be designed to serve mankind and that data protection does not constitute an absolute right.[39] The GDPR have to be balanced against other factors, such as its functioning in society as well as other fundamental rights.[40] This allow the rule of proportionality to be applied to the GDPR.[41] An example of the kind of consideration between fundamental rights is found in the case *Google Spain, Google Inc. V. Agencia Española de Protección de Datos (AEPD), Mario Costeja González* from the European Court of Justice (ECJ), a case which will be examined further below. The rule of proportionality is established to make sure that the data protection regulation serves mankind and that the implementation of said regulation is not unproportionate in reference to the negative result on the functioning of society as well as the economic growth of the internal market.[42]

2.2.1 Scope of application

The most vital part of the GDPR is its scope of application. This is structured into a *Material scope* and a *Territorial scope*. These two criteria of the GDPR's applicability will be introduced in the following chapters.

2.2.1.1 Material scope

The GDPR applies to any form of processing of personal data by automated or partly automated means, as well as processing by means other than by automated if it forms or is intended to form a filing system.[43] The material scope is by design very wide. When constructing the GDPR, a technological neutral approach was sought.[44] The goal was to make it impossible or extremely difficult to circumvent the regulation through technological measures.[45]

[36] Rec. 139 GDPR
[37] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 197
[38] Rec. 2 GDPR
[39] Rec. 4 GDPR
[40] Ibid.
[41] D. Frydlinger et al., *GDPR: Juridik, organisation och säkerhet enligt dataskyddslagen*, p. 30
[42] Ibid.
[43] Art. 2.1 GDPR
[44] D. Frydlinger et al., *GDPR: Juridik, organisation och säkerhet enligt dataskyddslagen*, p. 63
[45] Rec. 15 GDPR

There are some exceptions as to the material scope of application which constitutes the processing activity done by competent authority, for example in public security.[46] Another important exception from the material scope is the exception for natural persons in the course of a purely personal or household activity.[47] The household exception should however be interpreted narrowly as the inclusion of business together with private information in the processing will make the exception inapplicable.[48] It is furthermore important to note that the GDPR is not applicable for a deceased person.[49]

2.2.1.2 Territorial scope

The territorial scope of the GDPR is divided into two parts. *The first part* addresses when the processing of personal data occurs in the context of the activities of an establishment of a controller or processor within the Union.[50] The definition of establishment implies the effective and real exercise of activity through stable arrangement.[51] Observe that the processing itself does not have to take place within the Union.

The second part addresses the processing of personal data of data subjects that are within the Union by a processor or controller that is not established within the Union.[52] Through the GDPR, the territorial scope was extended by introducing the principle of *lex loci solutionis*.[53] The processing made by a processor or controller that is not established within the union is still included in the territorial scope as long as processing activities are relating to the offering of goods or services to data subjects, the monitoring of the data subjects behavior as long as their behavior takes place within EU.[54] It is important to note that in the context of territorial scope, the GDPR does not differentiate between processing executed by the controllers or the processors, but instead sets up the same territorial scope for both.[55]

The transnational application of the GDPR was a way to guarantee comprehensive privacy of individuals and supply fair competitive conditions on the EU internal market. The GDPR is

[46] Art. 2.2(d) GDPR
[47] Art. 2.2(c) GDPR
[48] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 16
[49] Rec. 27 GDPR
[50] Art. 3.1 GDPR
[51] Rec. 22 GDPR
[52] Art. 3.2 GDPR
[53] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 26
[54] Art. 3.2 GDPR
[55] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 22

also intended to remedy the phenomena where companies choose their place of business according to the lowest national level of data protection standard, so-called *forum shopping*.[56]

2.2.2 Personal data

The definition of personal data is naturally the most central in the GDPR.[57] The material scope of the GDPR clearly states that it must concern the processing of *personal data* for it to fall within the scope.[58] The GDPR defines personal data as:

"… any information relating to an identified or identifiable natural person (data subject)…"[59]

This definition of personal data could be considered to be structured in four building blocks.[60] These are "any information", "related to", "an identified or identifiable" and "natural person". The intention of the definition of personal data is supposed to have a wide interpretation.[61]

2.2.2.1 "Any information"

The first element is as aforementioned "any information". This concludes that from the perspective of the nature of the information, it covers anything from objective to subjective, true, or false data.[62] Personal data includes information *stricto sensu* that is relating to the individual private, and family life, but it also includes information about whatever types of activity that is undertaken by said individual.[63] The format or medium of the personal data which the information contains bears no change in the consideration if it is to be seen as "any information" or not.[64]

2.2.2.2 "Relating to"

The second element of the definition is "relating to". To put it in general terms, the information in question is considered to be "relating to" an individual if it concerns said individual.[65] In the normal case this is quite straight forward. However, not always, e.g., when the information is

[56] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 22.
[57] D. Frydlinger et al., *GDPR: Juridik, organisation och säkerhet enligt dataskyddslagen*, p. 41
[58] Art. 2.1 GDPR
[59] Art. 4.1 GDPR
[60] Art. 29 Data Protection Working Party. (2007) Opinion 4/2007, on the concept of personal data, WP 136, p. 6
[61] Ibid., p. 4
[62] Ibid., p. 6
[63] Ibid.
[64] Ibid., p. 7
[65] Ibid., p. 9

relating to an object and not an individual. This is referred to as information being indirectly relating to an individual.[66] The Working Party recommends that in these cases a content, purpose or result element should be considered to determine if the data is relating to an individual or not.[67] The *content element* states that information is "related to", when it is about that person. The *purpose element* states that when data is used or likely to be used with the purpose of evaluate, treat in a certain way or influence the status or behavior of an individual, it is considered to be "related to". While evaluating this, all circumstances surrounding the case should be considered.[68] In the case where both the content and the purpose element is absent, the *result element* can be applied. Such is the case when the result is likely to have an impact on the certain rights and interests of the individual. According to the Working Party, this does not have to be a major impact. If the result is that the individual is treated differently due to the data, it is considered to be "related to".[69] The Working Party has further given out a working paper on the issues regarding Radio Frequency Identifier (RFID) Technology.[70] The Working Party concluded that in the case of RFID[71] tags, if the data which the tag refer to constitute personal data, so will the tag.[72]

2.2.2.2.1 Identifiability

The natural person does not have to be identified already, the mere possibility of identification will render the data personal.[73] In general terms, a person is considered to be identified if the individual could be distinguished from a group of persons.[74] This can be made through the combination of different pieces of information to make a possible identification.[75] These pieces of information are called "identifiers" which are different characteristics which hold a particular privileged and close relationship with the particular individual.[76] These characteristics are the expression of a physical, physiological, psychological, genetic, economic, cultural, or social identity.[77] These different pieces of information does therefore not

[66] The Working Party, Opinion 4/2007, p. 9
[67] Ibid., p. 10
[68] Ibid.
[69] Ibid., p. 11
[70] Art. 29 Data Protection Working Party. (2005) Working document on data protection issues related to RFID technology, WP 105.
[71] RFID is giving an object a tag with an identity which could be communicated to a reader over radio frequency.
[72] The Working Party, *RFID technology*, p.8
[73] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 12
[74] The Working Party, Opinion 4/2007, p. 12
[75] Ibid.
[76] The Working Party, Opinion 4/2007, p. 12
[77] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 11

have to identify the natural person on its own, but in combination with other identifiers allow identification.[78]

The definition of personal data in the GDPR states that a person can be identified "directly" or "indirectly".[79] When a "directly" identified or identifiable person is discussed the said individual's name is the most common example and what the notion of an identified person implies.[80] However the name of a person is sometimes not enough. Additional information is often necessary such as date of birth, a picture or the name of the individual parents.[81]

"Indirectly" is instead a category of cases where the identifiers *prima facie* does not allow anyone to be identified but that information together with other might allow the individual to be distinguished.[82] This kind of identification may occur when the use of identifiers allows the individual to be singled out.[83] This has become an increasing problem due to the use of internet and our presence on it. The use of web traffic surveillance and similar tools have made it increasingly easier to learn and map the behavior of a user online.[84] The hypothetical possibility to single an individual out is however not enough for it to automatically being considered identifiable per the GDPR. When determining if the data subject is identifiable, one must consider all the means reasonably likely to be used by the controller or another person to acquire additional information from whatever source.[85] Here the GDPR allows additional factors to be considered. Such factors are e.g., the cost of conducting the identification, the interests at stake for the individual and the risk of organizational disfunctions etc.[86] Additional focus must be added regarding available technology and the possible technological developments.[87] The Working Party has concluded that this assessment should be based on the possibilities for technological advancements during the period for which the data is processed.[88]

Another important factor to determine "all the means likely to be used" is the purpose of the processing.[89] Objections from controllers are normally that only scattered pieces of information

[78] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 11
[79] Art. 4.1 GDPR
[80] The Working Party, Opinion 4/2007, p. 13
[81] Ibid.
[82] Ibid.
[83] Ibid.
[84] Ibid.
[85] Recital 26 GDPR
[86] The Working Party Opinion 4/2007, p. 15
[87] Rec. 26 GDPR
[88] The Working Party Opinion 4/2007, p. 15
[89] Ibid., p. 16

are being processed without reference to a name or other identifiers. The Working Party argues that in this kind of situation, the processing of this information only makes sense if it allows identification of specific individuals. It should therefore be assumed that where the purpose of processing is as such, they shall be considered to have the means "likely to be used" for identification.[90]

Under the Data Protection Directive, it had been controversially discussed whether relative or absolute criteria had to be used to establish "reasonable likeliness of identifiability".[91] A absolutist approach would result in that as soon as anyone would have the possibility to connect the processed data to an individual, the definition of personal data would have been met.[92]

Following two chapters will bring up cases from the ECJ which has affected the interpretation of identifiability.

2.2.2.2.1.1 Scarlet Extended

In the case Scarlet Extended[93], the proceeding was regarding that internet users were accessing and downloading copyright protected work from the company SABAM. The case of Scarlet Extended is one of copyright and the balancing of copyright as well as the freedom of information and therefore not interesting in the context of GDPR. However, in Scarlet Extended, the ECJ states that IP addresses constitutes personal data. They back this up with the motivation that the IP addresses[94] allows for the users to be precisely identified.[95]

2.2.2.2.1.2 Breyer V. Germany

In the case Breyer V. Germany[96], the question about identifiability was lifted. Mr. Breyer had accessed several websites which was operated by the German Federal Institutions. These websites had, with the aim of preventing attacks, chosen to store the information on all access operations in logfiles. In these logfiles, the IP address of the computer which accessed the website was sought. Mr. Breyer brought an action before the court seeking an order restraining

[90] The Working Party Opinion 4/2007, p. 16
[91] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 12
[92] Ibid.
[93] Case C-70/10, *Scarlet Extended SA V. Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)*
[94] IP addresses are series of digits assigned to networked computers to facilitate their communication over the internet. This could be a "static" or a "dynamic" IP address. The dynamic IP addresses do not enable a link to be established.
[95] Case C-70/10, *Scarlet Extended*, para. 51
[96] Case C-582/14, *Patrick Breyer v Bundesrepublik Deutschland*

the Federal Republic of Germany from storing or arranging for third parties to store, after consultation of the websites accessible to the public, the IP address of the applicant's host system except in so far as its storage is unnecessary to restore the availability of those media in the event of a fault occurring.

The ECJ first stated that the dynamic IP addresses[97] do not constitute information directly relating to an "identified natural person" since it does not in itself reveal the identity of the natural person.[98] The ECJ continued to discuss the implication of the wording "indirectly" in the Data Protection Directive in the context of that all the information enabling of the data subject must not be in the hand of one person.[99]

The fact that the data necessary to identify the user was not held by the same organization did not implicate exclusion of dynamic IP addresses being personal data.[100] The ECJ further concluded that the existence of channels which allowed the competent authority to get the information needed to make the identification[101] implicate that these dynamic IP addresses would fall within what constitutes a means likely reasonably to be used to identify the data subject.[102]

This would not be the case if such a link were unable be established, establishing such a link were prohibited by law or establishing such a link would have been a disproportionate effort in terms of time, cost and man-power. In such a case it would not have constituted means likely reasonably to be used to identify the data subject and thus, not constitute personal data.[103]

2.2.2.2.2 Natural person

The GDPR is only applicable for natural persons, that is "human beings".[104] However, the GDPR is not applicable for data that are relating to a deceased human being.[105] Even though

[97] A dynamic IP address is an IP address which changes each time, thus minimizing the risk of being linked to the user.
[98] Case C-582/14, *Breyer*, para. 38
[99] Ibid., para. 43
[100] Ibid., para. 44
[101] In this situation, the competent authority was allowed to get the information in the events of cyber-attacks with the purpose to bring criminal proceedings.
[102] Case C-582/14, *Breyer*, para. 47
[103] Ibid., para. 46
[104] Art. 4.1 GDPR
[105] Rec. 27 GDPR

the GDPR is not applicable for personal data relating to a deceased person, the data of said person can be personal data for its descendant in the case of e.g., hereditary diseases.[106]

2.2.3 Processing

According to the GDPR, processing means any operation or set of operations that are performed on personal data or on sets of personal data, whether or not by automated means.[107] I.e., processing comprehends any treatment of data.[108] However, if the processing is done manually, it falls within the material scope of the GDPR as long as it forms part of a filing system or is intended to form part of a filing system.[109] As aforementioned, the definition of processing as stated by the GDPR is by design very wide. The intention of this wording is to make it hard to circumvent and to make it independent from technological change.[110]

The processing of personal data must be lawful.[111] One of the grounds for processing to be lawful is if the consent of the data subject has been given to the processing of the data subjects personal data.[112]According to the GDPR, consent is when any freely given, specific, informed, and unambiguous indication given by the data subject signifies agreement regarding this processing of personal data.[113] As an example of what is not considered to be freely given consent is when an employer demands consent to data processing by its employee. The power imbalance between the two parties makes it impossible for the GDPR to accept the consent given by the employee to be freely given, unless very special circumstances apply.[114] The requirement for unambiguousness can be said to require it to be excluded that the data subject have not consented to the given purpose of processing. This assessment requires to factor in aspects such as how the consent was given and the distinctness in the information which the data subject been given.[115]

The data subject always has the right to withdraw its consent to the processing of its personal data. Regarding this right, the GDPR has explicitly written that it should be as easy to withdraw

[106] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 11
[107] Art. 4.2 GDPR
[108] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 9
[109] Art. 2.1 GDPR
[110] Chapter 2.2.1.1
[111] Art. 5.1 (a) GDPR
[112] Art. 6.1 (a) GDPR
[113] Art. 4.11 GDPR
[114] D. Frydlinger et al., *GDPR: Juridik, organisation och säkerhet enligt dataskyddslagen,* p. 147
[115] Ibid., p. 148

the consent as it was given. When this right is exercised, the lawfulness of the processing based on said consent prior to the withdrawal shall not be affected, but any further such processing would.[116]

2.2.4 Controller

The definition of a controller as stated by the GDPR is:

"… the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purpose and means of the processing of personal data…"[117]

This definition is built on three main components, namely "a natural or legal person, public authority, agency or other body", "That alone or jointly with others", "determines the purposes and means of data processing". As shown in the definition the legal form is not decisive for being considered responsible for the legal obligations under the GDPR.[118] It is the controller who is responsible and liable for any processing of personal data carried out by itself or on its behalf.[119]

As stated in the definition, the controllership can be jointly with others. Joint controllers exist when two or more determine the purpose and means of processing.[120]

The last component of the definition states that it is the one who determines the purpose and means of the processing that shall be the controller.[121] Controllership does not depend on the execution of the data processing but upon decision-making power.[122] The controller decides which data shall be processed, for how long someone shall have access and what security measures need to be taken. It is a factual approach towards the determination of which one is controller. It is not the one which can be considered to lawfully determine, rather the one which actually determines the purpose and means of processing shall be considered to be controller.[123]

---

[116] Art. 7.3 GDPR
[117] Art. 4.7 GDPR
[118] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 18
[119] Art. 24 GDPR
[120] Art. 26 GDPR
[121] Art. 4.7 GDPR
[122] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 19
[123] Art. 29 Data Protection Working Party. (2010) Opinion 1/2010, on the concepts of "controller" and "processor", WP 169, p. 9

2.2.5 Processor

Besides controller, the GDPR establish data protection obligations on the "processor". The GDPR defines the processor as:

"…*a natural or legal person, public authority, agency or other which processes personal data on behalf of the controller*."[124]

As aforementioned the controller is allowed to delegate the processing to a processor under certain conditions.[125] The processor is a separate legal entity or individual which acts on behalf of the controller.[126] Thus, the bare existence of the processor depends on a decision taken by the controller. The processor must be governed through a contract or another legal act.[127] The processor is not however a third party according to the GDPR.[128] This allows the processor to take a privileged position since its involvement with the controller does not require additional consent or statutory justification.[129]

2.2.6 Pseudonymization

One way of ensuring an adequate level of data protection is the use of pseudonymization. This is when the processing of personal data is done in such a manner that the personal data cannot be attributed to a specific data subject without the use of additional information.[130] This is done through replacing the name or the characteristics of the data with certain indicators. The use of pseudonymization does not imply that the natural person no longer is identifiable and is still considered to be personal data.[131] Instead it is considered to constitute information of individuals which are indirectly identifiable.[132] Mainly, the usage of pseudonymization is to reduce the risk of the data subject in question and help the controller and processor meet their data-protection obligation.[133] The effectiveness of the pseudonymization procedure depends on a number of factors e.g., when it is implemented and how secure it is against reverse tracing.[134]

[124] Art. 4.8 GDPR
[125] Such conditions are established in Art. 28 of the GDPR and includes delimitation in what is supposed to be processed, adopting equal security measures as the controller etc.
[126] Art. 28 GDPR
[127] Art. 28.3 GDPR
[128] Art. 4.10 GDPR
[129] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 80
[130] Art. 4.5 GDPR
[131] Rec. 26 GDPR
[132] The Working Party, Opinion 4/2007, p. 18
[133] Rec. 28 GDPR.
[134] The Working Party, Opinion 4/2007, p. 18

An example of pseudonymization is *key coded data*, most used within clinical trials in medicine. To be able to identify the natural persons in key coded data you have to use the key together with the pseudonymized data. In such an example, the researchers are the only one with the "key" to the pseudonymized data and other parties, such as the pharmaceutical companies, only has access to the data in pseudonymized form without access to the key. In this situation, is the result data still considered to be personal data? The Working Party argues, as aforementioned, that the mere hypothetical possibility of identification is enough for it to constitute personal data.[135] The Working Party continues regarding this example that the fact that the pharmaceutical companies do not have access to the key for the data does not affect the assessment. Since it is within the purpose of the processing to identify the data subject it shall be considered personal data. Due to this, the key coded data in the example constitutes personal data, as all pseudonymized data do.[136]

2.2.7 Anonymization

Anonymization is the alteration of personal data which results in the no existence of a strong link between the data and an individual.[137] There are two kinds of anonymized data, namely data that has no relationship to an identified or identifiable individual and personal data that has been changed in such a way that the natural person is not or no longer identifiable. It is desirable to achieve anonymization since it will render the GDPR not applicable.[138]

Anonymization can be achieved through several different techniques. However, most of these techniques fall within two main categories. The first category is called "Randomization" which is when the accuracy of data is altered in order to remove a strong link between the data and the individual. The second category is called "Generalization" which is when you generalize/dilute the attributes of the data subjects by altering the respective scale or order of the data.[139]

The Working Party highlights three risks that are essential to anonymization. The first of these risks are "Singling out". This is the possibility to isolate some or all records which identify an individual in the data set. The second risk is "linkability" which is the ability to link two or more records concerning the same data subject or group of data subject. The third risk is

[135] The Working Party, Opinion 4/2007, p. 15
[136] Ibid., 19
[137] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 13
[138] Rec. 26 GDPR
[139] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 13

"inference" which is the possibility to deduce, with significant probability, the value of an attribute from the value of a set of other attributes. The solution for all these three risks would provide, according to the Working Party, robust protection against re-identification performed by the most likely and reasonably means used by the data controller or anyone else.[140]

According to the Working Party, the use of pseudonymized data as anonymized data is one of the most common pitfalls regarding anonymized data.[141] The Working Party argues that merely replacing one or more attributes to a dataset does not make it anonymous. Quasi-identifiers are combinations of attributes relating to a data subject, or a group of data subjects.[142] As long as these quasi-identifiers remain in the dataset or if the values of other attributes are still capable of identifying the data subject, it cannot be anonymized data.[143] The Working Party adopts an approach towards anonymization which is achieved through processing of the personal data to irreversibly prevent identification and achieve irreversible de-identification.[144]

The aforementioned stance by the Working Party regarding the definition of anonymous data is something that has been criticized.[145] Stalla-Bourdillon and Knight proposes a different view on personal data and anonymization in contrast to what the Working Party has concluded.[146] They propose a dynamic approach to anonymization which combines elements from a harm-based, risk-based and procedure-based approach. *The harm-based approach* focuses on assessment of specific privacy harms due to poor de-identification, *the risk-based approach* advocates that the definition of anonymous data should be based on a case-by-case assessment by utilizing an ex-ante evaluation of the potential risk of re-identification. The last aspect of Stalla-Bourdillon and Knight suggestion is *the procedure-based approach* which argues that the law could be designed around the processes that are necessary to lower the risk of re-identification and sensitive attribute disclosure.[147] They argue that the line between personal

[140] Art. 29 Data Protection Working Party. (2014) Opinion 5/2014, on Anonymisation Techniques, WP 216. p. 12
[141] The Working Party Opinion 5/2014, p. 10
[142] Ibid., p. 12
[143] Ibid., p. 21
[144] Ibid., p. 3
[145] See S. Stalla-Bourdillon and A. Knight, *Anonymous data V. personal data – a false debate: an EU perspective on anonymization, pseudonymization and personal data,* (2017), Wisconsin International Law Joural 34 (2), and M. Finck and F. Pallas, *They who must not be Identified – Distinguishing Personal from Non-Personal Data under the GDPR*, Max Planck Institute for Innovation and Competition Research Paper Series. 2020
[146] S. Stalla-Bourdillon and A. Knight, *Anonymous data V. personal data – a false debate: an EU perspective on anonymization, pseudonymization and personal data*
[147] Ibid., p. 307 ff.

and non-personal data should be fluid and change over time. Thus, allowing data to be anonymous but acknowledging that said anonymous data can become personal data in the future, making the definition depend on the context.[148] Furthermore, they argue that as we evolve into a more open data world and thus, to make this sustainable, we have to move away from the release-and-forget model.[149]

Finck and Pallas argues that the Working Party provides an absolutist approach toward anonymization that is more far reaching than the relative risk-based approach provided by recital 26 of the GDPR.[150] Finck and Pallas continues to argue that ECJ has established the risk-based approach in the GDPR through the Breyer case since the court evaluated the actual risk of identification.[151] They continue to highlight that if an absolutist approach were adopted it could effectively rule out the existence of anonymous data since there will always be parties able to combine a dataset with additional information that may re-identify it.[152] Something that is backed up in the article by Stalla-Bourdillon and Knight.[153] According to Finck and Pallas, anonymization can be fashioned in such a way as a means of reducing the risks which data processing generates in regards to the freedom and rights of the individual and at the same time highlighting the lack of absoluteness regarding anonymization.[154] Thus, enhancing the possibilities for personal data to be deemed as anonymous in a world with ever enhancing possibilities of re-identification of natural person through the use of anonymous data. This approach, according to Finck and Pallas, would be better suited in the dynamic world of technology and a risk-based approach would give more incentive for controllers to strive after anonymization if this is based on the risk of re-identification rather than the possibility of re-identification.[155]

[148] S. Stalla-Bourdillon and A. Knight, *Anonymous data V. personal data – a false debate: an EU perspective on anonymization, pseudonymization and personal data*, p. 318
[149] Ibid., p. 320
[150] M. Finck and F. Pallas, *They who must not be Identified – Distinguishing Personal from Non-Personal Data under the GDPR*, p. 7
[151] Ibid., p. 14
[152] Ibid., p. 15
[153] S. Stalla-Bourdillon and A. Knight, *Anonymous data V. personal data – a false debate: an EU perspective on anonymization, pseudonymization and personal data*, p. 320
[154] M. Finck and F. Pallas, *They who must not be Identified – Distinguishing Personal from Non-Personal Data under the GDPR*, p. 45
[155] Ibid., pp. 45-46

2.2.8 Rights of the Data Subject

The GDPR gives the data subject certain rights regarding its personal data. These are stated in the third chapter of the GDPR. Data Processing can negatively impair the rights and freedoms of the data subject. This is especially true where the processing is unlawful or where it involves incorrect or incomplete data. To remedy the possible negative effect of incorrect data processing, the GDPR provides different rights of data subjects which permits them to limit or influence the processing activities carried out by the controller.[156] Two of these rights that sprung out of the aforementioned factors are the Right of Rectification (ROR) and the Right of Erasure (ROE).

2.2.8.1 Right of Rectification

The ROR gives the data subject the right to rectify inaccurate personal data concerning the data subject.[157] This right reflect the principle of accuracy of the GDPR according to which the processed data, at any given time, shall reflect reality. An example of where this comes into play and how it affects the data subject is when the creditworthiness data of an individual is saved incorrectly and as a result, said individual is denied credit. As stated, the ROR only applies to inaccurate data, and it is the data subject which bears the burden of proof.[158] The personal data is deemed to be inaccurate based on facts. It is unclear whether value judgements can be deemed to be inaccurate. Here, a tension between the interests and rights of the data subject and the freedom of opinion by the controller occurs. In these cases, a balancing of the different interests has to be carried out and judged case by case.[159] The ROR also applies to incomplete data. When the data is incomplete, the data subject has the right to have the data completed through a supplementary statement.[160]

2.2.8.1.1 Peter Nowak V. Data Protection Commissioner

The Nowak case[161] was concerning whether answers given on an examination constitute personal data and if so, how the rights of the individual were to be applied in this situation. The

[156] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 154
[157] Art, 16 GDPR
[158] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 154
[159] Ibid., p. 155
[160] Art. 16 GDPR
[161] Case C-434/16, *Peter Nowak v Data Protection Commissioner*

ECJ concluded that the written answers constitute personal data.[162] The ECJ through the Nowak case established a purposive approach regarding the ROR. The ECJ concluded that "in the assessment whether the personal data is accurate and complete must be made in the light of the purpose which the data was collected."[163]

2.2.8.2 Right of Erasure

According to the ROE, the data subject has the right to demand the erasure of its personal data if one of several grounds laid out by the GDPR applies.[164] One of these grounds of erasure is if the data subject withdraws its consent of which the processing is based.[165] Since the ROE establishes a right for the data subject to demand the erasure of its personal data it also establishes an obligation for the controller to erase said personal data. It is the relationship of the corresponding right and obligation which establish that it is the data subject which has the burden of proof regarding the grounds of erasure.[166]

There are however exceptions to the ROE established by the GDPR. Such exemptions include, but are not limited to, situations regarding the exercise of the right of freedom of expression and information, public interests in the area of public health or compliance with legal obligations.[167]

There is no legal definition of erasure provided in the GDPR. The Oxford dictionary establishes that the term "erasure" is defined as "the removal of writing, recorded material or data" and "the removal of all traces of something".[168] In the case *Google Spain*[169], the delisting of result from the search engine was considered to be enough to achieve "erasure".[170] In this specific case the erasure from the result list of the search engine was however all that was requested by the claimant.[171] In the case *Nowak* it would appear that the ECJ indicated that erasure implies the destruction of personal data.[172] In its core, erasure consists of making the data unusable for

[162] Case C-434/16, *Nowak.* para. 62
[163] Ibid., para. 53
[164] Art. 17 GDPR
[165] Art. 17 (b) GDPR
[166] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 159
[167] Art. 17.3 GDPR
[168] https://www.lexico.com/definition/erasure
[169] Case C-131/12, *Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja Gonzaléz*
[170] Ibid., Para. 99
[171] Ibid., Para. 15
[172] C-434/16, *Nowak*, Para. 55

the controller, processor or a third party from accessing, processing, and reading out the data. If the data is erased, whether it be by physically or technologically destroying it, it is of paramount importance for it to achieve the level needed to be deemed erased, that it cannot be restored without excessive effort. However, the mere theoretical possibility of restoring the data is not enough to render it not erased.[173]

The following case brought the *Right to be forgotten* to great public attention. The Right to be forgotten has since then been strengthened through the GDPR.[174] It entails that when a controller has made the personal data public and the controller is later obliged to erase said data, the controller is furthermore obliged to take reasonable steps to inform other controllers that the data subject has requested to have the personal data erased.[175]

2.2.8.2.1 Google Spain V. Mario Costeja González

In the case of *Google Spain*[176], a Spanish citizen, Mario Costeja Gonzáles (Mr. González), lodged a complaint against Agencia Española de Protección de Datos (AEPD), a daily newspaper named La Vanguardia as well as against Google Spain and Google Inc. The complaint was based on the fact that Mr. González, when entering his own name in the Google search engine, obtained links to two pages of La Vanguardia's Newspaper on which Mr González name appeared together with the real-estate auction connected to Mr. González social security debts. Firstly, Mr. Gonzáles requested that La Vanguardia be required either to remove or alter those pages so that the personal data relating to him no longer appeared. Secondly, Mr. González requested Google Spain or Google Inc. to be required to remove or conceal the personal data relating him. The AEPD rejected the part regarding the request towards La Vanguardia but upheld the part relating to the request toward Google Spain or Google Inc and went to Audiencia Nacional (National High Court) in Spain which asked the ECJ for a preliminary ruling.[177]

Through its ruling, the ECJ established the role of the operator of a search engine as a controller of personal data since it is the search engine which determines the purpose and means of the processing.[178] Furthermore it would be contrary to the objective of the Directive to interpret the

[173] P. Voigt and A. Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide*, p. 161
[174] Ibid., p. 156
[175] Art. 17.2 GDPR
[176] Case C-131/12, *Google Spain*
[177] Ibid.
[178] Ibid., Para. 32

definition of a controller as to exclude the operator of a search engine.[179] The ECJ continued to establish, in contradiction to what Google submitted regarding the virtues of the principle of proportionality, that the operator of a search engine are obliged to upon request from the data subject, remove results displayed following a search made on the basis of a person's name as long as those webpages contains information relating to that person.[180] The ECJ ruled that in the light of the data subjects fundamental rights under Articles 7 and 8 of the Charter, the request that the information regarding that data subject be made no longer available to the public trumps the economic interest of the operator of a search engine.[181] The interest of the data subject also trumps the interest from the general public in having access to that information upon using the data subjects' name in a search engine.[182]

3 Blockchain

3.1 Introduction to Blockchain

Blockchain technology does not have a definite definition. Blockchain is not a form of software but instead it is supposed to constitute a class of technologies which operates on a spectrum of different technical and governmental structures.[183] Blockchain is a working peer-to-peer system and was first achieved and introduced by Satoshi Nakamoto through the publication of the paper "*Bitcoin: A Peer-to-Peer electronic cash system*" in 2008.[184] Nakamoto was not the originator of many mathematical discoveries used in the paper. Nearly all the technological components of Nakamoto´s Bitcoin originated from academic research that date back to the 1980s and 1990s.[185] What Nakamoto set out to, and succeeded, to create was an electronic payment system based on cryptographic proof instead of trust.[186] Therefore allowing transactions to be made in a low-trust environment without the need for a trusted third party, resulting in an achievement of a decentralized transactional system.[187]

The potentiality of Blockchain cannot be questioned. According to Deloitte's 2021 Global Blockchain Survey, more than 76% of the respondents believes that digital assets, which are

[179] Case C-131/12, *Google Spain.*, Para. 34
[180] Ibid., Para. 88
[181] Ibid., Para. 99
[182] Ibid.
[183] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* p.3
[184] S. Nakamoto, *Bitcoin: A Peer-To-Peer Electronic Cash System*, (2008), https://bitcoin.org/bitcoin.pdf
[185] A. Narayanan and J. Clark, *Bitcoin's Academic Pedigree*, 60 Communications of the ACM p. 36
[186] S. Nakamoto, *Bitcoin: A Peer-To-Peer Electronic Cash System*, (2008), P. 1
[187] Ibid.

based on Blockchains, will serve as a strong alternative to traditional fiat currencies in the next 5-10 years.[188] The potentiality can be exemplified through the Blockchain based "Non-Fungible Tokens" (NFT).[189] In March 2021 the total sale of NFT was almost 400 million US dollars, constituting a 400% increase over the last months numbers.[190]

The EU goal to be a leader within Blockchain technology was declared in the "Blockchain Strategy" policy including making EU an innovator of Blockchain and home to significant platforms, applications, and companies.[191] The Commission has further established a "gold standard" including aspects of environmental sustainability, cybersecurity, and data protection for the Blockchain to be compatible with. This golden standard by the Commission establishes the minimal requirements of which the Blockchains of Europe has to be compatible with. This includes compatibility with the strong data protection regulations of the EU[192]

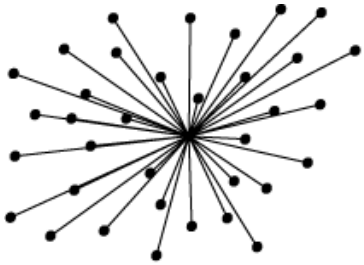
3.2 The architecture of the Blockchain

When designing a software, one of the most fundamental decisions regarding the implementation of the system is in which way its components are organized and related to one another, i.e., its architecture.[193] When talking about the architecture of the implementation of a software, the two major categories are 1) centralized and 2) distributed. In a distributed architecture none of the components are connected with all other components directly but

[188] Deloitte's 2021 Global Blockchain Survey: *A new age of digital assets*, 2021, https://www2.deloitte.com/content/dam/insights/articles/US144337_Blockchain-survey/DI_Blockchain-survey.pdf, p. 5
[189] NFT are certificates that prove that you own something digital.
[190] Future of Blockchain: How will it revolutionize the world in 2022 & Beyond, The European Business Review, 2021, https://www.europeanbusinessreview.com/future-of-blockchain-how-will-it-revolutionize-the-world-in-2022-beyond/
[191] The European Commission: Shaping Europe's digital Future, *Blockchain Strategy*, https://digital-strategy.ec.europa.eu/en/policies/blockchain-strategy
[192] The European Commission: Shaping Europe's digital Future, *Blockchain Strategy*
[193] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, Main, APress, 2017, p. 10

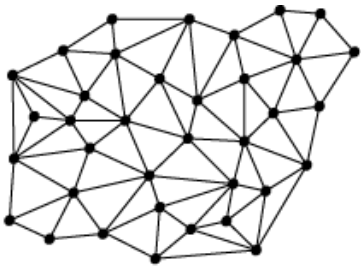
instead indirectly, while in a centralized architecture every component is connected to one central component.[194] Both architectural types are illustrated in figure 1 and 2.

Figure 2



centralised

Figure 1



distributed

The Blockchain structure can have different reading and writing access. It can be public or private, permissioned or permissionless. The Blockchains which are relevant for this thesis are public permissionless Blockchains. These kinds of Blockchains are open for everyone to read, write and create transactions.[195]

3.2.1 Peer-to-Peer

A certain kind of a distributed system is called Peer-to-Peer (P2P). Such a system consists of several individual computers (nodes). These nodes constitute the systems computational power (e.g., storage capacity, processing power etc.) and make it available for every member of that specific system, without the need of a central point of coordination.[196] Nodes are equal towards each other regarding their rights and roles in the system. Within these purely distributed P2P systems, all the nodes perform the same tasks, acting both as providers and consumers of resources and services without any element of control or coordination.[197] P2P is connected to Blockchain as the P2P system uses Blockchain as a tool to achieve and maintain integrity in a trustless environment.[198]

[194] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 11
[195] Ibid.., p. 216
[196] Ibid., p. 14
[197] Ibid., p. 23
[198] Ibid., p. 24

3.3 Ownership in a Blockchain

The documentation of ownership within a Blockchain can be contributed to two major aspects. These are 1) describing the transfer of ownership and 2) maintaining the history of transfers.[199]

3.3.1 Transfer of ownership

A transaction of ownership relies on data that describe the intended transfer. Data that contain all the necessary information for the transaction to be able to be executed.[200] The transfer of ownership within a Blockchain can be resembled to the transaction of funds between bank accounts. The bank requires you to provide the information necessary to perform the transaction, similar requirements are set within a Blockchain. One will have to provide an identifier of the account that is to hand of ownership to another account, an identifier of the receiving account, the amount of the goods to be transferred, the time the transaction is to be done, a fee to be paid to the system for executing the transaction and finally a proof of the owner of the account that hands of ownership indeed agree with that transfer. The main difference between a transfer within a bank and a transfer of ownership within a Blockchain is that a bank will have a central fee schedule, while within a Blockchain each user has to tell the system in advance how much it is willing to pay to have the transaction executed.[201]

3.3.2 Maintaining the history of transfers

Proving the ownership within a Blockchain is designed differently than "normal" banks and other institutions. Within a Blockchain, the ownership is not proved through information that is about the current state of affairs. Instead, the Blockchain utilizes the transaction history, which is stored in "ledgers" to correctly identify the ownership.[202] Executing a transaction means, in this context, to add the transaction to the ledger which is then used to clarify ownership within the Blockchain.[203]

To be able to back up the ownership with the use of transaction data, Blockchain maintains the whole history of all transactions that have happened. They do this through storing the transaction data in the Blockchain-data-structure in the exact order which they have occurred.[204]

---

[199] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps,* p. 65
[200] Ibid.
[201] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*., p. 66
[202] S. Nakamoto, *Bitcoin: A Peer-To-Peer Electronic Cash System*, p. 1
[203] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 66
[204] Ibid.

Once a node get information about a new block of transaction which is supposed to be added to the Blockchain, it checks first if its valid. If the block is valid the node alters its own local copy of the Blockchain to add the new block and then broadcast this change to the other nodes.[205] This way, they continue to have an updated version of the ownership within the Blockchain.

As shown, the transaction history is in the very heart of the Blockchain structure. It is therefore imperative to protect the integrity of the transaction history to be able to provide a true statement of ownership.[206]

3.4 Distributed ledgers

As aforementioned, Blockchain works in a low trust environment without the possibility of centralized exercised control or authentication. As such, the Blockchain prove ownership through the transaction history. This form of mapping between owners and objects is typically done with a ledger, which has to be constantly updated as new information accrue. The ledger both fulfill as the means to prove ownership, as well as documenting any transfer of ownership.[207]

The ledger can be compared to witnesses in a court of law. Unless you have a witness, who provenly speak the truth or what all parts of the court know is the truth (centralized control), the more witnesses you have the more reliable your story is. Applying this on the ledgers we utilize in the context of a Blockchain that operates in a low trust environment, there is a certain need for many witnesses. As aforementioned, one of the main features of Blockchain is its decentralized structure which implies that there is no need for a trusted third party. The Blockchain substitute the third party with purely distributed ledgers across the P2P system. Every node which stores one version of the ledger fulfill the same role as the "witnesses", and the ledger with the history of transfers which the majority of the "witnesses" agree on, shall constitute the correct history of transfers.[208] These ledgers are distributed and stored in different capacities of the entire Blockchain, on the different nodes.[209] This aspect of the distributed

[205] J. Bacon et al, *Blockchain Demystified: A technical and legal introduction to Distributed and Centralized Ledgers,* 25 Rich. J.L. & Tech ., no. 1, 2018 P. 19
[206] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps,* p. 67
[207] Ibid., p. 45
[208] S. Nakamato, *Bitcoin: A Peer-to-Peer Electronic Cash System,* p. 2
[209] J. Bacon et al, *Blockchain Demystified: A technical and legal introduction to Distributed and Centralized Ledgers,* p. 19

ledgers allow the Blockchain to achieve strong resilience. If one or several nodes fail or become destroyed, the ledger and the Blockchain will remain unaffected due to the existence of several ledgers.[210]

3.5 Hashing

Blockchains use a technique called "hashing" for many different aspects. One of these are to show the integrity of the Blockchain.[211] The hash, or cryptographic hash value, can be seen as the digital equivalent to fingerprints due to its collision resistant features. The fact that the hash function is collision resistant implies that it is highly improbable to find two or more distinct pieces of data for which it yields the identical hash value. Such a collision is the digital equivalent to having two people with identical fingerprint.[212]

Hashing is when you put the contents of a data item through a so-called "Hashing function", which creates a string of digits which is unique for that data input. The output of the hashing function is called "hash value" and it is practically impossible for two different data items to have the same hash value.[213] Hashing is a deterministic system which means that the same input always yields the same output.[214] It is also pseudorandom, meaning that the hash value returned by a hash function changes unpredictably when the input data are changed.[215]

This can be illustrated with following example:

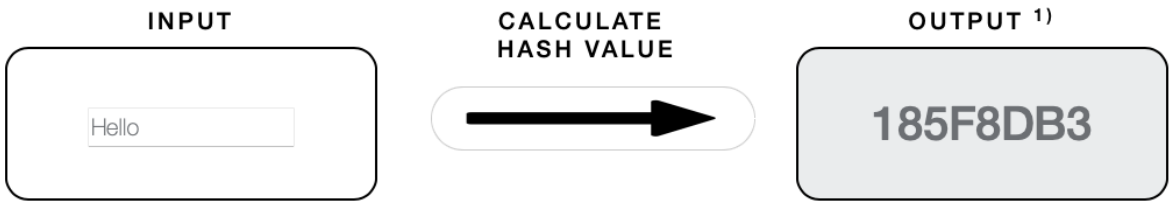


Figure 3

[210] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* p.3
[211] J. Bacon et al, *Blockchain Demystified: An introduction to blockchain technology and its legal implications,* Queen Mary University of London, School of Law Legal Studies Paper No. 268/2017 p. 7
[212] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps,* p. 73
[213] Bacon, Michels, Millard, Singh, *Blockchain Demystified: An introduction to blockchain technology and its legal implications,* p. 6
[214] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* p. 29
[215] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps,* p. 73

As shown in figure 3, the input “Hello” into a hash function becomes the output “185F8DB3”. If we then change the input by just adding an exclamation point, we get the following:

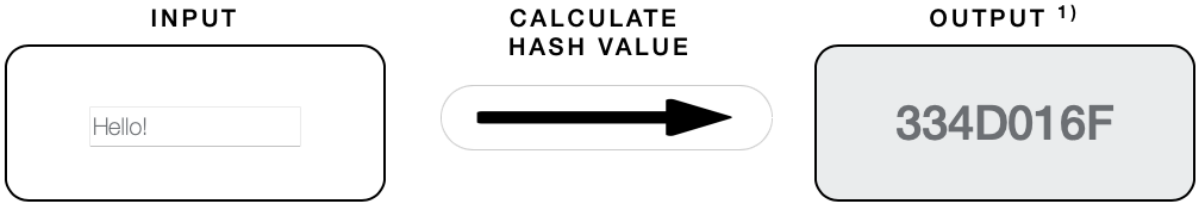


Figure 4

The output in figure 4 changes significantly from the output in figure 3. From the previous hash “185F8DB3” to “334D016F” by only adding an exclamation point. This factor allows hashing to be used to prove integrity of the input data due to that even a change in a single character or space will produce an unrelated hash output.[216] In the context of Blockchain, this works by the creation of the cryptographic hash value of the data that are supposed to stay unchanged within the Blockchain. When you need to verify given data to make sure that it has not been changed, you enter the original data once again in the hash function and then compare the two hash-outputs.[217] If they do not match, the data have been altered.

3.5.1 Change-sensitive storing

The hash function can also be used to store data in a change sensitive manner. This is done through the use of so called “hash references” that point to other data, that also contains a hash reference to another part of data etc.[218] These hash references (or hash pointers) can be used to prove integrity in a string of documents. [219] A Blockchain utilizes two structures, one of these has a chain like structure and the other has a tree like structure.

[216] J. Bacon et al., *Blockchain Demystified: An introduction to blockchain technology and its legal implications,* p. 7
[217] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 82
[218] Ibid., 86
[219] J. Bacon et al., *Blockchain Demystified: An introduction to blockchain technology and its legal implications,* p. 7

3.5.1.1 The Chain

The chain is structured in such way that the first entry of data has no hash reference, the second entry of data has a hash reference to data entry one and so on as exemplified in the figure 5.[220]

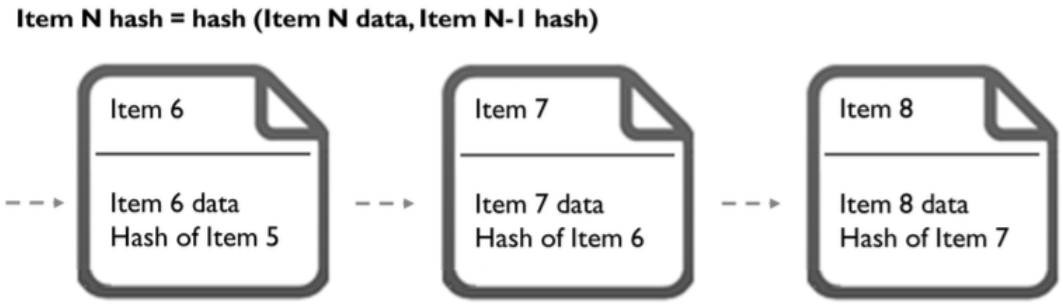


Figure 5

As shown in figure 5, the data of each item together with the hash reference of the previous item becomes the input data in the hash function. The whole chain starts with only the "item" as it is referenced to in the example meaning no hash references. The second item has its own data as well as the hash reference to item one and so on. A chain like structure is useful when the data that is supposed to be stored and linked together is not available all at once, but instead is added gradually.[221]

3.5.1.2 The Merkle Tree

The second way transactions can be linked together is through a tree like structure. Such a structure is called a "Merkle Tree".[222] The special significance of the Merkle Tree is the allowance of the possibility to group distinct pieces of data or transactions that are available at the same time and make these accessible through a single hash reference.[223]

[220] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 87
[221] Ibid.
[222] Ibid., p. 88
[223] Ibid.

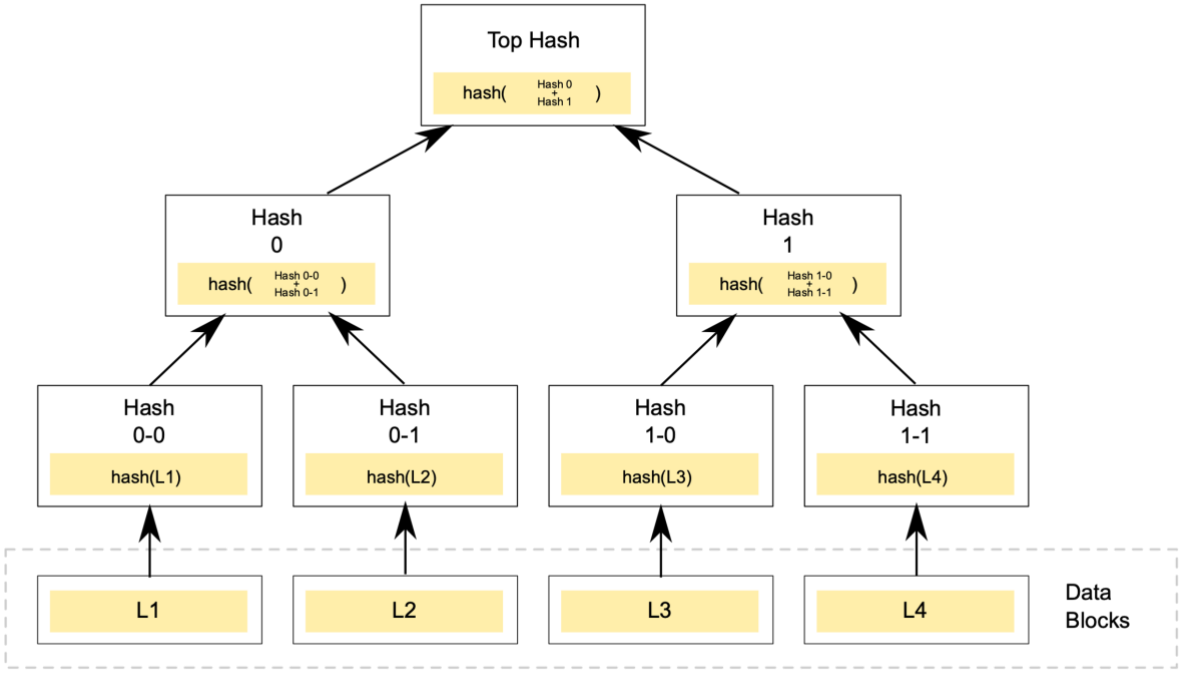


Figure 6

As shown in figure 6, you start with four different entries of data, an individual hash reference is then produced for each data entry. These hash references are then grouped together and another hash reference to the group of the two hash references is created. This is later done one last time through the grouping of the top two hash references into the "Top hash" as referenced in figure 6. The Top hash is also called the root of the Merkle Tree.[224]

These are two kinds of data structures which store data in a change-sensitive fashion. This since the hash references which both structures are built upon are change sensitive. If any data which these hash references are based upon changes, it will have the result of destroying the link throughout the structure.[225]

3.5.2 Storing Transaction in a Blockchain

The Blockchain store transactions in a tamper evident way through its combination of the two structures of chain-sensitive storing. Each block contains a "block header" and the block header contains the hash references. If one of the blocks in this example is the first block of a Blockchain, the block header will contain the Merkle Tree root which, as explained above,

[224] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 87
[225] Ibid., p. 88.

contains hash references to several different data/transactions.[226] The second block has the same structure, a block header with the hash reference which is the Merkle Tree root. But, as within the chain structure of hash references, because it is the second block of the Blockchain, it contains a hash reference to the anterior block. Therefore, the second block header has a hash reference to the Merkle Tree, and a hash reference to the previous block. This system follows for every subsequent block with a hash reference to the transaction data which is represented within the Merkle Tree, and a hash reference to the previous block.[227] Within the block header there is also metadata e.g., a timestamp. The timestamp works so that everyone within the Blockchain can verify that the data must have existed at that time and therefore make it possible to provide a correctly organized transaction history.[228]

3.5.2.1 Off-chain data storage

As shown above, storing the data on Blockchain implicates that it is all referenced to each other through hash references. An alternative solution to storing data on the Blockchain is to do so off-chain instead. This is referred to as *off-chain data storage*. In this solution, the data at hand would be stored in an off-chain, encrypted, and modifiable database instead of on the Blockchain.[229] Within the Blockchain there would be a hash reference such as in the Merkle Tree instead of the data, which in turn will reference to the data stored off-chain. [230] The employment of off-chain data storage is however a step away from one of the core fundamentals of Blockchain, namely the no necessity of a trusted third party, which would have to be employed to maintain the encrypted database.[231]

One benefit of utilizing an off-chain data storage is the possibility of altering the data which is stored in the off-chain data storage. This allows for the Blockchain to be immutable since the data which is on the Blockchain will not be altered.[232] The hash reference will remain unaltered

[226] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 121
[227] Ibid.
[228] S. Nakamato, *Bitcoin: A Peer-to-Peer Electronic Cash System,* p. 2
[229] M. Finck, *Blockchains and Data Protection in the European Union,* European Data Protection Law Review, 2018, p. 23
[230] Ibid.
[231] G. Zyskind, O. Nathan and A. Pentland, *Decentralizing privacy: Using Blockchain to Protect Personal Data*, IEEE Security and Privacy Workshops, 2015, p. 181
[232] See further in chapter 3.6

in the Blockchain even if the data it is referring to is erased.[233] It is important to note that everything within a Blockchain cannot be stored off-chain. Such is the case with public keys.[234]

3.5.3 Asymmetric cryptography – Personal and Private keys

As stated, Blockchain constitutes a public permissionless P2P system. Because of this, the issue of protecting property of the specific user arises. It is the transaction history which confirms ownership within a Blockchain. One crucial part of ownership is the exclusivity, that it is only the owner who can transfer ownership. A problem to protect the property assigned to accounts arises when doing this without limiting the open architecture of the public, permissionless P2P system. If not addressed, an attacker would be able to pose as a third party and propose new transactions to an account the attacker would control.[235] One alternative to remedy this issue is the utilization of asymmetric cryptography.

To be able to understand how this works, a short detour into cryptography must be made. One can simplify cryptography by viewing encryption as a lock and decryption as the key to that lock. Encrypted data is called cypher text and looks just like a pile of useless letters for anyone who cannot decrypt it. To be able to decrypt it you must have the key to that specific encryption. This decrypted cypher text is identical to the original data. There are two forms of cryptography, 1) symmetric and 2) asymmetric. In the case of symmetric cryptography, the same key is used both for encryption and decryption of the data.[236]

Asymmetric cryptography uses two complementary keys in its structure. To exemplify this, the first key will be named "key one" which is used in this example to encrypt data. Unlike keys in symmetric cryptography, key one cannot be used to decrypt that data. This cypher text can only be decrypted with the second key, named "key two". This also works the other way around, key two can encrypt data, but it is only key one which is able to decrypt it. It will naturally also result in two different types of cypher texts depending on which key is used to encrypt.[237]

Asymmetric cryptography is normally referred to as public-private-key cryptography even if there are no one public or private key in asymmetric cryptography per se. The reference is due

[233] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* p. 32
[234] M. Finck, *Blockchains and Data Protection in the European Union,* p. 25
[235] J. Bacon et al., *Blockchain Demystified: An introduction to blockchain technology and its legal implications,* p. 14
[236] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps,* p. 96
[237] Ibid., p. 97

to the roles assigned to the two different keys. The public key is given out to the public, regardless of their trustworthiness or any other factor while the private key is kept safe. Asymmetric cryptography can be done in two different ways, 1) Public to Private, 2) Private to Public. The first alternative is when information flows from the private key towards the public. The second alternative is instead when the information is flowing from the public, where it is encrypted, to the private where it is decrypted.[238] The second alternative can be resembled to a mailbox where everyone can put letters in but only the owner is able to access the mail inside the mailbox. This way of using asymmetric cryptography is focused on sending information in a secured fashion to the owner of the private key.[239]

Blockchain utilize the asymmetric cryptography Public-to-Private for identifying user accounts and transferring ownership between them. The account numbers within the Blockchain constitutes the public cryptographic keys, the Blockchain then identifies these public keys involved in a transaction and utilizes them in a similar to how one would utilize a mailbox.[240]

3.5.3.1 Stealth addresses

In its publication, Nakamoto recommended the users of Bitcoin to use a new pair of public keys for each transaction in order to put up another safeguard to guarantee privacy.[241] This practiced is referred to as using "stealth addresses", which in reality implies that for each transaction the total amount of that account must be emptied.[242] It is also recommended that every user, when utilizing stealth addresses, should generate a new private key with each new account.[243] The purpose of this method is to reduce the linkability of the public key. This can be exemplified. Imagine there exist 0.0001 Bitcoins (BTC) on a public key, and 0.00003 BTC of these are being transferred to another public key. At the same time these 0.00003 BTC are being transferred, the remaining 0.00007 BTC will be transferred to a new account which the originator of the transaction controls.

[238] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*., p. 99
[239] Ibid.
[240] Ibid.
[241] S. Nakamato, *Bitcoin: A Peer-to-Peer Electronic Cash System*, p. 6
[242] M. Finck, *Blockchains and Data Protection in the European Union*, p. 25
[243] V. Buterin, *Privacy on the Blockchain*, (2016), https://blog.ethereum.org/2016/01/15/privacy-on-the-blockchain/

3.6 The Immutability of Blockchain

The importance of the transaction history cannot be mentioned too many times, certainly not as it is used to prove ownership within the Blockchain. Since Blockchain operates in a public permission P2P system, there is a real possibility that dishonest users would try to manipulate the transaction history in their favor.

The main idea with the way Blockchain addresses this issue is making the transaction history immutable, since data that is considered to be immutable cannot be changed once it is entered.[244] As discussed, the Blockchain utilizes a structure that store data in a change-sensitive manner. If an attacker or a dishonest user would change the data entry into one block, due to the collision resistant features of the hash function, this would yield a whole different output of the hash function. Thus, if any of the data inside one block would have been changed, it would effectively break the link between the different blocks from the invalid change and forward.[245] This way, it is impossible to change or manipulate the data that is within the Blockchain without anyone noticing it.[246] This change-sensitive storage solution utilized by Blockchain mandates that if a fraudulent change would be made, all subsequent items need to be re-hashed for the link to be upheld.[247]

In addition to this "all or nothing requirement" by the Blockchain, it utilizes a system which makes the computational costs high enough for a manipulation of the transaction history to be unprofitable to perform.[248] This is accomplished through the use of adding a hash puzzle, i.e., a form of computational problem to every block.[249] This implies that in addition to "only" re-hashing the whole chain, an attacker or malicious node would also have to conduct the proof-of-work[250] which these hash puzzles constitute. As such, the computational cost to manipulate the transaction history makes it severely unattractable to manipulate, and as such, allowing the ledgers to become append-only.[251]

[244] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 137
[245] J. Bacon et al., *Blockchain Demystified: An introduction to blockchain technology and its legal implications*, p. 7
[246] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 138
[247] J. Bacon et al., *Blockchain Demystified: An introduction to blockchain technology and its legal implications*, p. 8
[248] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 138
[249] S. Nakamato, *Bitcoin: A Peer-to-Peer Electronic Cash System*, p. 3
[250] Proof-of-work is a form of consensus protocol of which the acceptance of new blocks is based upon.
[251] D. Dresher *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, p. 141

However, there is still a possibility to alter data that has been added to the Blockchain in a non-malicious way. This requires that a majority of nodes in the system accept the change. The result of this would be that the majority of the nodes would have to verify the legitimacy of every block backwards to the original changed block. [252] In other words unbuild the entire Blockchain block by block and then rebuild it afterwards. A process that is very similar to what a dishonest node or user has to go through. During this time, the Blockchain would be blocked to do its most basic function, adding new transactions to the Blockchain.[253]

4 The relationship between append-only ledgers, the GDPR and society

4.1 The Right of Rectification and Erasure and the Blockchain

The ROR and ROE constitutes two of the fundamental rights given to the data subject by the GDPR. Furthermore, the ROE also includes the right to be forgotten. These rights implicate the alteration or deletion of data as the data subject put forward such demands to the controller.[254] However, the structure given by the Blockchain is one which is embossed by the immutability of data. Where a single alteration would depend on the acceptance from the majority on a multitude of nodes.[255] Even if the alteration would be accepted, during the time of which the part of the Blockchain which has been broken due to the alteration, is re-hashed, the Blockchain would be blocked from performing any activities such as accepting new block to it.[256]

Purely technologically, it is not impossible to comply with the ROE and ROR since every node has the possibility to alter their own ledger.[257] However, this is not straight forward. As shown above, everything within the Blockchain is encrypted.[258] This provides the difficulty for the individual node, upon request, to identify and rectify the data in question. Even if it was possible to identify the data, it is unsure if this would be effective enough to satisfy the requirements set by the GDPR.[259] In theory, all the nodes are able to accept a change in the transaction history for the Blockchain and thus, be able to alter or even delete certain parts of the chain to comply

[252] M. Berberich and M. Steiner, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* 2 European Data Protection Law Review, (2016), p. 426
[253] Ibid.
[254] Chapter 2.2.8
[255] Chapter 3.6
[256] M. Berberich and M. Steiner, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* p. 426
[257] J. Bacon et al, *Blockchain Demystified: A technical and legal introduction to Distributed and Centralized Ledgers*, p. 77
[258] Chapter 3.5
[259] M. Berberich and M. Steiner, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* p. 424

with the ROR and ROE. This is something that becomes increasingly more diffuclt as the network at hand grows larger and more nodes are involved. The level of coordination required, to allow such alteration of the ledger and with such requests is very difficult to achieve in a Blockchain involving a large sum of nodes. As shown, the core issue with the ROR and ROE is in the bare context of Blockchain. The pure nature of the immutability of the Blockchain provides major difficulties. The previously shown issues provides the "traditional" alternatives for the Blockchain to be compliant with the ROR and ROE. These alternatives are unlikely to be functional in reality. This point of view is backed by Berberich and Steiner. According to Berberich and Steiner, the technological and operational implications[260] renders these alternatives not feasible as possibilities to comply with the ROR and ROE.[261]

Following chapters will continue to discuss the individual difficulties and possible solutions in the context of ROR and ROE.

4.1.1 The Right of Rectification – more specific

The purpose of the ROR is to mirror the principle of accuracy of the GDPR.[262] One way of achieving the results of accuracy without changing the original data is providing a supplementary statement per the GDPR.[263] According to the GDPR, this is however only possible when the data is incomplete, and not incorrect as a whole. In the context of the distributed ledgers of a Blockchain, the addition of a supplementary statement does not provide an immediate problem. Anyone that has the possibility to add new data to the Blockchain is able to provide the supplementary statement and therefor comply with the ROR in regard to incomplete data.[264]

In the *Nowak* case, the ECJ stated that whether the data is accurate or not must be assessed based on the purpose of which the data was collected.[265] What the ECJ has established here is a purposive approach toward the ROR which deems that a supplementary statement might not

[260] Chapter 2.2.8.1
[261] M. Berberich and M. Steiner, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* p. 426
[262] Chapter 2.2.8.1
[263] Art. 16 GPDR
[264] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law? p*. 73
[265] Case C-434/16 *Nowak*, Para. 53

always be satisfactory in regard to the ROR.[266] An example of such a situation is where it is not enough to add supplementary data but the original data has to be removed and replaced. In such situations, the problem between the GDPR and the ROR still exists.

4.1.2 The Right of Erasure – more specific

Due to what has been stated previously, the combination of Blockchain and the ROE seems to be unfeasible. There is however leeway given by the GDPR regarding one of the grounds for requesting erasure. The wording of Article 17.1 (a) of the GDPR may allow for the functioning principle of Blockchain to be considered.[267] The core requirement of Blockchain to continuously process the personal data might constitute a legal ground for processing.[268] This could result in that a request of erasure on the basis of Art. 17.1 (a) of the GDPR is unapplicable for data stored on Blockchain.[269] However, the legality of such processing is unclear.

The GDPR has not specified the meaning of erasure, raising the questions of what actions or results can be considered to fall within the meaning. This is highly interesting in the context Blockchain since the technical hindrance of its compliance with the ROE. In *Google Spain*, the delisting of result from the search engine was considered to be enough to achieve "erasure".[270] In this specific case the erasure from the result list of the search engine was however all that was requested by the claimant.[271] This case has been used as a ground to argue that the GDPR obliges the controller to everything they are able to, in order to secure a result as close as possible to the destruction of data within the limits of their own factual possibilities.[272]

The French *Commision Nationale Informatique Libertés* (CNIL) has given out a recommendation in regard to be compliant with the ROE within Blockchain technology.[273] CNIL said, while pointing out that the pure compliance with the ROE in the context of

[266] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* p. 73
[267] M. Berberich and M. Steiner, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* p. 426
[268] How the GDPR views the processing of personal data within blockchain indefinitely is uncertain. This is however not something that will be discussed further in this thesis.
[269] M. Berberich and M. Steiner, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* p. 426
[270] Case C-131/12 *Google Spain*, Para. 99
[271] Ibid., Para. 15
[272] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* p. 76
[273] Commision Nationale Informatique et Libertés, *Solutions for a responsible use of the blockchain in the context of personal data*, (September 2018)

Blockchain is technically impossible, that the destruction of the secret private key would move closer to the desirable and intended effect of the ROE and therefore might be a satisfactory solution.[274] As shown, the asymmetric cryptography in Blockchain makes the public key unusable without the private key.[275] To use the mailbox reference, all of the mails inside the mailbox (the public key) will be inaccessible without the key to the mailbox (the private key). The solution provided by the CNIL is therefore to leave the hashed personal data on the Blockchain, but to make it inaccessible by destroying the secret private key. The CNIL continues to show that for this to be possible, the information has to be deleted in other systems where it has been processed.[276] This last addition by the CNIL highlights the issue of when the private key has been published publicly or when it has not been kept private. These situations are making it problematic for the CNIL solution to be viable. What the CNIL does not consider is however if the public key is deemed to constitute personal data and if that constitution changes when the private key has been destroyed. Something which will be further discussed below.

The suggestion by the CNIL serves as an example of what could be considered to be "erasure" in the meaning of the GDPR even though the data in itself would not be deleted, merely made un-verifiable. The deletion of the private key is called "*Crypto-shredding*".[277] If done correctly, the only way to access the data in question in a de-crypted manner is through the destruction of the encryption. Such destruction, depending on the state of the art of the encryption, can be deemed almost impossible.[278]

4.1.3 Technical alternative to comply with these rights (Off-chain storage)

The main problem with the ROR and the ROE in the context of Blockchain is the way the data-storage of the Blockchain is designed. As shown, the immutability of what has been added to the Blockchain is a strict requirement in order for the Blockchain to be able to function in the decentralized manner it is supposed to. There are however discussions regarding the possibility of alternative technological solutions that might remedy this conflict.

[274] CNIL *Solutions for a responsible use of the blockchain in the context of personal data*, p. 8
[275] Chapter 2.5.3
[276] CNIL, *Solutions for a responsible use of the blockchain in the context of personal data*, p. 9
[277] seald, *Data destruction using crypto-shredding*, https://www.seald.io/blog/data-destruction-using-crypto-shredding
[278] The possibility to destroy the encryption will increase as further technical development is introduced. Thus, there is a possibility that in the future, it will be considered to be possible to destroy the encryption.

One example of such a technological advancement that would *possibly* make it in line with the GDPR is the off-chain data storage structure.[279] The distinction of data as personal or not is not affected in with this solution. However, the possibility for the specific Blockchain to comply with the ROR and ROE increases.[280] Finck points out the off-chain storage solution as the best possibility for a Blockchain to be compliant with these rights of the individual.[281] Through it, the controller is able to alter, rectify and erase the personal data without affecting the Blockchain as the hash reference will remain intact.[282] The off-chain data storage solution will solely facilitate the personal data that is within the transaction data. Any personal data that derives from the Blockchain and not the transaction data will therefore not be remedied by the off-chain solutions. Other information, such as public keys, cannot be stored in an off-chain manner.[283]

4.1.4 Public keys and hash references as personal data

When utilizing off-chain storage, the hash reference will continue to stay unaltered in the Blockchain and will do so even if the data it is referring to is erased. This is also the case with the public key after the private key has been destroyed. It is therefore necessary to inquire if this remaining hash enjoys the status as personal data before the personal data is erased and if that status changes after. Furthermore, a discussion whether the public key will continue to constitute personal data after the deletion of the private key must be held.

4.1.4.1 Does Public Keys constitute personal data?

According to recital 30 of the GDPR, a natural person can be associated to online identifiers that may, combined with others, be used to create profiles of the data subject and identify them. Public keys, as utilized by the Blockchain, hides the identity of the user unless they are combined with other factors. Public keys constitute such identifiers as mentioned in recital 30 of the GDPR.[284] The Working Party has stated in its opinion that in cases where the *prima facie* extent of these identifiers does not allow for identification alone, the natural person can still be identified as this information combined with other would lead to identification.[285] In the *Breyer*

[279] Chapter 3.5.2.1
[280] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* p. 32
[281] M. Finck, *Blockchains and Data Protection in the European Union*, p. 23
[282] Ibid., p. 29
[283] Ibid.
[284] Ibid.., p. 26
[285] Working Party, Opinion 4/2007, p. 13

Case, the ECJ determined that information can constitute personal data even when only a third party has the additional data necessary to identify the person.[286] In such cases, the ECJ determined that what's separates personal data from anonymous data, is if the possibility of combining the two sources constitutes "means reasonably to be used" to identify the data subject.[287] Regarding means reasonably to be used, the Working Party has presented that the technology available for identification when the data starts to be processed and the technological development during the time the data is being processed should be considered.[288] In the context of Blockchain this would suggest that you consider any technological development possible in the future since the Blockchain will process the public keys indefinitely.

A theoretical example provided by Berberich and Steiner where additional information can identify a natural person through public keys are when a transaction for off-chain goods is made.[289] Imagine a person using a Blockchain based currency to pay for a coffee. While the person pays, the cashier, other employees as well as other customers enjoy the possibility of acknowledging the usage of that exact Blockchain based currency. The knowledge of the exact time and which Blockchain based currency the person used, is allowing for the observers to access the public ledger and through the usage of timestamps on the transactions,[290] identify which specific public key was used for that transaction.

In addition to this theoretical example by Berberich and Steiner, there have been practical situations where identification has been made through combining public keys with other information. Such instances are the voluntary release of information such as the public key to receive funds, or when additional information is gathered in accordance with regulatory requirements such as tax regulation.[291] There are other instances where law enforcements have used public keys to identify individuals. Academic research further showns that there is a possibility of identifying an individual through the combination of the public key and IP addresses.[292] These examples, both theoretical and practical testify that there are possibilities where the natural person behind a public key, is identified.

[286] C-582/14 *Breyer,* para. 39
[287] Ibid., para. 45
[288] The Working Party, Opinion 4/2007, p. 15
[289] M. Berberich and M. Steiner, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* p. 424
[290] See chapter 3.5.2
[291] M. Finck, *Blockchains and Data Protection in the European Union,* p. 24
[292] Ibid.

The question that arises is if public keys should be deemed as pseudonymous due to the fact that natural persons have been identified through public keys? It is the possibility of the natural person reasonably likely to be identified in accordance with recital 26 of the GDPR which determines the public keys status a personal, or non-personal data. This will render, as many issues regarding Blockchain, different results depending on the distinct situations due to the different structures and safety measures that might be implemented. As an example, regarding pure on-chain transactions, identifications are said to be made very difficult due to encryption. Without the off-chain transactions the encryption would make it improbable to link the public keys to individuals.[293] Such a situation might render the natural person behind the public key not reasonably likely to be identified.

As stated, one way of reducing the linkability of the public keys is to utilize the method referred to as Stealth addresses.[294] By utilizing this method, you are not able to, by looking at the public ledger, deduct which is the recipients and which is the originators new public key and as such, severely reducing the linkability of the public key.[295] There is however research which demonstrate that the use of stealth addresses are possible to be reverted in a way which would identify the different public keys as controlled by the same person. According to the research, this can be done with reasonably certainty, but not actually "proving" this as a fact.[296]

Public keys are pseudonyms in its traditional sense. They are made up of quasi-random numbers that are linked to a user through the private key, thus resembling a secretly hold pseudonym table.[297] Without additional information these public keys cannot be attributed to a certain data subject. Public keys are therefore as a rule, pseudonymous and not anonymous, and therefore constitute personal data in accordance with the GDPR.[298] This view on public keys as personal data is backed by AEPD[299] However, neither the AEPD nor Finck[300] in his study addresses how the conclusion of public key as personal data is affected if the private key is destroyed, as the CNIL recommends in order to comply with the ROR and ROE.

---

[293] M. Berberich and M. Steiner, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* p. 424, note 15
[294] Chapter 3.5.3.1
[295] M. Finck and F. Pallas, *They who must not be identified – Distinguishing Personal from Non-Personal data under the GDPR*, p. 36
[296] Ibid., p. 37
[297] Ibid., p. 35
[298] Art. 4.5 GDPR
[299] Agencia Española protección datas, *Encrytpion and Privacy V: The key as personal data*, https://www.aepd.es/en/prensa-y-comunicacion/blog/encryption-and-privacy-v-the-key-as-personal-data
[300] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?*

As previously shown, the public key is by rule considered to constitute personal data. It stands without a doubt that the deletion of the private key does not affect the existence of the public key in the Blockchain. As stated, Blockchain utilizes an asymmetric encryption which implies that without the private key, there is not a possibility for verifying or accessing the data, which is encrypted with the public key, without destroying the encryption as aforementioned. As the CNIL argues, this leaves the personal data on the Blockchain, but you are not able to prove or verify which information that has been hashed.[301] In the discussion whether public keys are personal data or not, the issue was not the possibility for others to gain access to the private key which determined it to be personal data or not. Instead, it was the possibility to combine the public key with other identifiers and thus make the natural person identifiable. The deletion of the private key would not remarkably affect this since it only affects the possibility to accessing or verify the data. This would argue for the remaining public key to continue to constitute personal data after the destruction of the private key. This is in line with the approach which is advocated by the Working Party. According to the Working Party, anonymized data are data where identification is no longer possible.[302] Furthermore, the Working Party has stated that you have to bear in mind all possible technological developments that will occur during the processing of the data.[303]

The public key will remain on the Blockchain indefinitely and as such it will be processed indefinitely. According to the Working Party, this would imply that when defining public keys as personal data or not, you would have to consider possible technological developments that could occur during this indefinite processing. It would seem unfeasible that the public key would be unaffected by these possible future technological developments. As such, it would seem impossible for the public key not to be considered as personal data. The fact that the accessibility and verifiability of the data is rendered highly improbable due to the destruction of the private key does not affect this conclusion. Furthermore, an implementation of the aforementioned stealth addresses would not affect this conclusion. The Working Party has concluded that if there is a possibility for de-identified data to be re-identified, then it cannot constitute anonymous data.

[301] CNIL, Solutions for a responsible use of the blockchain in the context of personal data, p.8
[302] The Working Party, Opinion 05/2014, p. 8
[303] Chapter 2.2.2.2.1

However, if a risk-based approach is at hand, as the one argued by Finck and Pallas[304], the outcome might differ. As stated, the possibility to access and verify what data has been encrypted through the public key is almost non-existent when the private key is destroyed. Seen in the light of what Finck and Pallas has argued, public key with a correctly destroyed private key should have a significant lower risk in affecting the rights and freedoms of the data subject. The possibility of accessing the personal information would be very low. Contrary to what the Working Party advocates, the risk-based approach allows the status of the data as personal or non-personal to change over time. This allows the issue regarding possible technological advancements over time affect the interpretation of the public key as personal or not. Thus, allowing it to be deemed as anonymous data if the risk of it affecting the rights and freedom of the data subject is low. If stealth addresses were to be implemented as well, the level of linkability would be increasingly lower and as such, rendering the risks of it affecting the individual rights and freedom, improbable. With this line of reasoning, it could be concluded that where a risk-based approach was to be adopted, and seen in the light of the context, public keys with destroyed private keys could be rendered anonymous.

4.1.4.2 Does Hash reference constitute personal data?

Hash references occur on the Blockchain in cases where off-chain data storage has been utilized. They reference to personal data that is stored off-chain in a way to circumvent the issues regarding storing personal data within the Blockchain. For this to be a viable solution, certain issues must be discussed. For instance, the question whether the hash reference itself constitutes personal data and if this status changes as the personal data, to which it refers to, has been deleted.

For something to be considered personal data, it must constitute any information that are related to an identified or identifiable natural person. Since "any information" is not constricted to a specific form, medium or sort of information, a hash reference is considered to constitute any information per the GDPR. Because of the data the hash reference is supposed to refer to, is personal data, it stands without question that the "natural person" criteria are fulfilled. However, is a hash reference related to an individual? The hash reference is "pointing" towards another dataset which is by the very definition "relating to" an individual. The Working Party has concluded in its working document regarding RFID technology that the RFID would be deemed

[304] See chapter 2.2.7

to be "relating to" if the data it linked to constitutes personal data.[305] The Working Party has continued by stating that when determining if something is "relating to" an individual, you have to consider how the data will eventually be used. If the data in question is likely to be used with the purpose to evaluate or in any way influence the status or behavior of the individual, then the data is considered to be "related to" an individual.[306] The personal data within a Blockchain are used to provide a transparent transaction history of which ownership is based. The hash references are used as a substitute to the personal data in the Blockchain and therefore it should have the same status as the data it refers to. Since the personal data is used to evaluate or influence the status or behavior of the individual the same should apply for the hash reference and as such, in accordance with the opinion of the Working Party, the hash reference should be considered to "related to".

The last point to discuss is whether the hash-reference is linked to an identified or identifiable natural person. It is shown that the off-chain data storage solution is merely a way of allowing the personal data to be altered and even erased without disrupting the Blockchain. It is also shown that the hash reference itself is merely a hash function of the encrypted data stored in the off-chain data storage.[307] Whether hashed data is considered to be personal data is subject for ongoing debate, but hashed data will more often result in pseudonymous data than anonymous.[308] According to the Working Party, the use of hash functions would reduce the linkability of the dataset and is therefore a useful security measure but not a method of anonymization.[309] Therefore, the hash reference shall be considered to be related to an "identified or identifiable" natural person. Due to what has been shown above, the hash reference shall constitute personal data.

The question that remains, is what happens when the personal data in the off-chain storage is erased. This is a question which Finck has pointed out would need regulatory guidance.[310] Per the definition of personal data, it has to relate to an individual. The hash reference in itself is not considered to be personal data, it is the data which it refers to which renders it as personal

[305] The Working Party, *RFID technology*, p. 8
[306] The Working Party, Opinion 4/2007, p. 10
[307] Chapter 3.5.2.1
[308] M. Finck, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?* p. 30
[309] Ibid.
[310] Ibid., p. 32

data. The question is if the hash reference which still remains within the Blockchain continue to constitute personal data even after the data which rendered it personal is erased?

The criteria regarding "any information" is still applicable. It is still the same data as in the case where the personal data is not erased. This since the hash itself is still intact within the Blockchain and is connected to public keys which has shown are as a rule pseudonymous and thus personal data. The criteria "identified or identifiable" and "natural person" can therefore be considered to be fulfilled. Is it still considered to be "related to" and identified or identifiable natural person? If we consider the opinion delivered by the Working Party regarding RFID technology again, it will seem as in the light of this it will not fulfill the criteria to be "related to" since there is no personal data which the hash is referring to. Thus, rendering the remaining hash as non-personal data.

As aforementioned, the purpose of having an immutable ledger in the Blockchain is to be able to provide a correct proof of ownership. It could therefore be argued that the hash which is left within the Blockchain will still have an effect of influencing or evaluating individuals. The hash could be utilized or having the purpose of proving ownership which remains unaffected after the deletion of the personal data. Thus, in the light of this it could be argued that the remaining hash would have the purpose of influencing or evaluating the behavior of an individual, thus fulfilling the criteria of being "relating to" according to the Working Party.[311]

4.2 Society, GDPR and Blockchain

As shown, the relation between GDPR and Blockchain is a difficult one. Even with new possible technological aspects and pure Blockchain-related remedies there are still a lot of question marks regarding the ROR, ROE and Blockchain. Voices have been raised regarding the interpretation of GDPR, especially focusing on the interpretation of personal data and when it is deemed to be anonymous.

Recital 26 of the GDPR establishes the approach of the GDPR regarding anonymization. The recital puts forward a risk-based judgement regarding if something is considered to be pseudonymous or anonymous data. It is important to note that the recitals of EU legislation are not binding in themselves but should serve as highly persuasive interpretations of the provision within the regulation. The Working Party established its position regarding anonymization

[311] The Working Party, Opinion 4/2007, p. 10

through their opinion on anonymization techniques of which some statements seem sympathetic to the risk-based approach.[312] The Working Party warns about the difficulties to create a truly anonymous dataset.[313] Furthermore, they recommend that to be able to assess the risk re-identification, the risk posed by a hacker with the aim of re-identify should be considered.[314] As for these two statements it would seem that the Working Party encourage a risk-based approach regarding what is supposed to be anonymous data and what is not. Despite this, the Working Party's stance in this regard are instead a zero-risk approach. They state inter alia that "anonymized data would therefore be anonymous data that previously referred to an identifiable person, but where the identification is *no longer possible*".[315] Furthermore, as previously stated, the Working Party has encouraged an interpretation of recital 26 that you should consider all possible technological developments that possibly would happen during the time of processing.[316] As technology advances, the introduction of more powerful computational powers as well as AI and other software evolves, it would render the possibility for something to be considered to be anonymized personal data to zero.

This approach which is argued by the Working Party would hinder the development of Blockchain in many aspects. When applying this zero-risk approach it would appear that even with off-chain data storage as well as deletion of private keys would fall short in terms of Blockchains compliance with the GDPR. This is problematic for the EU since one of the Commissions digital strategies is the supremacy within Blockchain technology.[317]

As aforementioned, the right of data protection is not an absolute right and thus must be balanced against the rights and freedom of others.[318] The differentiation between personal and non-personal data should be carried out through the assessment of what means that are likely to be used to identify the natural person.[319] Pursuant to this, Finck has concluded that "Where personal data never related to a natural person or that is no longer reasonably likely to be attributed to a natural person, it qualifies as "anonymous information" and eschews the Regulation's scope of application."[320] Per the wording of the recital together with the comment

[312] The Working Party, Opinion 05/2014
[313] Ibid., p. 3
[314] Ibid., p. 4
[315] Ibid, p. 8
[316] Chapter 2.2.2.2.1
[317] Chapter 3.1
[318] Art. 52, *The EU Charter of Fundamental Rights*
[319] Rec. 26 GDPR
[320] Finck, Pallas, *They who must not be identified – Distinguishing Personal from Non-Personal Data under the GDPR*, p. 36

from Finck, it would seem that the GDPR sympathizes with a risk-based approach in this aspect. Finck continues to bring up that the ECJ has ruled on the risk-based approach in the *Breyer* case and that the ECJ established this approach.[321] It is worth noting that the wording by the ECJ in the *Breyer* case does not express that a risk-based approach should be adopted literally, but Finck argues that pursuant to their reasoning, it would seem that this approach is what the ECJ has adopted.

One of the factors which constituted the need for the GDPR was the lack of the desirable harmonizing effect of the Directive. It was seen as a threat for both the personal data, but namely the growth and competitive strength of the internal market. This reason for adopting new legislation regarding data protection has been coherent since the adoption of the guidance by the OECD.[322] This has also been escribed in the GDPR through the recitals which underlines the importance of the strengthening and the convergence of the economies within the internal market.[323] This shows the other underlining reason for data protection regulation throughout the history, and furthermore it determines the other objective of the GDPR. Data protection is arguably the most important, but the economic factor seems to play a large part as well. Bearing that in mind, the width of the negative implication on the internal market by data protection regulation should affect the interpretation of the regulation. However, regarding the ROR and ROE, the ECJ has established that the interest of the individual exceeds, as a rule, the economic interest.[324]

Pursuant to the EU Blockchain strategy[325] it is imperative not to restrict the development and establishment of Blockchain technology within the EU, to be able to have a competitive internal market. If the definition of personal data and non-personal data were to be interpreted such as argued by the Working Party, it would severely restrict the development of Blockchain within the EU as Blockchains such as those addressed in this thesis would fall outside of what would be compliant with the ROR and ROE. This could have effects regarding the internal market further growth and the strength of the internal market on the international market. The fact that the ECJ has ruled that the economic interest is exceeded by the interest of the individual cannot be given any credit in this aspect. In the *Google Spain* case, it was the economic interest of the

[321] Finck, Pallas, *They who must not be identified – Distinguishing Personal from Non-Personal Data under the GDPR*, p. 14
[322] Chapter 2.1
[323] Rec. 2 GDPR
[324] Case 131/12 *Google Spain*, para. 99
[325] The European Commission: Shaping Europe's digital Future, *Blockchain Strategy*

company which, as a rule, was exceeded by the interest of the individual.[326] Not the economic interest of the internal market which in the circumstances of this thesis, should take precedence.

One key aspect of the GDPR is that it is designed to be neutral to technological advances.[327] When the technology changes, the direct need for a new regulation should not exist, it should have a stance which is unaffected by this development. This is what has been strived after in the design of the regulation. Pursuant this, the GDPR should not hold back technological development either. There has to exist some leeway to circumvent the need for an updated regulation when the technological development requires it. That the GDPR, in its effort to be technological neutral, also has to be able to be adaptable as the technological environment changes.

With this in mind, the regulation should always be focused on protecting the personal data of natural persons, but it must also be a functioning part of the society it affects.

Arguably, the interpretation regarding the concept of anonymization as advocated by the Working Party would have a negative effect on the goal of the commission regarding Blockchain. Such an interpretation would also affect the competitive advantage of the internal market since it would severely limit the legal application of Blockchain within the EU. This could result in further negative economic effect on the EU since we would be restricted from exploiting the possible competitive advantages that would lie within this line of technology. In the light of the discussion above, such an interpretation would result in the need of a new regulation to be able to remedy this negative effect of the GDPR.

Due to this, the zero-risk approach which the Working Party has advocated seems outdated. An absolutist approach which zero-risk demands is not feasible in this day and age. Thus, to be able to avoid the need for further regulation in order for the data protection regulation not to be a restriction regarding the development of Blockchain within the EU, a risk-based approach as advocated by Finck and Pallas is proposed to be implemented.

[326] Chapter 2.2.8.2.1
[327] Chapter 2.2.1.1

4.2.1 Hash reference in the light of the risk-based approach

The remaining hash in the off-chain data storage solution after deletion of the personal data, should in the light of the arguments presented above constitute anonymous data. As aforementioned, the GDPR sets to protect the personal data of the individuals and to ensure that the protection of personal data does not implicate a unproportionally negative effect on the internal market. An interpretation which would give the remaining hash-reference status as personal data after deletion cannot be considered to be proportional to the GDPR purpose. Such an interpretation would limit the possibilities of allowing the establishment and further development of Blockchain within the EU. The off-chain data storage solution provides an alternative which would allow the rectification and erasure of the personal data. A conclusion which would result in the remaining hash reference constituting personal data would severely restrict the possibilities for Blockchains to be compliant with the GDPR. This would result in a disadvantage for the internal market regarding its competitiveness. Such aspects have historically been a reason for adopting new data protection regulation.

The argument that the hash reference still fulfills the requirement "relate to", since it will still influence the behavior cannot be given any credit. The remaining hash reference might still be used to influence the behavior of a natural person which in accordance with the Working Party, would render it personal data. This factor has to be put against the risk of the hash reference to affect the rights and freedoms of the data subject per the risk-based approach. Since the hash reference is no longer referring to any personal data, it is the hash reference mere existence which will have to provide a risk against the rights and freedoms of the data subject for it to constitute personal data. Since the hash reference in itself cannot be considered to constitute personal data, the risk it poses against the rights and freedoms of the individual has to be considered insignificant. It is this thesis suggestion that the hash reference should in such cases constitute anonymous data.

4.2.2 Public key in the light of the risk-based approach

The arguments for public keys to constitute personal data even after the deletion of the private key are quite strong. The only things that change after deletion is the mere possibility of accessing and evaluating the personal data. Without the private key, this possibility is rendered improbable. The conclusion of the public key still constituting personal data is in line with what the Working Party has been advocating. The examples brought up in the previous discussion

are concerning the different situations where re-identification has been made were not based on the existence of the private key.[328] The possibilities of re-identifications are identical as before the destruction of the private key.

If such an interpretation of the public key, regardless of the private key destruction, were to be made it would entail that there is a no-existing possible for open permissionless P2P Blockchains to be established within the EU. This would result in the need for new regulation in order for the Commission to be able to act accordingly to their Blockchain strategy.

As aforementioned the GDPR does not provide a legal definition of *erasure.*[329] The alternative proposed by CNIL for Blockchain to be compatible with the ROR and ROE has the effect of making the data inaccessible. Arguably this alternative could amount to erasure since it has the same effect. As shown, this does not affect the public key which remains on the Blockchain. However, when the private key has been correctly deleted, the probability for the remaining public key to affect the rights and freedoms of the data subject to which it refers has to be considered slim. This since the only way to gain access to the encrypted information would be by destroying the encryption, something that is not probable. Furthermore, the risk-based approach admits a fluidity for the concept of anonymous data. This opens up for the possibility of something to be considered as anonymous even if there is a probability that future technological advancements would allow re-identification.

Even with the risk-based approach, a clear interpretation does not exist. It could be argued that since the data is inaccessible, the risk for the rights and freedom of the individual are supposedly low. Therefore, should the remaining public key constitute anonymous data. However, the public key is still remaining on the Blockchain and as shown above there have been instances where e.g., law enforcement has identified an individual through the use of public keys. This highlights the need for regulatory guidance in the matter. This issue could be partly solved through the usage of stealth addresses, which would elevate the effort needed for e.g., law enforcement to successfully identify the owner behind the public key.

This thesis suggests that even though there is a possibility of re-identification, that the basis of it constituting personal data or not should be heavily contextualized. Even if there have been instances where public keys have been utilized for re-identification, that if combined with e.g.,

[328] Chapter 4.1.4.1
[329] Chapter 2.2.8.2

off-chain data storage and stealth addresses, the public key should not constitute personal data. As shown, the already existing risk of being identified through the use of solely public keys are not impossible but also not probable. In itself, a public key's possible effect on the rights and freedom of the individual does not constitute such a risk that it would be impossible for it to be rendered anonymous. This aspect in a context where the private key has been deleted and other efforts for securing the rights and freedoms of the individual, e.g., off-chain data storage and stealth addresses, would render the risk of it affecting the rights and freedom of the individual, negligible. This thesis suggests therefore, that public keys, depending on the context for which they exist in, could be considered as anonymous data.

4.3 Final remarks

Shown within this thesis is the difficulties that occurs when the Blockchain that contains personal data has to be compliant with the ROR and ROE of the GDPR. Both the regulation as well as the technology are novel. The different contradictions between the GDPR and Blockchain puts new strains on the regulation which was not factored in the design of the GDPR. This is especially highlighted in the contradictions shown in this thesis while bearing in mind the Blockchain strategy of the Commission. The GDPR demands everything containing personal data to be mutable while Blockchain utilizes an immutable, append-only ledger.

It is this thesis suggestion that while utilizing off-chain storage for the personal data together with stealth addresses and allowing the destruction of the private key to sum up to erasure, the Blockchain can be compliant with the ROR and ROE. This is, as shown in the discussion above, however highly contextualized. The conclusions reached in this thesis provides an alternative for the GDPR to remain intact but without restricting the possibilities for Blockchain to be established and thrive within the EU. An interpretation without the effects proposed in this thesis will risk having severely negative effect on the possibility for the EU to achieve a competitive internal market. Such an interpretation would also directly impede the possibility for the EU to achieve supremacy in Blockchain technology. As aforementioned these are factors which has to be brought in mind when applying the GDPR.

The relation between Blockchain and GDPR is interesting since it is a clear example of the disruptiveness of innovative technology to society. As the rate of technological advances increases, so will the need of more adaptable regulation which affect these developments.

The aim of this thesis was to highlight the issues with compatibility of Blockchain and GDPR and to investigate eventual solutions to the issue. This thesis is by no means the final solution to these problems of compatibility. However, the conclusion proposed by the thesis provide one alternative solution to a very complicated problem. As shown, there is a vital need for regulatory guidance regarding the questions discussed in this thesis.

List of references

Statues, conventions and documents of the European Union

Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data No. 108,

Consolidated Version of the treaty on the Functioning of the European Union, 2012, OJ C 326/47.

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 on the free movement of such data, 1995, OJ L 281/31

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016, OJ L 119/1

The Charter of Fundamental Rights of the European Union (2012/C 326/02), 2012, OJ C 326/391

Case law

Case C-131/12, Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja Gonzaléz, (2014), ECLI:EU:C:2014:317, Court of Justice of the European Union

Case C-434/16, Peter Nowak v Data Protection Commissioner, (2017), ECLI:EU:C:2017:994, Court of Justice of the European Union

Case C-582/14, Patrick Breyer v Bundesrepublik Deutschland, (2016), ECLI:EU:C:2016:779, Court of Justice of the European Union

Case C-70/10, Scarlet Extended SA V. Société belge des auterus, compositeurs et éditeurs SCRL (SABAM), (2011), Court of Justice of the European Union

Article 29 Working Party Documents

Article 29 Data Protection Working Party, (2005), Working Document on data protection issues related to RFID technology, WP 105

Art. 29 Data Protection Working Party. (2007) Opinion 4/2007, on the concept of personal data, WP 136

Art. 29 Data Protection Working Party. (2010) Opinion 1/2010, on the concepts of "controller" and "processor", WP 169

Art. 29 Data Protection Working Party. (2014) Opinion 5/2014, on Anonymisation Techniques WP 216

Literature and articles

Bacon. J, Michels., J, Millard. C, Singh. J, *Blockchain Demystified: A technical and legal introduction to Distributed and Centralized Ledgers,* 25 Rich. J.L. & Tech ., no. 1, 2018

Bacon. J, Michels. J, Millard. C, Singh. J, *Blockchain Demystified: An introduction to blockchain technology and its legal implications,* Queen Mary University of London, School of Law Legal Studies Paper No. 268/2017

Berberich. M And Steiner. M, *Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?* 2 European Data Protection Law Review, 2016

Dresher. D *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, Main, APress, 2017

Frydlinger, D, Edvardsson. T, Olstedt Carlström. C, Beyer. S, *GDPR: Juridik, organisation och säkerhet enligt dataskyddsförordningen*, Norstedts Juridik AB, Stockholm, 2018

Finck. M, *Blockchains and Data Protection in the European Union,* European Data Protection Law Review, 2018

Finck. M, *Blockchain and the General Data Protection Regulation – Can distributed ledgers be squared with European data protection law?,* European Parliamentary Research Service, Panel for the Future of Science and Technology (STOA), 2019

Finck. M and Pallas. F, Max, *They who must not be Identified – Distinguishing Personal from Non-Personal Data under the GDPR*, Planck Institute for Innovation and Competition Research Paper Series, 2020

Jareborg, N. *Rättsdogmatik som vetenskap*, SvJT, (2004)

Kleineman. J, "Rättdogmatisk metod", In M, Nääv and MJ, Kleineman. ”Rättsdogmatisk metod”, In M, Nääv and M, Zamboni (eds.), Juridisk Metodlära, 2nd edn., Lund, Studentlitteratur AB, 2018

Narayanan, A. and Clark, J., *Bitcoin's Academic Pedigree´60 Communications of the ACM 36*

Olsen, L, *Rättsvetenskapliga perspektiv*, SvJT, (2004)

Sandgren. C, *Är. Rättsdogmatiken dogmatisk?* Tidsskrift for rettvitenskap, Tfr 118(4-5), 2005

Stalla-Bourdillon. S, and Knight. A, *Anonymous data V. personal data – a false debate: an EU perspective on anonymization, pseudonymization and personal data*, Wisconsin International Law Joural 34 (2), 2017

Voigt, P, and Bussche, A. *The EU General Data Protection Regulation (GDPR): A Practical Guide*, Springer International Publishing, Cham, 2017

Zyskind. G, Nathan. O, Pentland. A, *Decentralizing privacy: Using Blockchain to Protect Personal Data*, IEEE Security and Privacy Workshops, 2015

Other sources

Agencia Española protección datas, *Encrytpion and Privacy V: The key as personal data*, 2021, https://www.aepd.es/en/prensa-y-comunicacion/blog/encryption-and-privacy-v-the-key-as-personal-data, accessed 9th of December 2021

Buterin. V, *Privacy on the Blockchain*, 2016, https://blog.ethereum.org/2016/01/15/privacy-on-the-blockchain/, accessed on 17th December 2021.

Commision Nationale Informatique et Libertés. *Solutions for a responsible use of the blockchain in the context of personal data*, September 2018

Deloitte's 2021 Global Blockchain Survey: *A new age of digital assets*, 2021, https://www2.deloitte.com/content/dam/insights/articles/US144337_Blockchain-survey/DI_Blockchain-survey.pdf, accessed 18th December 2021.

Nakamoto. S, *Bitcoin: A Peer-To-Peer Electronic Cash System*, 2008, https://bitcoin.org/bitcoin.pdf, accessed on 14th November 2021.

Oxford Lexico, https://www.lexico.com/definition/erasure, accessed on 3rd December 2021

Seald, *Data destruction using crypto-shredding*, https://www.seald.io/blog/data-destruction-using-crypto-shredding, accessed on 9th of December 2021.

The European Business Review, *Future of Blockchain: How will it revolutionize the world in 2022 & Beyond*, 2021, https://www.europeanbusinessreview.com/future-of-blockchain-how-will-it-revolutionize-the-world-in-2022-beyond/, accessed 18th December 2021.

The European Commission: *Shaping Europe's digital Future, Blockchain Strategy*, https://digital-strategy.ec.europa.eu/en/policies/blockchain-strategy, accessed on 20th November 2021