
CYBERLAW
by CIJIC

CYBERLAW by CIJIC

EDIÇÃO N.º VIII – SETEMBRO DE 2019

REVISTA CIENTÍFICA SOBRE CYBERLAW DO CENTRO DE INVESTIGAÇÃO JURÍDICA DO CIBERESPAÇO – CIJIC – DA FACULDADE DE DIREITO DA UNIVERSIDADE DE LISBOA


CYBERLAW
by CIJIC

CYBERLAW by CIJIC

ISSN 2183-729

NOTAS DO EDITOR:

Nesta nova edição da revista do Centro de Investigação Jurídica do Ciberespaço da Faculdade de Direito da Universidade de Lisboa, dada a pertença do CIJIC ao grupo do Network of Centers (https://networkofcenters.net/center/cyberlaw-research-centre-university-lisbon-school-law-cijic), a obrigação identitária desta comunidade, persuade-nos a publicar artigos em inglês. Traremos, portanto, duas investigações em anglo-saxónico.

Na oportunidade presente da publicação desta VIII Edição e dos actos legislativos nacionais em curso, foi nossa opção trazer uma visão jurídica sobre o poder, eventualmente, manipulativo da democracia através das redes sociais.

O contexto é o da eleição presidencial de 2018, no Brasil, mas o modo como se desenvolve, desde uma engenharia social mais dissimulada a uma difusão de *fake news* ou *deep fakes*, permitem utilizar tais distorção de forma globalizada. Sendo certo que carece de maior investigação o real efeito da *realidade* das redes sociais *versus* o do "*quotidiano não digitalizado*" e o resultado concreto disto em sede de apuramento final dos resultados de eleições livres e universais, parece já possível concluir que, mesmo ante esta condicionante ainda não determinada, a realidade democrática pode, efectivamente, ser *hackeavel*.

Não obstante, por princípio, a clarificação dos conceitos de *fake news* e *deep fakes*, deveria afastar-se do radical "notícia" que lhe dá a alma. Porque uma notícia corresponde a um acto jornalístico, exercício com tutela constitucional, que conclui um dado conteúdo factual, relatando acontecimentos de interesse geral da comunidade com

o maior grau de objectividade possível. Uma notícia identifica-se pela clareza, simplicidade, exatidão, e pelo bom uso da língua em que é escrita. Compreende contraditório, ou a possibilidade deste, suporta-se em fontes credíveis. Há todo um ónus ético e deontológico que sopesa uma notícia assinada por um jornalista. Toda esta súmula é uma notícia. Comentário, mesmo televisivo, liberdade de opinião, todos os outros "*fenómenos*", não se identificam com este radical conceptual. Logo, porque continuamos a insistir em querer colar uma qualquer liberdade opinativa ao conceito de "notícia"?

Não vos soa ridículo o exercício de contínuo *fact-check* a exercícios de liberdade de opinião? Desde quando é que mentira foi legalmente proibida? Mas, pelo contrário, uma notícia que veicule um facto falacioso, de cariz subjectivo, não é fortemente sancionável? Desde logo pelos poderes de regulação, pela sindicância da própria classe, pelo público?

Será assim tão difícil perceber as diferenças?

Noutro plano, em efeméride do décimo aniversário da Lei do Cibercrime portuguesa, a Lei n.º 109/2009, de 15 de Setembro, olhamos para a perspectiva da aptidão do enquadramento legal, num contexto nada fácil, de obtenção de resultados eficazes em tempos, da acção *contra-legem versus* investigação, demasiado assíncronos. Qual a razão que explica a falta de enquadramento legal nacional para o agente (digital) encoberto, quando dezenas de outras polícias de investigação, congéneres, já o fazem?

Se há disciplina onde a soberania das fronteiras físicas acabou é no digital. Outrossim, pela fragilidade dos "muros" digitais e das deficiências do enquadramento jurídico-penal nacional, abordaremos ainda o fenómeno do *Ransomware*. Dez anos volvidos da Lei do Cibercrime, e em apologia à vanguarda em que já estivemos nos idos do início da década de 90 do século passado, impõe-se no presente, em 2019, o revisitar a especialidade da lei do cibercrime. O contexto presente de *leaks* de índole variada e processos mais ou menos mediáticos, reclamam prudência. A digitalização do Estado, por outro lado, impõem mudanças assertivas. Ademais, quer a falta da criminalização do roubo de identidade digital[1], quer a complexidade jurídico-penal do

1 Atente-se por exemplo no Considerando (14) da Directiva: "(…) *A adoção de medidas eficazes contra a usurpação de identidade e outras infrações relacionadas com a identidade constitui outro elemento importante de uma abordagem integrada contra a cibercriminalidade. A necessidade de intervenção da*

Ransomware, quer a própria transposição da Diretiva 2013/40/UE do Parlamento Europeu e do Conselho, de 12 de agosto de 2013 (esgotado o prazo de transposição no ano de 2015), quer a protecção do Estado digital (e não só) reivindicam melhores ferramentas, desde logo legais, que bem que poderiam servir de impulso necessário ao dormente legislador nacional.

Por fim, tema que não sai das agendas, o Regulamento geral de protecção de dados. Desta vez, as fricções que a ferramenta *blockchain*, cada vez mais usada no contexto das relações entre particulares e organizações, compreende face ao RGPD mas, e também, a melhor consecução dos objectivos proclamados pelo RGPD que esta ferramenta pode ajudar a alcançar.

Por fim, mas antecipando o futuro, atendendo ao propósito identitário da revista, passaremos nas próximas edições a publicar artigos de investigação dos alunos do Mestrado em Segurança da Informação e Direito do Ciberespaço, trabalhos estes desenvolvidos nas cadeiras que frequentarem.

Resta-me, neste final, agradecer a todos quantos contribuíram para mais uma edição da Revista, pelo esforço e pelo trabalho, enereçando, em nome do Centro de Investigação Jurídica do Ciberespaço – CIJIC – da Faculdade de Direito da Universidade de Lisboa, um reconhecido:

- Muito Obrigado.

Cyberlaw by CIJIC, ***Direito: a pensar tecnologicamente*.**

Boas leituras.

Lisboa, FDUL, 29 de Setembro de 2019

Nuno Teixeira Castro

União contra este tipo de comportamento criminoso poderá também ser ponderada no contexto da avaliação da necessidade de um instrumento transversal e abrangente da União."

CYBERLAW

by CIJIC

DOUTRINA

AN INTRODUCTION TO BLOCKCHAIN TECHNOLOGY FROM A LEGAL PERSPECTIVE AND ITS TENSIONS WITH THE GDPR

DIOGO GUERREIRO DUARTE [1]

1 Research Trainee at Portucalense Institute for Legal Research (Oporto, Portugal). Contacts: diogo.gue.duarte@gmail.com

ABSTRACT

In this paper, we provide a brief overview of blockchain technology from a legal perspective, and its legal tensions with the General Data Protection Regulation (GDPR). The purpose of our study is to provide a first approach to help legal professionals, researchers, and students to better understand what is the blockchain technology and how it works, and what are its implications on data protection requirements, particularly in the allocation of responsibilities and in the data subject's rights. This study primarily focusses on the decentralized and immutable features of blockchain technology and the complexities and uncertainties it creates in respect to the centralized way in which the GDPR operates. Consequently, we also present a few solutions that can be implemented into the design of blockchain-based applications to achieve some of the GDPR's objectives.

Keywords: Blockchain; Distributed Ledger Technology; Encryption; Data Protection; General Data Protection Regulation (GDPR); EU Law.

RESUMO

No presente artigo pretendemos abordar numa breve visão geral a tecnologia blockchain de uma perspectiva legal e as suas tensões jurídicas com o Regulamento Geral de Proteção de Dados (GDPR). O objetivo do nosso estudo é procurar fornecer uma primeira abordagem para ajudar profissionais do mundo jurídico, investigadores e estudantes a compreenderem melhor o que é a tecnologia e como a *blockchain* funciona e quais são suas implicações nos requisitos de proteção de dados, particularmente na alocação de responsabilidades e nos direitos do titular dos dados. Concentrar-nos-emos, principalmente, nos recursos descentralizados e imutáveis da tecnologia *blockchain* e nas complexidades e incertezas que esta cria em relação à maneira centralizada pela qual o Regulamento Geral de protecção de dados (RGPD) opera. Concomitantemente, apresentaremos ainda algumas soluções que podem ser implementadas no *design* de aplicativos baseados em *blockchain* para alcançar alguns dos objetivos do RGPD.

Palavras-chave: *Blockchain*; Tecnologia Distributed Ledger; Criptografia; Proteção de Dados; Regulamento Geral de Proteção de Dados (RGPD); Legislação da UE.

TABLE OF CONTENTS

1. INTRODUCTION

Since the General Data Protection Regulation (GDPR) came into force, numerous questions have emerged in relation to its applicability to blockchain technology. At first sight, this innovative class of new technologies seems to be unable to comply with GDPR's requirements, due to its very immutable, decentralized and transparency-based nature, thus restraining its own development and, consequently, endangering the European digital market and its technological development.[1] At the same time, being the respect for human rights one of the most important core values of the European Union,[2] the protection of natural persons with regard to the processing of personal data is expressly established in the most relevant European Union's instruments, namely under the article 8 (1) of the Charter of Fundamental Rights and the article 16 (1) of the Treaty on the Functioning of the European Union (TFEU). In this respect, the development of the internal market and the promotion of human rights need to find a fair balance, allowing the European Union to achieve its economic objectives, without sacrificing the protection of human rights, and vice-versa.[3]

As we will observe through this study, GDPR implicitly assumes that data is controlled or processed by identifiable actors, in a centralized manner. On the contrary, blockchain-based applications where designed to operate in a decentralized manner, with multiple actors and participants within a widely distributed network. The non-linear operation of blockchain-based applications, in relation to the GDPR, gives rise to several tensions, which led to the idea that theirs is an incompatible relationship. In order to detail the nature of these tensions, the main focus of this study is identifying the key features of blockchain technology that might pose a challenge to the GDPR's requirement, in particular to the data subjects' rights and freedoms. Additionally, we will explore how blockchain-based applications can be used to help achieve GDPR's objectives.

To accomplish this analysis, we will firstly provide an overview on blockchain technology, highlighting its main characteristics both from a technical and a legal perspective.

1 Article 173 (1) of Treaty on the Functioning of the European Union.
2 Article 2 of the Treaty on European Union.
3 The objectives of the internal market are described, in a general manner, in the Article 3 (3) of the Treaty on European Union.

Once we have identified its main elements of blockchain-based applications, we will examine the different types of blockchains and the roles its participants can assume in each one of them. Subsequently, we will study in further detail the existent complexities and uncertainties this technology introduces in relation to the GDPR's requirements. Finally, we explore the technological solutions that may be incorporated into blockchain-based applications to comply with the GDPR and to help achieve the GDPR's objectives. In this particular regard, we provide two solutions that, once embodied into blockchain-based applications, may allow natural and legal persons to take full advantage of the blockchain technology, aiming to achieve a fair balance between the promotion and protection of human rights and the digital market development.

A final note must be addressed to state that compliance with the GDPR is not about technology itself, but rather, it is about how technology is used.[4] Despite recognizing the need to conduct a case-by-case analysis, this study aims to provide a general overview of the application of the GDPR's requirements to the various types of blockchain-based applications.

4 *See* Ibáñez, Luis-Daniel, O'Hara, Kieron and Simperl, Elena (2018), "On Blockchains and the General Data Protection Regulation", EU Blockchain Forum and Observatory, p. 29.

2. BLOCKCHAIN

In its historical context, the starting point of the blockchain technology can be traced back to 2008, when an individual (or a group) writing under the pseudonym Satoshi Nakamoto published a whitepaper intitled '*Bitcoin: A Peer-to-Peer Electronic Cash System*'.[5] In this whitepaper, Bitcoin emerges as a '*purely peer-to-peer version of electronic cash*'[6], that uses a decentralized network to enable irreversible transactions.[7] In this new open-source online currency system, based on a peer-to-peer network, transactions can be made between the holders of the currency directly with one another, without going through any intermediaries such as financial institutions.[8] As this system does not rely on third-parties to validate, safeguard, and preserve transactions, payments can be made immediately and without the extra fees that typically increase the cost of the transactions.[9] Additionally, Bitcoin also makes non-reversible payments possible, which is a distinct feature of its technology, considering that financial institutions and other intermediaries cannot avoid mediating disputes between transacting parties, which is why non-reversible transactions are not possible within a centralized trusted entity model.[10]

Although Bitcoin was not the first manifestation of the idea of a digital currency[11], it was the first realization of this concept, and the first digital payment system that successfully allowed its participants to make direct online transactions, without placing any trust towards a central authority, and also solved the '*double-spending*' problem without relying on a trusted third party.[12] (Briefly, on blockchain cryptocurrency's applications, the digital files

5 *See generally* Chang, Henry, (2017) "*Blockchain: Disrupting Data Protection?*", Privacy Law and Business International Report, November 2017; University of Hong Kong Faculty of Law Research Paper No. 2017/041.
6 *See* Nakamoto, Satoshi, (2008) "*Bitcoin: A Peer-to-Peer Electronic Cash System*", www.bitcoin.org, p. 1.
7 *See* Maurer, B., Nelms, T. C., & Swartz, L. (2013), "*When perhaps the real problem is money itself! the practical materiality of Bitcoin*", Social Semiotics, 23(2), p. 261.
8 *Ibid*.
9 This is what Nakamoto calls '*the cost of mediation*'. *See* Nakamoto S., (2008) supra note 3, p. 1.
10 *Ibid*.
11 The idea of a cryptographic currency dates to 1983. In his article, entitled '*Blind Signatures for Untraceable Payments*', the author David Chaum proposes an untraceable-payments system based on a blind-signatures system. This system can be described as follows: "*A single note will be formed by the payer, signed by the bank, stripped by the payer, provided to the payee, and cleared by the bank*". During the 90's, the initial blind-signatures system obtained some important contributions, such as allowing payments without the bank being online at the purchase time; allowing coins to be divided into smaller unites; and improving its overall efficiency. *See* D. Chaum (1983) "*Blind Signatures for Untraceable Payments, Advances in Cryptology*", Proceedings of the Springer-Verlag Crypto'82, Vol. 3, p. 202.
12 Before the creation of Bitcoin, several companies, such as DigiCash and Peppercoin, attempted to implement electronic cash protocols. *See* Crosby, M., Nachiappan, Pattanayak, P., Verma, S., and Kalyanaraman, V., (2016), "*Blockchain Technology: Beyond Bitcoin*", Applied Innovation Review, Issue No 2, p. 2. *See also* Nian, L., Chuen, D. (2015), "*Introduction to Bitcoin*", Handbook of Digital Currency, Chapter 1, pp. 9-11.

representing a cryptocoin can be duplicated or falsified, thus being able to potentially be used more than once – this flaw is termed *double-spending*.) More broadly, the importance of Bitcoin lies on its technological structure, as it materializes the first usage of blockchain technology.[13] In fact, Bitcoin was the first ever application of blockchain technology[14] and, for that reason, the two concepts are often confused with one another, although they differ in many other aspects. For instance, Bitcoin is a cryptocurrency that was basically created to simplify and increase the speed of transaction, without relying on the intervention of a central organization or a third party. As we will observe later in this section, blockchain technology is not limited to transactions of cryptocurrencies, as it can be used to transfer any type of data or information, and can be easily adapted to different types of business and purposes.[15] Overall, the relationship between these concepts is easily understood if we consider that '*blockchain is Bitcoin's backbone technology*'.[16]

However, defining and circumscribing the concept of blockchain technology is not a straightforward task. In the absence of a unique and consensual definition in the blockchain literature, many authors tend to use different criteria to define blockchain technology.[17] For instance, some authors define blockchain by stressing out its technical characteristics and core components; others try to comprise the essential features of blockchain into a generic definition; while there some who, based on a Bitcoin blockchain generic definition, introduce some of the most recent developments of this technology.[18] In our view, in order to define and

13 See Fabiano, N. (2018), "Blockchain and Data Protection: The Value of Personal Data", J. Systemics, Cybernetics & Informatics, p. 49.

14 As Sater refers, '*Bitcoin, a cryptocurrency and a protocol, was the first decentralized and permission-less peer-to-peer payment system to implement blockcha*in'. *See* Sater, Stan (2017), "Blockchain and the European Union's General Data Protection Regulation: A Chance to Harmonize International Data Flows", Social Science Research Network, p.19. *See* also Bonneau, J., Miller, A., Clark, J., Narayanan, A., A. Kroll, J., and Felten, E., (2015) "SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies", IEEE Symposium on Security and Privacy, p. 2; and Schwerin, Simon (2018) "Blockchain and Privacy Protection in the Case of the European General Data Protection Regulation (GDPR): A Delphi Study", The Journal of The British Blockchain Association, Vol. 1, Issue 1, p. 20.

15 '[Contrary to Bitcoin] *Blockchain technology itself is non-controversial and has worked flawlessly over the years and is being successfully applied to both financial and non-financial world applications*', see Crosby, M., Nachiappan, Pattanayak, P., Verma, S., and Kalyanaraman, V., (2016), supra note 12, p. 8.

16 *Ibid.* p. 17.

17 As Schwerin identifies, the blockchain technology is currently under development, which makes it difficult to establish a precise and clear definition of its concept. In the author's view, it is possible to decompose the definition of blockchain into three layers: 1) the datalogical layer, which refers to the cryptographic functions that are used to store all transactions; 2) the infological layer, which perspectives the blockchain definition as a series of inputs and outputs between accounts that are stored in a public ledger; and 3) the essential layer, which sees transactions as commitments and economic events. *See* Schwerin, S., (2018), supra note 14, p. 21.

18 For a technical definition of blockchain and its core components, *see* Cate, Fred H.; Kuner, Christopher; Lynskey, Orla; Millard, Christopher; Ni Loideain, Nora; and Svantesson, Dan Jerker B., (2018), "*Blockchain versus Data Protection*", International Data Privacy Law, Volume 8, Issue 2, p. 103. For a broader definition of blockchain, *see* Ibáñez, Luis-Daniel, O'Hara, Kieron and Simperl, Elena (2018), supra note 4, p. 1. For a Bitcoin blockchain based definition and its subsequent developments, *see* Wright, Aaron and De Filippi, Primavera,

explain what blockchain is, we must begin by acknowledging that there is not one single blockchain technology, but, on the contrary, there is an entire class of technologies that present different technical and governance structures.[19] Thus, any attempt to define the concept of blockchain, if it goes beyond the core components common to all varieties of this technology, will fail to recognize the existence of other blockchain types. As a mere example, the typical definition of blockchain includes a reference to the resource-intensive consensus mechanism, which is used by miners to validate pending datasets and form news blocks on the chain.[20] While this feature is common among DTLs (distributed ledgers technology), the same cannot be said about the centralized trusted third-party models like private blockchain-based applications in which there is only a single entity that manages the entire blockchain. In this context, a more rigorous and realistic approach to the concept of blockchain technology requires, in the first place, the consideration of its core components. Only then it is possible to introduce and analyze the different technical and governance structures blockchain can assume, which are crucial to measure the different impact those structures have on data protection[21].

2.1 Core components of blockchain technology

Taking into account the aforementioned difficulties of producing a unanimous definition, blockchain can be generally described as a specific type of database that '*uses certain cryptographic functions* [mathematical functions/algorithms used in cryptography, *i.e.* the study and construction of protocols that prevent third parties from accessing private communications and transactions] *to achieve the requirements of data integrity and identity*

(2015) "Decentralized Blockchain Technology and the Rise of Lex Cryptographia", Social Science Research Network, pp. 1-4. Available at SSRN: https://ssrn.com/abstract=2580664

19 *See* Finck, Michèle (2019), "*Blockchain and the General Data Protection Regulation: Can distributed ledgers be squared with European data protection law?*", Panel for the Future of Science and Technology, European Parliamentary Research Service, Scientific Foresight Unit, PE 634.445, p. 3.

20 As we will see in the sections below, miners or validating nodes are one of the three participants on blockchain, which are responsible for assembling datasets into blocks and broadcast those blocks to other nodes across the peer-to-peer network, in order to validate the new block and add it on the chain.

21 In this issue, we follow the approach taken by Jean Bacon et al., who expressly acknowledge that this approach is more useful as "*(…) a lot of existing material assumes that readers are familiar with the underpinning technologies. Further, some sources fail to distinguish between the core components of blockchain and the various ways in which the technology could be applied*". *See* Bacon, Jean and Michels, Johan David and Millard, Christopher and Singh, Jatinder, (2017) "Blockchain Demystified", Queen Mary School of Law Legal Studies Research Paper No. 268/2017, p. 3.

authentication'.[22] These two core components of blockchain technology – data integrity and identity authentication – allow it to create a persistent and tamper-evident record of the dataset and authenticate the parties associated with it.[23] In this section, in order to explain how blockchain works and what are its core components in a simple manner, we will use the train's metaphor, in which the train represents the whole blockchain, each carriage characterizes a block of the chain, and the passengers symbolize single data items.

2.1.1. Data Integrity

The early applications of blockchain, such Bitcoin, were created to operate in a trustless environment, where the blockchain should not be managed by any central party, but instead stored in a distributed manner across the peer-to-peer network, in which each node holds an updated copy of the ledger of transactions.[24] In the same way the distributed peer-to-peer network is essential to overpass the inexistence of a central entity, the cryptographic hash functions are essential to safeguard the integrity of the transactions.[25] In practice, hash functions not only create a tamper-evident record of the transactions, but also guarantees that they are "*computationally impractical to reverse*".[26] Bitcoin and other blockchain applications use hash functions to generate a unique hash value to the input data item, which consists of a string of digits with a fixed length.[27] The hash value is used to prove the integrity of a data item, in that any change to the original data item will generate a different and unrelated hash value that allows blockchain's participants to detect if any attempt to tamper the data has occurred. For this particular reason, it is commonly said that the hash value of a data item works as its fingerprint, in that this value is unique.[28] Besides this particular characteristic, the hash functions are irreversible, in the sense that is not possible to use the hash value to recreate the original input of a particular data item. Revisiting the train's metaphor, the hash functions

22 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 4.
23 In order to include all the possible types of data blockchain can incorporate, the term 'dataset' is used in this study in a broad manner. Regarding the early applications of blockchain, such as cryptocurrencies, this term refers to the ledger of transactions. However, as other applications of blockchain are being currently explored, the term dataset can also include different types of data, such as land registers.
24 *See* Nakamoto, Satoshi, (2008), supra note 3, p. 1.
25 Bitcoin uses SHA-256. This cryptographic hash function requires validating nodes (or miners) to solve a cryptographic puzzle, in which they need to find a block, whose SHA-256 hash is less than a target value. In the Bitcoin context, the miners try random nonces (an arbitrary number that can be used just once in a cryptographic communication) until they find a solution, that is then broadcasted to the entire network in order to be confirmed by the other nodes. For more details, *see* Bonneau, J., Miller, A., Clark, J., Narayanan, A., A. Kroll, J., and Felten, E., (2015) supra note 14, pp. 2-5.
26 *See* Nakamoto, Satoshi, (2008), supra note 3, p. 2.
27 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 6.
28 *See* Wright, Aaron and De Filippi, Primavera, (2015), supra note 18, p. 7.

works, in this specific context, as a way to verify that passengers have a valid ticket to enter the train and that the ticket was not tampered or falsified in any way, as it is unique and only can be used by those particular passengers.

Beyond single transactions and data items, the hash functions also play an important role in making large data structures, which contain multiple transactions or data items, tamper-evident, by using hash pointers. On blockchain, those structures are commonly known as blocks, and each block contains a record of numerous individual transactions or other data items. In order to prove the integrity of the blocks, including its content and sequence, hash pointers link the blocks together, by putting into a hash function the combination of the data of each block with the hash value of the previous block. This creates a block's hash value that will be included in the next block alongside with a list of transactions or datasets and other metadata.[29] The result is a tamper-evident chain of blocks, in which any attempt to modify a block's content, will immediately break the link between blocks, allowing any fraudulent interfering to be easily spotted.[30]

In our train's metaphor, in which a carriage represents a block of transactions or other data items, the function of the 'hash pointers' is to link all the individual carriages that form the train in their proper order. In order to ensure the integrity of the train, each carriage includes a number that represents the previous one, which is generated through the combination of the carriage number and the passengers' tickets numbers. In case a modification occurs, whether regarding to passengers' tickets or to the carriage itself, the link with the other carriages will break automatically, showing that something wrong happened with any particular carriage or a passenger's ticket.

Early blockchain applications were conceived to be practically immutable and irreversible, recording and linking all the transactions into a chain of blocks. From an early stage, a concern about storage space was emerged, since the more transactions occur, the more the database grows.[31] The use of a Merkle tree provided a solution to both storage space and data verification[32]. In general, a Merkle tree is a hash-based data structure that contains the

29 As Jean Bacon *et al.* correctly identify, a block consists of two main parts: a '*block body*', that contains a list of all transactions that a block holds; and a '*block header*', that is composed by the hash of the previous block and some metadata, such as a timestamp. *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, pp. 7-8. *See* also Bonneau, J., Miller, A., Clark, J., Narayanan, A., A. Kroll, J., and Felten, E., (2015), supra note 14, pp. 4-5.

30 However, this only mitigates the risk of data being modified by an attacker, as it is virtually possible to re-hash the blocks and reconstruct the chain if the attacker holds the majority of the computational power on the network.

31 *See* Nakamoto, Satoshi, (2008), supra note 3, p. 4.

32 *Ibid*.

combination of the hash values of the individual transactions. In practice, the hash values of individual transactions are paired and put into a hash function in order to generate new hash values. This process is successively repeated until the last hash value – also known as the Merkle root – is found.[33] Each block contains a Merkle root, which represents a summary of all transactions a block holds. Besides requiring less space to store data and using fewer resources, the Merkle tree system makes it easier to verify the integrity of transactions and to check if a transaction has been included in a block without having to download the entire ledger of transactions.[34]

2.1.2. Identity Authentication

Presently, a large number of transactions related to the most diverse economic activities are still executed using financial intermediaries, such a financial institution or a bank. In this context, one of the main duties of a financial institution is to correctly identify the parties involved in a transaction and to ensure the content of the transaction is accurate. As mention above, early applications of blockchain, such as Bitcoin, were designed to operate in a trustless environment, *i.e.* without the intervention of a trusted third-party. However, blockchain still needs to identify and authenticate the parties involved in any transaction, before storing it into a block. To achieve this purpose, blockchain technology relies on a security method known as public key infrastructure (PKI).[35] This security method is used to implement strong authentication by generating a key pair containing a public and a private key, a signing algorithm, and a validation function that checks the digital signatures' validity.[36]

---

33 In order to observe how a Merkle root is obtained, let us consider the following example. Imagine that a block holds eight transactions (Tx), in which eight persons sent a certain amount of Bitcoin to other eight persons. As we mentioned above, each transaction (Tx1, Tx2, …, Tx8) is put into a hash function generating a unique hash value (h1, h2, …, h8). These hash values are then paired and put into a new hash function: (h1+h2) = h12; (h3+h4) = h34; (h5+h6) = h56; and (h7+h8) = h78. As a result, we have compressed the eight transactions into four hash values. However, to achieve a final hash value – a Merkle root – the process must continue. Thus, the four hash values are paired and put into a new hash function: (h12+h34) = h1234; (h56+h78) = h5678. The process repeats itself one last time, resulting in a Merkle root: (h1234+h5678) = h12345678. This hash value (h12345678) is then added to the block header. In case a single transaction is modified or tampered in any way, this will generate a completely different hash value, modifying, by consequence, the Merkle root. In this way, it becomes easier to detect any change in the block. For a visual explanation of this concept, *see* Bashir, I., (2017), "*Mastering blockchain*", Packt Publishing Ltd., pp. 174-177.

34 Although storage space is not considered to be a problem on blockchain, as the Moore's Law predicts that the computer capacity will be enough to store all the data inside blockchain, the use of the Merkle tree is undoubtedly a more manageable way to process large amounts of data. *See* Nakamoto, Satoshi, (2008), supra note 3, p. 4.

35 *See* Nian, L., Chuen, D. (2015), supra note 12, pp. 15-17.

36 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 9.

As all the transaction records on blockchain are signed before being included in blocks, the PKI is mainly used to establish the digital identity of a user and to create digital signatures. The public and private key can be used to encrypt and decrypt data that has been respectively encrypted or decrypted using one of these keys. Thus, to prove its identity or to sign a transaction or any other data item, a user can encrypt data using her private key and provide the associated party with the public key. If the associated party can successfully decrypt the data using the public key provided by the user, she can be confident that the transaction, or any other data item, originated from that particular user.[37]

The train's metaphor we used to explain the data integrity component, when applied to the identity authentication component, works in the following way: imagine it is only possible for a passenger to enter the carriage with an coded ticket that must be acquired on a specific platform. When creating a profile on that platform, the passenger receives a private password to login and a public password that she must give alongside the ticket before entering the carriage. An officer then uses the public password to confirm that the ticket belongs to that passenger. In case the public password provided by the passenger allows the officer to successfully scan the train ticket, her entity is proved, and she is accepted on the carriage.

2.2 Types of blockchain

On its historical context, cryptocurrencies were the first application of blockchain technology, which came into existence to surpass the '*inherent weaknesses of the trust based model*'[38], namely the fraud percentage that is accepted as unavoidable on such a model and, more specifically, the transactions' time-length and costs that costumers support when using third parties to process electronic payments. Consequently, the early applications of blockchain technology were designed to operate without a trusted third party. However, an important question arises from this paradigm: who controls the blockchain?

The question can be divided into the following two questions: who can store copies of the blockchain, and who can propose new blocks to be added to the blockchain?[39] As different

37 Unless the private key has been compromised by any attack on the user account or computer devices. *Ibid*.
38 *See* Nakamoto, Satoshi, (2008), supra note 3, p. 1.
39 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 10.

answers can be given, the design of the platforms using blockchain technology can vary significatively, creating different types of blockchain applications. Although the blockchain's design can present a wide spectrum,[40] to make the concept easier to understand blockchain types are usually segmented into private or public, and permissionless or permissioned databases.[41]

The criterion to differentiate private or public blockchains can be found by observing the way participants join the network. In this respect, while public blockchains are open to any person or entities that desire to join the peer-to-peer network, on another hand, private blockchains only allow pre-selected participants to join their peer-to-peer network. The pre-selected criterion is also used to differentiate permissionless and permissioned blockchains. In the first, any person or entity can participate in the consensus mechanism, having the possibility to add new blocks into the chain. On permissioned blockchains, only the pre-selected entities are authorized to add new blocks into the chain. [42]

For the purposes of our study, our analysis will focus next on the structure of both public, private and consortium blockchains, as the role of its participants are crucial to consider the impact these blockchain types have on data protection.

2.2.1 Public and permissionless blockchains

Cryptocurrencies are perhaps one the most widely known application of blockchain technology. As Jean Bacon *et al.* (2017) affirm, Bitcoin *'shaped the public perception of what a blockchain is*'.[43] Indeed, Bitcoin has an enormous importance to the development of blockchain technology, not only because it was the first digital currency to be successfully implemented, but more importantly, it allowed direct transactions to be made without the intervention of a trusted third party.[44] In order to operate in a trustless environment, Bitcoin's design relies in the combination of three main components: a decentralized peer-to-peer network (P2P network); a consensus mechanism; and a series of cryptographic functions.[45]

40 In its book, the author Irman Bashir provides a list of different blockchain types, which includes public; private; and semi-private blockchain; sidechains; permissioned ledger; shared ledger; fully private and proprietary blockchains; tokenized blockchains; and tokenless blockchains. *See* Bashir, I., (2017), supra note 33, pp. 32-34.
41 *See* Schwerin, Simon, (2018), supra note 14, p. 25.
42 *Ibid*.
43 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 5.
44 *See* Schwerin, Simon, (2018), supra note 14, p. 20.
45 *See* Schwerin, Simon, (2018 supra note 14, p. 22. *See* also Bonneau, J., Miller, A., Clark, J., Narayanan, A., A. Kroll, J., and Felten, E., (2015) supra note 14, pp. 2-9.

In the beginning of this section (*2.1. Core components of blockchain technology*), we analyzed how cryptographic functions work in blockchain environment. As this core component is common to the different types of blockchain, we are going to examine the other core components of public and permissionless blockchains, using Bitcoin as example.

Although the P2P network is often described as the least innovative of the three main components of Bitcoin, using our train's metaphor, this component is the engine of blockchain.[46] There are three type of participants in Bitcoin's blockchain: user, nodes and miners. Each one of these participants has a different role on the P2P network. *Users* are the persons or the entities that use Bitcoin to make transactions, *i.e.* to buy or sell bitcoins. On the user's level, Bitcoin is open and permissionless, which means that anyone can participate by simply buying a bitcoin hardware wallet, running an open source code on the computer, or using online software wallet services.[47] The *nodes*' role is of fundamental importance on Bitcoin's platforms, as they not only accept and validate transactions broadcast by the miners, but they also discover and maintain connections with other nodes to whom they send an update copy of the ledger.[48] On the nodes' level, Bitcoin is also open and permissionless, as anyone can download and run the Bitcoin's appropriate software and start storing the blockchain archive into the computer. Finally, the *miners* are the Bitcoin's participants, who assemble transactions into blocks and broadcast those blocks to the entire P2P network, according with a consensus mechanism that we will analyze next. Regarding the mining process, an interesting feature of Bitcoin architecture is that it incentivizes miners to perform the task of adding new blocks to the blockchain through an economic reward, which is also the way Bitcoin puts new coins into circulation.[49] On the miner's level, Bitcoin platform is, once again, open and permissionless, which means that anyone can be a miner.

As the open and permissionless types of blockchain do not rely on a single centralized party, the P2P network is crucial to maintain the integrity of the ledger, in the sense that even if any interference with the ledger occurs, the rest of the P2P network still has a valid copy of

---

46 *Ibid*.

47 Regardless of the way the user has chosen to join the P2P network, a public-private key pair is generated, allowing her to start using Bitcoin's platform. A user can start trading by either receiving bitcoins from another user or buying bitcoins from online exchanges. *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 11.

48 See Bonneau, J., Miller, A., Clark, J., Narayanan, A., A. Kroll, J., and Felten, E., (2015), supra note 14, p. 5, ('By default, each node attempts to make eight outgoing connections and is prepared to receive up to 125 incoming connections').

49 The miners can also be rewarded through transactions fees, which are normally used by the users to incentive miners to prioritize their transactions. In practice, a transaction fee is the difference between the input and output values that users allow miners to retain. *See* Nakamoto, Satoshi, (2008), supra note 3, p. 4.

the ledger, which will be used by the majority of the nodes and miners to create, validate and add new transactions and blocks into the chain. Because the P2P network operates in a decentralized manner, it increases the resilience of the blockchain platform, as there is no single point of failure that can be targeted with a denial of service attack.[50] However, to operate properly, namely to allow new blocks to be added to the blockchain, it is required that all the nodes and miners in the network hold an updated and synchronized copy of the ledger.[51] To this end, open and permissionless blockchain such as Bitcoin have implemented a consensus mechanism.

The Proof-of-Work (PoW) is one of the most well-known and the most used consensus mechanisms on public and permissionless blockchains. PoW is a protocol that is used to validate the data (or transactions) and form new blocks on the chain[52]. One key feature of the PoW is its asymmetry, as the work is difficult to produce, since it requires increasing amounts of computational power to decipher the cryptographic puzzle specifically created to be solved by the means of brute force calculation,[53] but is easy to be verified by all the other nodes, who can create consensus on the solution broadcasted to the whole P2P network by the first node that solved that particular cryptographic puzzle. Once the solution has been confirmed by all the other nodes and consensus has been achieved, the new block can then be appended to the longest chain.

Although the PoW is the most common consensus mechanism, there are many other types of mechanisms, such as the Proof of Stake (PoS), Byzantine fault-tolerant variants (BFT), Proof of Elapsed Time (PoET), and Algorand.[54] Apart from the specific particularities of each one of them, the consensus mechanisms are used to serve two different purposes simultaneously. On one hand, by checking the current state of the blockchain on a regular basis, the mechanisms perform a process that aims to mitigate the creation of forks into the blockchain

50 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, pp. 12-13.
51 *See* Lyons, T., Courcelas, L., and Timsit, K. (2018), "Blockchain and the GDPR", European Union Blockchain Observatory and Forum, p. 14.
52 The SHA-56 is one of the most used proof-of-work schemes and it was introduced by Bitcoin. This cryptographic hash function requires validating nodes (or miners) to solve a cryptographic puzzle, in which they need to find a block, whose SHA-256 hash is less than a target value. In the Bitcoin context, the miners try random nonces (an arbitrary number that can be used just once in a cryptographic communication) until they find a solution, that is then broadcasted to the entire network in order to be confirmed by the other nodes. For more details, *see* Bonneau, J., Miller, A., Clark, J., Narayanan, A., A. Kroll, J., and Felten, E., (2015), supra note 14, pp. 2-5.
53 *See* also Ibáñez, Luis-Daniel; O'Hara, Kieron; and Simperl, Elena (2018), supra note 4, pp. 2-3.
54 *See* Truong, N., Sun, K., Lee, G., Guo, Y. (2019) "GDPR-Compliant Personal Data Management: A Blockchain-based Solution", IEEE transaction on information forensics and security, p. 2.

and their frequency. This process is frequently referred to as the '*fork choice rule process*'.[55] On the other hand, the consensus mechanisms are used to ensure the majority of the nodes on the network agree on the legitimacy and validity of the transactions that a proposed new block contains, avoiding the malicious nodes to broadcast their own blocks. In general, the consensus mechanism allows nodes to audit the entire blockchain by constantly checking all the transactions, which significantly reduces the risk of various attacks.[56]

Once added to the blockchain, each block is computationally impractical to modify, which means transactions are recorded into blockchain on a permanent basis. In order to successfully modify a block, a validating node would need the majority of the computational power within the P2P network to do it.[57] Even if a node could successfully modify a dataset on a block, it would need to re-hash the subsequent blocks in the chain, since any modification to the dataset would automatically break the blockchain.[58] Additionally, and since the chain with the most combined computational difficulty is considered the valid one, an attacker would needed to control the addition of new blocks and, thus, use the PoW much faster than the rest of the P2P network. For these reasons, public and permissionless blockchains are considered to have a strong security feature and a *'51% attack*' is unlikely to happen, although the '*mining pools*' (*i.e.* a group of two or more miners that work together) could be virtually able to concentrate 51% of the computational power of the P2P network. [59]

2.2.2 Private and permissioned blockchains

As we mention above, the public and permissionless blockchains were designed to operate in a trustless environment, where anyone can participate either by proposing, verifying or adding new data to the blockchain. Although these features fit the purposes of early

55 From time to time, two blocks can be created simultaneously, generating a fork on the blockchain. During a fork, one of the blockchain branches will be discarded since the validating nodes (or miners, in the case of Bitcoin) will converge on the other branch. During the time a fork subsists, the blocks of both branches will be apparently included in the longest chain. The '*length*' of the entire blockchain refers not to the one with the most blocks, but to the chain that has the most combined computational difficulty. This prevents some node from forking the chain and creating many low-difficulty blocks, which otherwise would be accepted by the network as the longest chain.
56 *See also* Nian, L., Chuen, D. (2015), supra note 12, pp. 22-25.
57 Since the attacker must have most of the computational power of the entire P2P network, the attack to a public and permissionless blockchain that uses the PoW became known as the '*51% attack*'. For more details on this kind of attacks and their feasibility, *see* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, pp. 17-18.
58 Nakamoto S., (2008), supra note 3, p. 1 ('The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work').
59 S*ee* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, pp. 17-18.

applications, such as Bitcoin and Ethereum, they are not suitable for many other economic activities and industries, where efficiency, confidentiality and control over the blockchain are required.[60] In such cases, and whereas a certain level of trust among the participants can be found, blockchain-based platforms do not need to operate in trustless environments and, thus, it is possible to avoid the use of costly consensus mechanisms by relying on trusted intermediaries, such as a single trusted third party or a defined number of nodes.[61]

As private and permissioned blockchains are fully controlled by a single third party (or by a group of nodes that come from a single party), they are often regarded as being completely centralized and closed. Indeed, although private and permissioned blockchain can be designed as open at the user level, meaning that anyone can propose new data to be added to the database, only the trusted third party can perform the role of nodes and miners, *i.e.* to store the copies of the database, and propose and add new blocks to the chain.[62]

As the private and permissioned blockchains allow a single entity to have a *de facto* control over the entire blockchain, these platforms have been explored for the use of both financial and non-financial actors. [63] During the last years, an increasing number of governments around the world have been engaged with blockchain technology, and some of its key uses across the public sector often includes: identity management (proof of identity); government records, which comprises personal records, land registration, and corporate registration; government activities such as electronic voting and tax records; and other similar health and social services.[64] Although private blockchain applications are more efficient, as they operate with a single validator, it must be observed that, due its limited number of participants, these types of blockchains can impose a higher risk to the integrity of data items when compared with public blockchains.[65]

---

60 *See* Bashir, I., (2017), supra note 33, p. 632.

61 S*ee* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p.19.

62 As an example, considering the use of blockchain technology for land registry purposes, the private and permissioned blockchain models allow any natural or legal person to propose the registry of their land on the database, but only a trusted third party, such as a government agency, can store the registry and add new data to it. In this sense, this trusted third party acts simultaneously as a node and a miner, having a *de facto* power and control over the blockchain.

63 *See* Crosby, M., Nachiappan, Pattanayak, P., Verma, S., and Kalyanaraman, V., (2016), supra note 12, pp. 13-14.

64 *See* Woods, Jorden (2019), "*Blockchain Revolution in the Power Sector*", https://www.blockchainbeach.com/blockchain-revolution-in-the-power-sector-part-1/ [accessed 19 August 2019]. *See also*, *See* Sater, Stan (2017), supra note 14, p. 38.

65 *See* Zheng, Z., Xie, S., Dai, H.N., Chen, X. and Wang, H., (2017), "*An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends*". *In* 2017 IEEE International Congress on Big Data (BigData Congress), p. 559.

2.2.3 Consortium blockchains

The consortium blockchains also operate within an environment where a certain level of trust among the participants can be found, but contrary to the private blockchains, instead of relying on a single trusted third entity, the consortium blockchain applications are structured around a defined number of nodes, called '*trusted nodes*'.[66] The consortium blockchains restrict control over the blockchain by giving the possibility to store a copy of the database and to add new blocks on the chain to only a small group of trusted nodes.[67] This signifies that, on the nodes and the miners' level, the consortium blockchains are commonly considered as closed and permissioned platforms. Even at the user's level, it is frequently observed that only authorized parties can join the network. That is case of the R3 Corda, one of the best-known examples of a consortium blockchain platform, which enables a consortium of more than 300 participants of the financial industry '*to transact directly and in strict privacy using smart contracts, reducing transaction and record-keeping costs and streamlining business operations*'.[68] Hyperledger, an open source and modular platform that allows customization and utilizes permissioned blockchain technology to build private business networks, is also another great example of a consortium blockchain platform.[69]

Besides limiting the participation in the network only to a small number of trusted nodes, another characteristic of consortium blockchains, which distinguishes them from both public and private blockchains, lies on its consensus mechanism. As we analyze above, public blockchains uses the PoW protocol to achieve consensus among the nodes and, on private blockchains, there is only a single entity that controls the entire blockchain. Differently, on consortium blockchain, only a small number of nodes participate in the consensus mechanism, which means that an absolute consensus must be achieved in order to add new blocks to the chain. The public blockchain consensus protocols are not adequate for consortium models, as they are costly and require higher amounts of energy and computational power. Thus, instead of relying on asynchronous consensus protocols, consortium blockchains applications typically rely on a traditional and synchronous consensus mechanism.[70] Similar to the private

66 Sometimes, consortium blockchains are designated as '*combined blockchains*'. *See* Fabiano, N. (2018), supra note 13, p. 49.
67 *See* Zheng, Z., Xie, S., Dai, H.N., Chen, X. and Wang, H., (2017 supra note 65, p. 559.
68 *See* https://www.r3.com/platform/ [accessed 19 August 2019].
69 *See* https://www.hyperledger.org/about [accessed 19 August 2019].
70 Synchronous consensus mechanisms are consensus protocols that keep all the nodes in the network synchronized with each other by imposing two requirements: firstly, all the nodes must have an updated copy of the database before moving to the next block; and secondly, before adding a new block to the chain all nodes must achieve consensus. By contrast, asynchronous consensus mechanisms, as the name indicates, do not synchronize

blockchains, the consortium blockchains also present a higher risk to the integrity of data and to the database itself, as the nodes consortium represent single points of failure, which makes them more vulnerable to denial of services attacks and other hacking attacks.[71]

2.3 Blockchain's control and governance

Who controls the blockchain platform? Who can change the platform's design and to what extent? These questions are important not only to conclude our introduction to the blockchain technology, but they are also crucial, as we will see in the following sections, to determine who can be regarded as a controller and/or a processor in the GDPR's perspective.

Each blockchain platform has its own governance rules and its own design and structure. Developers are responsible for producing the software which is used by nodes and miners to support the blockchain.[72] However, the way developers change the platforms' design by introducing changes to the application software can differ substantially, taking into count how platforms were designed in the fist place. For instance, some public blockchain platforms, such as Bitcoin and Ethereum, were developed using an open-source code, which could be used by other developers, rather than the core developers, to write a new version of the software and make it available for the P2P network participants. Typically, apart from the bug fixes, the changes introduced to the blockchain software are meant to achieve other functionalities or to modify the software's capability. Once introduced to the P2P network, nodes and miners can decide which software version they want to run. In case a developer successfully convinced miners and nodes to adopt a new version of the software, a hard fork on the blockchain will be created, originating two different blockchains, in which new blocks will be added subsequently. In the Bitcoin context, when a new version of the software is adopted by nodes

---

with the other nodes, which means that nodes in a network can move to the next block without waiting for an update copy of the database.

71 As the authors Jean Bacon *et al*. state, considering the characteristics of consortium blockchains platforms, it is accurate to see these platforms as a '*permissioned, 'narrowly distributed' platform with a 'shared' (as opposed to distributed) ledger*'. *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 21.

72 In the Bitcoin context, it is possible to split the developers' group in two other groups: the developers' group, who can propose new technical improvements; and the core developers' group, who has the power to introduce changes to the Bitcoin Core software. *Ibid*. *See* also Vetter, Greg R., (2004) ''Infectious' Open Source Software: Spreading Incentives or Promoting Resistance?', Rutgers Law Journal, University of Houston Law Center, No. 2004-A-11, pp. 77-88.

and miners a hard fork is created, giving rise to new blockchains, one that continues to track bitcoins, and a new one that now tracks a new crypto coin.[73]

On the contrary, on private and consortium blockchains, new versions of the software that supports the blockchain can be subjected to contractual provisions that were negotiated between the parties involved in the creation and development of a particular blockchain platform. In such cases, it can be argued that the developers' role is limited to the performance of contractual obligations, as they do not typically have the power or the means to change the software version by their own initiative.

Finally, there is a fifth intervenient group on the blockchain environment: the service providers. Normally, service providers intervein on blockchain platforms either by offering services related to online wallets or by offering '*Blockchain-as-a-Service*'.

Online wallets are digital wallets that work as an interface to a blockchain system, allowing users to manage crypto coins or other digital assets. When using an online wallet, the users are provided with a wallet ID, which is a unique identifier such as a bank account number. However, this wallet ID is completely different from the private and public key pair, which is generated by those services on the users' behalf.[74]

On other hand, '*Blockchain-as-a-Service*' is a service that allows the customer to leverage cloud-based solutions to create their own blockchain applications. The service providers' offer includes a wide range of task and activities that can include the management of the platform, the hosting of a certain number of nodes, or even the management of identity authentication. Since the service providers have a direct involvement on the blockchain platform creation and control, this might raise some questions related to the degree of power and the control the service providers have over the blockchain. As we will analyze below, these questions are important to determine the nature of service providers in the context of the GDPR, namely, to assess whether the service providers can determine the purposes and the means of the processing of personal data.

---

73 Currently, there are 105 Bitcoin forks of which 74 are considered active projects, and 34 are regarded as historic projects. For an overview of all Bitcoin forks, *see* https://forkdrop.io/how-many-bitcoin-forks-are-there [accessed 19 August 2019].

74 When using an online wallet, users must bear in mind that relying on an intermediary could jeopardize the security of their assets, as those service providers are not immune to cyberattacks and other risks that could result in the lost of the users' private and public key pair. Actually, phishing attacks are often directed against online wallets. *See* Khatri, Yogita, (December 28, 2018) "*Electrum Wallet Attack May Have Stolen As Much as 245 Bitcoin*" *in* https://www.coindesk.com/electrum-wallet-attack-may-have-stolen-as-much-as-245-bitcoin [accessed 19 August 2019].

3 IDENTITY OF THE BLOCKCHAIN PARTICIPANTS

As previously stated, the taxonomy (*i.e.* the structure, organic, etc.) of the blockchain applications has a different impact on the powers, rights, permissions and restrictions of the participants of a particular platform. Typically, any participant, or even the public, can consult the entire blockchain archive of a public and permissionless blockchain application. By contrast, on consortium blockchain applications, the archives' reading permissions can be limited to a certain number of participants, while on private and permissioned blockchain applications, the access to the blockchain archives can be denied or limited to a few blocks or certain data entries.[75]

This immediately raises two mains questions: can participants of the blockchain platforms be identified? If so, can any other participant access their data and transactions' history?

As stated above, blockchain applications use a PKI to authenticate the identity of their participants. In general, the public and private key do not reveal the participants real-world identify. Early blockchain applications, such as Bitcoin, took these concerns into account and, as Nakamoto (2009) explains: '*privacy can still be maintained by breaking the flow of information in another place: by keeping public keys anonymous. The public can see that someone is sending an amount to someone else, but without information linking the transaction to anyone*'.[76] Additionally, users can generate a new private and public key pair for each new transaction.[77] This level of pseudonymization ensures that even on public blockchains, where anyone can consult the blockchain archive, no one will be able to determine the real-world identity of the parties involved in a particular transaction.

However, the users' identity can be exposed on a voluntary basis, if the users decide to reveal their real-world identity, or on an involuntary basis, as it is the case of malicious attacks on online wallets, in which the attacker has obtained access to user information. In the same way, the real-world identity of a user can be indirectly revealed by linking different data elements. For instance, if a user uses bitcoin as a method of payment to buy goods or services, the other party might need the customer's name, email address, postal address, and other

75 *See* Zheng, Z., Xie, S., Dai, H.N., Chen, X. and Wang, H., (2017 supra note 65, p. 559
76 *See* Nakamoto, Satoshi, (2008), supra note 3, p. 9.
77 *Ibid*.

personal information that can lead to the identification of the user.[78] The IP addresses can also be used as a way to determine the users' identity by linking the users' private and public key pair to the locality from which the transaction was generated.[79]

Although the PKI ensures a certain level of protection to the users' identity, once their real-world identity has been revealed, anyone can access the entire transaction history associated with that user, especially in cases where the private and public key pair has not been changed. Although private and consortium blockchains normally operate in an environment where trust among participants can be found, and where sometimes the participants know each other, these blockchain types can limit the access level to the blockchain archive, ensuring that the users' identity remains properly protected.

78 *See* Reid, F. and Harrigan, M., (2013), "*An analysis of anonymity in the bitcoin system*", Security and privacy in social networks, Springer, New York, p. 15, *apud* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 27.

79 *See* Biryukov, A., Khovratovich, D. and Pustogarov, I., (2014), "*Deanonymization of clients in Bitcoin P2P network*" *in* Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (pp. 15-29), *apud* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 27.

4. Legal tensions between blockchain technology and the GDPR

The European Union's General Data Protection Regulation (GDPR) entered into force in May 2016 and became legally binding in May 2018, replacing the 1995 Data Protection Directive.[80] The GDPR establishes a homogenous legislative framework across the European Union, ensuring a high-level protection of natural persons and the removal of the obstacles to flows of personal data between all the Member States.[81]

The GDPR is an innovate legal framework that changed how data protection is perceived and how the processing of personal data should be regulated by the law.[82] By introducing new data protection rights and obligations, and enforcing a *'data protection by design*' approach, the reform of the European Union's legal framework on data protection has influenced the usage and development of new technologies such as blockchain. In this context, a common critique emerges among the data protection authors, who affirm that the European legislator has failed to take into proper consideration the emergence of new technologies, especially technologies that were under development when the GDPR draft was elaborated.

In fact, several tensions between GDPR and blockchain have been identified, revealing the difficulty GDPR has in keeping pace with blockchain technology.[83] Without prejudice to other factors, the tensions between GDPR and blockchain technology occur at two mains levels: firstly, the GDPR implicitly assumes that data is controlled or processed by identifiable actors;[84] and secondly, it also assumes that the data subjects' personal data can be rectified or erased in any case, to comply with the legal requirements set under articles 16 and 17 of the GDPR.[85]

In this section, after analyzing the GDPR territorial and material scope, and defining what is considered personal data in the context of blockchain technology, we will examine the legal tensions between the blockchain and the GDPR in further detail.

---

80 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) *in* Official Journal of the European Union, L 119/1, 4.5.2016, pp. 1-88.
81 Article 1(1) and Recital 10 of the GDPR.
82 *See* Giannopoulou, Alexandra and Ferrari, Valeria, (2016), "Distributed Data Protection and Liability on Blockchains", in Internet Science: 5th International Conference proceedings, Vol. 2. Workshops; Amsterdam Law School Research Paper No. 2019-06; Institute for Information Law Research Paper No. 2019-03. p 204.
83 *See* Finck, Michèle (2019), supra note 19, p. II.
84 *See* Lyons, T., Courcelas, L., and Timsit, K. (2018), supra note 51, p. 17.
85 *See* Finck, Michèle (2019), supra note 19, p. II.

4.1 The GDPR's applicability to blockchain-based platforms

Taking into account the objective of ensuring a consistent and homogenous protection of the natural persons with regard to the processing of their personal data, the GDPR's material and territorial scope is broad, covering a wide range of cases that also include the data processing activities taking place outside of the European Union territory.

With respect to its territorial scope, the article 3 (1) of the GDPR states that the regulation applies to all data controllers and processors established in the European Union, regardless of whether the processing takes place in the Union or not. The recital 22 of the GDPR clarifies that establishment of a controller or processor '*implies the effective and real exercise of activity through stable arrangements*', which suggests that the concept of establishment is not limited to its formal elements but, on the contrary, includes its functional elements as well.[86] In line with this approach is the case law of the European Court of Justice (ECJ), who in the *Weltimmo* case explained that '*the degree of stability of the arrangements and the effective exercise of activities (…) must be interpreted in the light of the specific nature of the economic activities and the provision of services concerned*'.[87] In the *Google Spain* case, the ECJ also stated that the '*the European Union legislature sought to prevent individuals from being deprived of the protection guaranteed by the directive and that protection from being circumvented, by prescribing a particularly broad territorial scope*'.[88] Thus, due to the flexible concept of '*establishment*', it is likely the territorial scope of GDPR will be fulfilled for most part of the blockchain operators established within the Union, unless the household exception under the article 2 (2) (c) of the GDPR applies.

In case the establishment criterion does not trigger the GDPR's application, the article 2 (a) and (b) of the GDPR extends its territorial scope to data controllers and processors not established in the Union, where the processing activities relates to the offering of goods or services to the data subjects who are in the Union,[89] and to the monitoring of the data subjects behavior, as far as their behavior takes place within the Union.

Due to its broad territorial scope, the GDPR is most likely to apply to a wide range of blockchain-based platforms and its operators. For instance, the GDPR will be applicable in relation to all blockchain operators who are established outside of the Union territory,

86 *See* Finck, Michèle (2019), supra note 19, p. 8.
87 *See* Case C-230/14, Weltimmo, EU:C:2015:639, 1 October 2015, para. 29.
88 *See* Case C-131/12, Google Spain, ECLI:EU:C:2014:317, 13 May 2014, para.
89 The reference of to the data subjects '*who are in the Union*' set under the article 3 (2) of the GDPR is related to the data subjects' location, not their nationality.

whenever they offer services to data subjects who are in the Union. In the same way, the operators of open and permissionless blockchain-based platforms are also subjected to comply with the GDPR rules, as it could be argued that those types of platforms offer services to data subjects who are in the Union. This is the case of Bitcoin, a platform that offers an electronic payment method to data subjects in the Union.[90]

Under the article 2 (1) of the GDPR, the regulation applies to the processing of personal data wholly or partly by automated means as well as personal data processing that relies on non-automated means, but forms part of, or is intended to form part of, a filling system. According to article 4 (2) of the GDPR, personal data processing is '*any operation or set of operation which is performed on personal data or on sets of personal data, whether or not by automated means*'. The general definition of '*processing*' includes the '*collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction*' of personal data. However, under its case law, in particular under the *Bodil Lindqvist* case, the ECJ noticed that these are mere examples of personal data processing, as the concept of '*processing'* is meant to be interpreted broadly.[91]

In this context, one can argue that the main functions of blockchain-based platforms are precisely to transmit, store and record personal data by automated means. For such reason, it could be said that blockchain participants are undoubtedly engaged in the processing of personal data, which, considering its material scope, triggers the GDPR's application, unless the household exception set under the article 2 (2) (c) of the GDPR is applicable. [92]

90 As we will analyze under the section 4.3 of your study, determining the data controllers and processors of a public and permissionless blockchain-based platform is not straightforward. In the context of the territorial scope of the GDPR, Bacon *et al.* suggest that nodes and miners are operators of the Bitcoin platform, as they support it collectively, and thus they are obligated to comply with the GDPR rules. Although we agree with the authors, it becomes clear that it is extremely difficult to identify them individually, which jeopardizes the GDPR's level of protection regarding the processing of personal data. *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 39.

91 *See* Case C-101/01, Bodil Lindqvist, EU:C:2003:596, 6 November 2003, para. 25.

92 *See* Giannopoulou, Alexandra and Ferrari, Valeria, (2016), supra note 82, p. 205.

4.2 Personal Data on Blockchain

In accordance with recital 26 of the GDPR, *the principles of data protection should only apply to any information concerning an identified or identifiable natural person*, which signifies that GDPR's applicability to the blockchain-based applications and their operators is, in any case, dependent on the qualification of the data stored and processed in the blockchain as personal data.

The article 4 (1) of the GDPR incorporates a wide definition of 'personal data'. It includes any information that directly or indirectly relates with an identified or identifiable natural person. In order to determine whether a natural person is identifiable, recital 26 of the GDPR, states that *all the means reasonably likely to be used (…) to identify the natural person directly or indirectly* should be taken into consideration. To ascertain what are the reasonable means likely to be used, objective factors should be taken into consideration. Recital 26 of the GDPR highlights some of those factors, which include: *i)* the costs of and the amount of time required for identification; *ii)* the technology that is available at the time of processing; and *iii)* the technological developments. The ECJ case law also reflects the broad interpretation of '*personal data'*. For instance, in the *Digital Rights Ireland* case, the ECJ has determined that the definition of 'personal data' is broad enough to qualify the metadata (*e.g.* the location of mobile communication equipment, IP address, *etc*.) as personal data, as the usage of this type of data makes it possible to identify a person, and it '*may allow very precise conclusions to be drawn concerning the private lives of the persons whose data has been retained*'. [93]

The online identifiers provided by the data subjects' devices, applications, tools and protocols, can also be used to directly or indirectly identify them. As the recital 30 of the GDPR recognizes, the online identifiers *'may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them*'.

The broad definition of personal data also includes the personal data which have undergone pseudonymization. Contrary to anonymization, the application of pseudonymization to personal data is deemed as a security measure that contributes to mitigate the risks to the data subjects in relation to the processing of personal data.[94] In fact, the Article

93 *See* Cases C-293/12 and C-594/12, Digital Rights Ireland, EU:C:2014:238, 8 April 2014, para. 26 and 27.
94 Recital 28 and Article 32 (1) (a) of the GDPR.

29 Working Party (hereafter WP29) expressly recognizes that '*pseudonymization is not a method of anonymization. It merely reduces the linkability of a dataset with the original identity of a data subject, and is accordingly a useful security measure*'.[95] In this sense, the usage of encrypted and hashed techniques to store and process data on blockchain-based applications are deemed to be qualified as a particular method of pseudonymization, considering that an important factor to qualify data as being anonymous is that the processing of re-identification of a natural person must be irreversible.[96] Bearing in mind the broad definition of personal data, blockchain-based applications are likely to process, at least, two types of personal data: public keys and transaction data.[97]

As previously explained, on blockchain-based applications, public keys are used essentially for identification purposes, while private keys are mostly used for authentication and encryption purposes. The private and public key pair, represented by a string of letters and numbers, is used to hide the real identity of the natural persons. As the WP29 sustains, the '*pseudonymisation is the process of disguising identities. The aim of such a process is to be able to collect additional data relating to the same individual without having to know his identity*'.[98] In this regard, and taking into consideration the provisions set under the article 4 (5) of the GDPR, it could be argued that the private and public key pair is likely to be qualified as a pseudonymization method, as the identity and other personal data can no longer be attributed to a specific data subject without using additional information. As the WP29 expressly recognizes, using a pseudonym means that it is still possible, under certain circumstances, to backtrack the individuals and discover their identities.[99]

Indeed, there are some practices and methods to determine the identify of the holders of a private and public key pair. Besides the voluntary disclosure of the private and public key pair, it is possible to identify a natural person when additional information is gathered in accordance with other regulatory requirements – as it is the example of the Anti-Money

95 *See* Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, p. 3.
96 As the WP29 acknowledges, '*anonymisation is increasingly difficult to achieve with the advance of modern computer technology and the ubiquitous availability of information. Full anonymisation would also require, for instance, that any reasonable possibility of establishing a link with data from other sources with a view to re-identification be excluded.*' *See* Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, p. 31.
97 *See* also Edgar, Laura, (2018), "*Blockchain and data protection: evaluating the legal compatibility of blockchain technology with the general data protection regulation*", Queen Mary University of London, Centre of Commercial Law Studies, p. 39.
98 *See* Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, p. 18.
99 *Ibid.*

Laundering duties – and then, combined with that, the specific private and public key pair.[100] On Bitcoin's platform, it is also possible to determine the identity of a data subject by linking its public key to its IP address.[101] The pattern of transactions can also be used to single out a particular data subject by using the 'transaction graph analysis' technique, which allows the determination of the identity of a certain unknown user by analyzing her transactional activity with a known user of a certain blockchain-based application.[102] In this context, one can argue that the private and public key pair should be regarded as person data in the terms of the article 4 (1) of the GDPR, since it could potentially lead to the direct or indirect identification of a natural person.

In many circumstances, the object of transactions – or the transactional data – can also be regarded as a personal data, as this type data can be linked to a real-world identity. Aside for the private and public key pair, the transactional data includes all the other categories of data that a transaction can contain. For instance, if a group of banks uses a consortium blockchain-based application to share Know Your Client information, the data contained in those transactions are deemed to be qualified as personal data, since such data concerns identified or identifiable natural persons. The transactional data can be used in plain text, in an encrypted form, or it can be hashed.

When transactional data is used in plain text, containing any information relating to an identified or identifiable natural person, there is no doubt concerning its qualification as a personal data.

As for encryption, as the WP29 correctly describes, it is one of the most used pseudonymization techniques.[103] As stated above, although this technique contributes to reduce the linkability of a particular dataset with the identify of a data subject, it is a useful security measure, but it cannot be considered an anonymization method.[104] Indeed, the holder of the private and public key pair can still be re-identified through the decryption processes. In this context, one can argue that personal data is still storage in a dataset that has been encrypted and, thus, encrypted data should be qualified as personal data.[105]

100 *See* Finck, Michèle (2019), supra note 19, p. 27.
101 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 40.
102 *Ibid.*
103 *See* Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, p. 20.
104 *Ibid*.
105 *See* Finck, Michèle (2019), supra note 19, p. 29.

Contrary to the use of encryption techniques, hash functions cannot be reversed, which means that once data has been put through a hash algorithm – such as the SHA-256 – that has transformed the input value into an output value with a fixed length, the hash function cannot run backward. Nevertheless, this does not automatically mean that hash functions are a method of anonymization,[106] as the linkability between a particular dataset and the hash function's output value can still be found. As the WP29 corroborates, in case the range of an input value is known, it can be replayed through a hash function, in order to achieve the accurate value of a particular dataset.[107] To Michèle Finck, a non-invertible hash function must ensure that the possible inputs are sufficiently large and unpredictable to prevent the option of trying all the possible combinations, but as the author recognizes, this is hard to achieve, especially if we are to consider the increasing power and decreasing cost of computing.[108] Therefore, following the WP29's opinion, hashing will generate pseudonymized data in most cases, even where hash functions with stronger privacy guarantees are used (*e.g.* salted hash, peppered hashes, keyed-hash functions with stored key, keyed-hash functions with deletion of the key, etc.).[109]

Without prejudice to a case-by-case analysis, it could be argued that encryption and hash functions are specific methods of pseudonymization that do not preclude the GDPR's applicability, considering the recital 26 test and the article 4 (1) and (5) of the GDPR.

4.3 Allocating responsibilities within blockchain platforms

Enhancing the protection of natural persons with regards to the processing of personal data is one of the two main objectives of the GDPR[110] and, for such reason, the accountability principle set under the article 5 (2) of the GDPR (which extends to processors) obliges controllers to take responsibility and demonstrate compliance with all the other principles set

106 As the WP29 states, the use of a (salted) hash function '*can reduce the likelihood of deriving the input value but nevertheless, calculating the original attribute value hidden behind the result of a salted hash function may still be feasible with reasonable means*.' *See* Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, p. 20.
107 The WP29 provides the following example: '*if a dataset was pseudonymised by hashing the national identification number, then this can be derived simply by hashing all possible input values and comparing the result with those values in the dataset*'. *Ibid*.
108 *See* Finck, Michèle (2019), supra note 19, p. 30.
109 *See* Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, p. 20.
110 Article 1 (1) of the GDPR.

under the same article. Indeed, controllers are obligated to implement appropriate technical and organizational measures in order to demonstrate that its data processing is performed in accordance with the GDPR.[111] When contracting with a processor to process personal data on the controller's behalf, the later shall also use processors who have provided sufficient guarantees that technical and organizational measures were implemented in accordance with the GDPR.[112]

The GDPR's structure defines the roles of controllers and processors in a clear and objective fashion, which is well adapted to scenarios where it is possible to find a central entity responsible for processing personal data, but remains inadequate to all the scenarios in which data is being processed in a distributed way.[113]

According to the article 4 (7) GDPR, the controller is any natural or legal person, which alone or jointly with others, determines the purpose and means of the processing of personal data. In the WP29's opinion, '*determining the purposes and means amounts to determining respectively the 'why' and the 'how' of certain processing activities*'.[114] The WP29 also clarifies that '*The concept of controller is a functional concept, intended to* ***allocate responsibilities where the factual influence is****, and thus based on a factual rather than a formal analysis*'.[115] Although the two elements, '*means'* and '*purposes*', appear to have an equivalent importance to determine who the controller is, in the WP29's opinion, the purposes criterion has primacy over the means criterion, as the '*determination of the "purpose" of processing is reserved to the "controller"*' and the '*determination of the "means" of processing can be delegated by the controller, as far as technical or organisational questions are concerned*'.[116] Aligned with this view, in the *Google Spain* case the ECJ also stated that, in order to ensure an effective and complete protection of data subjects, the concept of '*controller*' should be interpreted broadly.[117]

Finally, according to the article 4 (8) of the GDPR, the processor is any natural or legal person who processes personal data on behalf of the controller. The existence of a processor depends on the controller's decision, who might decide to delegate all or part of the processing

111 Article 24 (1) of the GDPR.
112 Article 28 of the GDPR.
113 *See* Ibáñez, Luis-Daniel, O'Hara, Kieron and Simperl, Elena (2018), supra note 4, p. 5.
114 *See* Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, p. 13.
115 *Ibid.* (our own emphasis).
116 *Ibid*., p. 15.
117 *See* ECJ, Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González, Case C‑131/12, 13 May 2014, para. 34.

activities to another natural or legal person. Therefore, two basic conditions must be present to qualify any person or entity as a processor: firstly, the processor must be a separate legal entity with respect to the controller; and secondly, the processing activities are conducted by that separate legal entity on the controller's behalf.[118]

As we will see below, identifying a controller or a processor in blockchain-based applications is not straightforward. In order to answer the question '*Who determines the purposes and means of data processing?*', it is not only necessary to consider the specificities of each case and the manner in which personal data is being processed, but also to examine the structure and governance design of the different blockchain platforms. Thus, considering the criterion provided by the WP29 and the relevant ECJ case-law, we analyze next the possible qualification of blockchain actors across the different blockchain types.

4.3.1 Allocating responsibilities within public blockchains

As mentioned under the section 2.2 of our study, there are three main actors on blockchain: the users, who propose new transactions; the nodes, who store copies of the distributed database; and the miners, who propose new blocks by executing a consensus protocol. Apart from these actors, we also have the developers, who produce the software which is used by nodes and miners to support the blockchain, and the wallet providers, who generate a private and public key pair on the users' behalf, providing them with a service that works as an interface to a specific blockchain platform, from which they can manage crypto coins or other digital assets. As we also explained before, open and permissionless blockchain-based platforms lack a central administrator, since the control over the platform is intentionally distributed. For this reason, determining who are the controllers and the processors has been a difficult exercise, as the definitions set under the article 4 (7) and (8) of the GDPR are ill-suited to these types of platforms.[119] Additionally, there is not a common understanding on the literature on DLTs and GDPR about which actor should be regarded as being the controller. [120]

The difficulty to determine who is a controller in an open and permissionless blockchain platform arises from two main factors: firstly, there is a wide number of the actors that influence

118 *See* Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, p. 25.
119 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 43.
120 *See* Finck, Michèle (2019), supra note 19, p. 43.

the means of processing personal data; and secondly, the purposes of processing personal data are also fragmented.[121]

To surpass these difficulties, Bacon *et al*. sustain that determining who is the controller requires an analysis that should be based on a micro-level perspective (*i.e.* the individual transactions), where '*the choice of the blockchain platform*' is the decisive criterion.[122] Thus, the macro-level perspective, which determines the controller by taking into account the blockchain infrastructure as a whole (*i.e.* as a service) should be rejected. In our view, this position should be adopted alongside with the criterion extracted from the WP29's guidelines and the ECJ case-law. Indeed, the DLTs are a mere infrastructure where blockchain applications, its design and governance structure can be developed and, since the processing of a specific item of personal data is more relevant to the GDPR, the micro-level perspective is more adequate to determine which actor can be considered a controller.

In this context, and without prejudice of a case-by-case analysis, it could be argued that when deciding to use an open and permissionless blockchain platform (*e.g.* Bitcoin) for a specific purpose (*e.g.* to make a transaction), the users determine the '*purposes*' and '*means*'. In such case, the user has opted for using the Bitcoin blockchain (the *'means*') to make a transaction (the '*purpose*'), when she arguably had the option to choose a different type of payment and another platform to make the transaction. Unless the user is a natural person that is using Bitcoin blockchain in the course of a purely personal or household activity, as defined under the article 2 (1) (c) of the GDPR, she should be considered a data controller.[123]

In this regard, although nodes and miners exercise significant control over the means (*e.g.* Bitcoin blockchain) by choosing to run a specific software and its embedded protocols, they usually do not determine the purposes, which is, as we analyze above, the main criterion to determine who is the controller. [124] Therefore, generally speaking, the nodes and miners are considered to be data processors. [125] However, it must be acknowledged that, in certain cases, nodes and miners can define their own purposes and set up their own means. For instance, these

121 *Ibid*.
122 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, pp. 41-45.
123 *Ibid*. *See* also Finck, Michèle (2019), supra note 19, pp.46-47; Giannopoulou, Alexandra and Ferrari, Valeria, (2016), supra note 82, pp. 208-210.
124 *See* Finck, Michèle (2019), supra note 19, p. 47.
125 In its guidance, the French supervisory authority (Commission Nationale Informatique et Libertés) considers that '*The miners limit themselves to the validation of the transactions submitted by the participants and do not intervene on the object of these transactions: they do not determine how the finalities and the means will be implemented'*. *See* Commission Nationale Informatique et Libertés, 'Premiers Éléments d'analyse de la CNIL: Blockchain' (September 2018), p. 2 (translated).

actors can access the public database stored on the blockchain to collect personal data for commercial purposes, or they can also change the rules of the blockchain-based platforms by creating a fork in the chain. In such cases, nodes and miners became joint controllers in the meaning of article 26 of the GDPR.

4.3.2 Allocating responsibilities within private blockchains

Contrary to public and permissionless blockchain-based platforms, closed and permissioned blockchain-based platforms are usually controlled by a centralized entity (*e.g.* a company, a public agency, etc.), who not only controls the means, but in many cases also determines the purposes of the processing of personal data. In such cases, authors like Michèle Finck tend to qualify the platform operator as a data controller, since the means and purposes of the processing are essentially determined by such entity.[126]

By contrast, authors like Jean Bacon *et al*. consider that such a conclusion is only possible from a macro-level perspective, which focuses on the blockchain infrastructure as a whole. However, taking into account the micro-level perspective, that focuses on individual transactions, these authors consider that users should be considered data controllers, whereas the centralized entities only act as a data processor.[127] Using the example of a land registry, this conclusion is supported by the idea that users insert personal data onto private and permissioned blockchain-based platforms for their own purposes (*i.e.* register or transfer titles of land) and, since they also chose those platforms as a medium to execute their transfers, they also determine the means of processing.[128]

With due respect to both positions, and without prejudice to a more detailed case-by-case analysis, we sustain that the identification of a data controller should take into account the criterion defined by the Article 29 Working Party (WP29), in which the allocation of responsibilities should be based on where the factual influence could be found.[129] In this sense, if the users have limited choice regarding the platform, it is not feasible to sustain that they have a factual influence over the purposes and means of the processing. For instance, if a

126 *See* Finck, Michèle (2019), supra note 19, p. 44. *See* also Giannopoulou, Alexandra and Ferrari, Valeria, (2016), supra note 82, p. 208; and, Ibáñez, Luis-Daniel, O'Hara, Kieron and Simperl, Elena (2018), supra note 4, p. 5.
127 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p.42
128 *Ibid*.
129 *See* Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, p. 9.

government agency implements a land registry blockchain-based platform, compelling the citizens to adopt it when registering and transferring titles of land, the users do not exercise any factual influence over the means and the purposes of the processing. One could argue that, in any case, the users decide to continue using the land registry platform for their own purposes. However, in cases such as this, their decision is greatly influenced by the government agency and, thus, the users' decision is limited *ab initio*. On another hand, if the users decide to use a closed and permissioned blockchain-based platform, where they could have chosen another means, they should be considered data controllers, since they truly determine the purposes (*e.g.* transfer a digital asset to other person) and means (*e.g.* using Blockchain-as-a-Service). In a case like this, the entities offering Blockchain-as-a-Service should be considered a data processor, unless they use the personal data for their own purposes.

4.3.3 Allocating responsibilities within consortium blockchains

As we previously observed, consortium blockchains are a permissioned, narrowly distributed platform, controlled by a small number of trusted nodes. The R3 Corda is one of the best-known consortiums blockchain-based platforms, where more than 300 financial entities participate by sharing information and settling payments among themselves. There are also other consortium blockchain-based applications in which banks and other financial institutions share information about their clients in order to comply with Know Your Client (KYC) and Anti-Money Laundering (AML) laws. These types of platforms are commonly designed as closed and permissioned, as only authorized participants can use the platforms and access the blockchain database.[130]

Similar to the private blockchain-based platforms, on consortium blockchains, the participating nodes act as a centralized entity, exercising a factual influence on the platform by determining the means and the purposes of the processing. In this sense, at the users' level, the participating nodes should be regarded as data controllers. Indeed, in such cases, it must be observed that the financial entities choose to use a certain mean (*e.g.* to use R3 Corda and similar platforms) to submit data about their clients and use the data on the blockchain database for their own purposes (*e.g.* to comply with AML and KYC laws). Regarding the other participants of a consortium blockchain-based platform, who process personal data as nodes and miners, they should be considered data processors, unless they use the personal data that

130 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 43.

is stored on the database for their own purposes, in which case they should be qualified as data controllers.

4.4 Data Subjects Rights

The allocation of responsibilities within a certain blockchain-based application is of crucial importance not only to comply with the GDPR's technical and organizational requirements, but also to allow data subjects to exercise their rights. The articles 15 to 22 of the GDPR incorporate specific rights of the data subjects, among which there are the right of access (article 15), right to rectification (article 16) and right to erasure (article 17). As we analyze below, blockchain-based applications, especially the open and permissionless blockchains, impose serious limitations to the exercise of some of the data subject's rights and freedoms, as the immutability feature of those applications is hard to combine with the desirable flexibility of a database, which seems necessary to comply with the data subjects' requests.

In accordance with the article 15 (1) of the GDPR, the data subject has the '*right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed'*, and, where that is the case, access the personal data and the information considering, *inter alia*, the purposes of the processing, the categories of personal data, the recipients to whom the personal data have been disclosed, the duration of storage, and the existence of automated decision-making. Under the article 15 (2) of the GDPR, data subjects also have the right to be informed about the adoption of the appropriate safeguards set under the article 46 of the GDPR, with respect to the transfer of personal data to third countries or international organizations. As we will analyze on the next point of our study, transfers of personal data to third countries can create a legal tension between the GDPR and some blockchain-based applications, considering that nodes located in the European Union probably share data with nodes that are located in other jurisdictions, without relying on a legal basis for such transfer.

The data subjects' right of access may be difficult to exercise in a context where personal data is processed on a blockchain-based application, especially on the open and permissionless blockchains. Since blockchain-based applications often rely on the use of encryption

techniques and hash functions to pseudonymize personal data, it may be difficult for nodes to know exactly which data is stored on a blockchain database and provide the data subject with information concerning the processing of her personal data.[131] A similar problem arises in relation to the provision established under the article 15 (3) of the GDPR, which entitles the data subject to receive a copy of her personal data undergoing processing.[132] Taking into consideration the open and permissionless blockchain-based applications, it may be impossible for nodes to provide a copy of the undergoing personal data processing not only because they are in no position to know which data is being processed, but also because they can only provide the data subject with their local copy of the blockchain database, which does not guarantees, *per se*, that is the copy other nodes on the P2P network are using to process data. On the contrary, on closed and permissioned blockchains, data controllers are in better position to facilitate the exercise of the data subjects' rights, as the users and nodes of these type of blockchain-based applications, who are regarded as being the data controller, have more control over the processing of personal data and, in general, over the platform.

The right to rectification, established under the article 16 of the GDPR, provides the data subject with the right to obtain, from the controller, the rectification of inaccurate personal data. The right to rectification is intrinsically related to the accuracy principle set under the article 5 (1) (d) of the GDPR, which determines that '*every reasonable step must be taken to ensure that personal data that are inaccurate (…) are rectified without delay'*.

However, even though data subjects themselves can be qualified as data controllers in many cases, exercising the right to rectification can be extremely challenging, considering the immutability characteristic of blockchain technology. Indeed, exercising the right to rectification on open and permissionless blockchain-based applications is tremendously impractical or almost impossible. First, the nodes only can alter their own local copy of the blockchain database and, as we analyzed before, such modification is irrelevant as nodes and miners tend to use the blockchain database version used by the majority of blockchain's participants.[133] Second, it may be impossible for a data subject to identify all the nodes, or to identify enough nodes (51%) to create a fork on the blockchain, in order to rectify her personal data. Third, even if enough nodes were identified, it would be extremely difficult to ensure

131 *See* Edgar, Laura, (2018), supra note 97, p. 46.
132 *See* Finck, Michèle (2019), supra note 19, p. 72.
133 *See* point 2.2.1 of our study.

such level of coordination.[134] On the other hand, the operators and data controllers of closed and permissioned blockchain-based applications, as it is the case of private and consortium blockchains, are in a better position to comply with the data subject's requests, as the centralized characteristics of these types of blockchains allows for better reversibility than the open and permissionless blockchains.[135]

Under the provisions established under the article 17 (1) of the GDPR, the data subject has '*the right to obtain from the controller the erasure of personal data concerning him or her without undue delay'*. In accordance with article 17 (2) of the GDPR, in case the controller has made the personal data public, as it is often the case of open and permissionless blockchains, the controller, taking into account the available technology and the cost of implementation, shall inform other controllers which are processing the personal data over which the data subject has exercised her rights. Similar to what we examined in relation to the right to rectification, the right of erasure is intrinsically related to the accuracy principle set under the article 5 (1) (d) of the GDPR.

As stated above, the immutable characteristic of blockchain-based applications makes it difficult (or near impossible) to change or delete any data stored into the blockchain database. In this sense, it could be argued that the creation of hard forks inside a blockchain could be a valid option to comply with the data subjects' right to erasure. However, besides the difficulties originated in trying to achieve such level of coordination among the nodes, the creation of hard forks is of a very exceptional nature, and it does not constitute a viable method to ensure the exercise of the right to erasure.[136] Additionally, such option could lead to the inoperability of blockchain-based applications, as hard forks invalidate all the subsequent blocks in the chain, forcing the nodes and miners to re-hash, validate and append all the other valid blocks into the chain, which would require long periods of time and it would be particularly costly in some cases.

In this context, it seems that blockchain-based applications cannot comply with the right to erasure (especially the open and permissionless blockchain-based applications). Nevertheless, it has been argued that the meaning of '*erasure*' is open to interpretation, as it can include, for instance, the simple removal of personal data from a search index and not the

134 *See* Berberich, Matthias and Steiner, Malgorzata (2016), "Blockchain Technology and the GDPR - How to Reconcile Privacy and Distributed Ledgers", 2 Eur. Data Prot. L. Rev. pp. 422-426.
135 *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 47.
136 *See* Finck, Michèle, (2017), "Blockchains and Data Protection in the European Union", Max Planck Institute for Innovation & Competition Research Paper No. 18-01, p. 31.

personal data itself, as the ECJ ruled, which is regarded as a '*soft version of the right to be forgotten*'.[137] Although we tend to agree that the meaning of '*erasure*' is open to interpretation and the inclusion of the expression '*available technology*', in the article 17 (2) of the GDPR, suggests that blockchain features shall be taken into consideration in relation to the exercise of the right to erasure,[138] it seems improbable, at current time, that the design and key features of early blockchain-based applications would comply with the requirements of the article 17 of the GDPR. For the reasons already presented in relation to the right to rectification, private and consortium blockchain-based applications are, in principle, in a better position to comply with the right to erasure.

In accordance with the article 19 of the GDPR, the controller shall communicate any rectification or erasure to every recipient who received a certain data subject's personal data, unless this proves to be impossible or in case it involves a disproportionate effort. Once again, considering the difference between open and permissionless and closed and permissioned blockchain-based applications, the provision of the article abovementioned may not be applicable to the former since it can involve a disproportionate effort or, in some cases, it may be virtually impossible.

Although the majority of the most well-known blockchain-based applications, such as Bitcoin or Ethereum, does not allow data subjects to fully exercise their rights under the GDPR, some solutions have been presented and proposed to surpass those difficulties. As we will analyze next, the off-chain repository solution provides a simple and easy answer to mitigate the legal tensions that occur at the data subjects' rights and freedoms level.

4.5 Personal Data transfer to third countries

In line with articles 44 to 49 of the GDPR, personal data can only be transferred to third countries on the basis of an adequacy decision; if the controller or processor has provided appropriate safeguards, and an equivalent level of protection can be found; or on the basis of a derogation.

137 *Ibid*.
138 *See* Edgar, Laura, (2018), supra note 97, pp. 48-49.

As the article 45 of the GDPR establishes, transfers of personal data to a third country may take place where the Commission has decided that the third country, a territory, or one or more specified sectors within that third country ensure an adequate level of protection. An adequacy decision should be based on clear and objective criteria and, in particular, on the elements of the article 45 (2) of the GDPR. When assessing whether a third country '*offer*[s] *guarantees ensuring an adequate level of protection essentially equivalent to that ensured within the Union*', the Commission should evaluate if that third country embeds the core principles of the GDPR and operates in the same spirt as the Charter of Fundamental Rights and other relevant international instruments.[139] Where an adequacy decision as been made by the Commission following the requirements of the article 45 of the GDPR, transfers of personal data do not require any specific authorization.[140]

In the absence of an adequacy decision, transfers of personal data to a third country are still possible in case the controller or processor has provided appropriate safeguards and, on the condition, that enforceable data subjects' rights and effective legal remedies for data subjects are available.[141] Such safeguards include legally binding and enforceable instruments between public authorities or bodies; binding corporate rules; standard data protection clauses, adopted by a supervisory authority and approved by the Commission; binding code of conduct together with enforceable commitments of the third country's controller or processor to apply these safeguards; and approved certification mechanisms together with enforceable commitments of the third country's controller or processor to apply these safeguards.[142] Where any appropriate safeguards are provided, and in case the requirements set under the article 46 of the GDPR are met, the controller or processor can transfer personal data to a third country, regardless of whether that transfer relies on the usage of blockchain-based applications or any other type of technology.[143]

As the article 49 of the GDPR implicitly acknowledges, there is a hierarchy between the legal grounds that allow the transfer of personal data to third countries. In this context, the abovementioned article only allows personal data to be transferred to a third country in case of

139 Recitals 104 and 105 of the GDPR.
140 At the time of this study, the European Commission has recognized Andorra, Argentina, Canada (commercial organizations), Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Switzerland, Uruguay and the United States of America (limited to the Privacy Shield framework) as providing adequate protection. To consult an updated list of the Commission's adequacy decisions: https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en, [accessed on the 24th of August of 2019].
141 Article 46 (1) of the GDPR.
142 Article 46 (2) of the GDPR.
143 *See* Finck, Michèle (2019), supra note 19, p. 90.

the absence of an adequacy decision or of appropriate safeguards, on the basis of one of the conditions established under the article 49 (1) lit. (a) to (g) of the GDPR. Those conditions allow personal data to be transferred to a third country if the data subject has provided explicit consent after becoming aware of the risks of the transfer; if the transfer is necessary for the performance of a contract concluded between the data subject and the controller, or if such contract has been concluded in the interest of the data subject; if reasons of public interest justify the transfer; in case it is necessary to establish, exercise or defend legal claims; if the transfer is necessary to protect vital interests of the data subject; or, finally, if the transfer is made from a register which, according to Union or Member State law, is intended to provide information to the public and which is open to consultation. Although in most cases article 49 (1) of the GDPR provides compelling reasons for a transfer of personal data to a third country to be regarded as lawful, recitals 111 to 113 expressly recognize that those conditions only apply *'in residual cases where none of the other grounds for transfer are applicable*', considering that the transfers are occasional and, in most cases, necessary to achieve a legitimate purpose. In such cases and taking into consideration the other requirements of the article 49 of the GDPR, the controller or the processor may transfer personal data to a third country. However, considering the residual and occasional characteristics of those derogations, the article 49 of the GDPR does not provide an appropriate legal ground for all the cases dealing with considerable amounts of transfers of personal data to third countries.

Whereas blockchain-based applications rely on a wide P2P network, transfer of personal data to third countries or international organizations can generate legal tensions between blockchain and the GDPR, as the location of nodes cannot be controlled. On the contrary, blockchain-based applications that operate in a centralized manner can have a better control over the location of nodes and miners and decide to transfer or not personal data to a third country. Nevertheless, even on private and consortium blockchain-based platforms, nodes and miners can be located outside the European Union, as some types of platforms include the participation of subsidiaries or group companies. In such cases, a legal ground for such transfers still needs to be found and, as established under the article 13 (1) (f) of the GDPR, where personal data relating to a data subject are collected from her, the controller shall provide the data subject with information related to the intended transfer of personal data to a third country and the legal ground that lawfully allows such transfer.

5 BLOCKCHAIN: A TOOL TO ENHANCE COMPLIANCE WITH GDPR

As we have demonstrated previously, blockchain-based applications stand in tension with some provisions of the GDPR. Among others, these tensions are related to the identification of the data controllers and processors, the data subjects' rights and freedoms (namely the right of access, the right to rectification and the right to erasure), and the transfers of personal data to a third country or to an international organization. In part, this is due to the implicit GDPR's presumption that a single actor or a specific group of actors can be perfectly identified in all cases and qualified as either data controller or processor. However, as we also observed, the technological innovation made possible by blockchain technology profoundly changed the dynamics of the personal data processing, as in many cases, especially on public blockchain-based applications, the data subjects are themselves involved with data processing by copying, changing, sharing, and moving their own data through a (sometimes) wide and open P2P network. In face of that, a common critique has emerged, stressing out that even before the GDPR has entered into force and application, it was already outdated in relation to the innovative technologies such as blockchain.[144]

Although we tend to agree with such criticism, since, in certain circumstances, the expected technological neutral characteristic of the GDPR is not fully adapted to the latest technological advances, it can be argued that blockchain technology can be used to achieve the GDPR's objectives.[145] In fact, one could argue that blockchain technology has emerged precisely to give individuals more control over their data.[146] In this sense, some authors have highlighted the fact that blockchain technology and the GDPR share common values and principles with one another, which makes it possible to bring technology and law together.[147]

In view of the legal tensions that we have identified above, there are some solutions the legal literature has been studying in order to surpass them and to improve the level of protection the GDPR concedes to data subjects. In this context, blockchain-based applications architecture shall incorporate the principles of data protection by design and by default set under the article

144 *See* Cate, Fred H.; Kuner, Christopher; Lynskey, Orla; Millard, Christopher; Ni Loideain, Nora; and Svantesson, Dan Jerker B., (2018), supra note 18, p. 103. *See also* Finck, Michèle (2017), supra note 136, p. 34.
145 As per the Article 1 of the GDPR, the protection of natural persons with regard to the processing of personal data and the free movement of personal data are the two main objectives of the GDPR.
146 *See* Nakamoto, Satoshi, (2008), supra note 3, p. 1.
147 *See* Wirth, C. and Kolain, M. (2018), "*Privacy by BlockChain Design: A Blockchain-enabled GDPR-compliant Approach for Handling Personal Data*", Wolfgang Prinz and Peter Hoschka (eds) Proceedings of the 1st ERCIM Blockchain Workshop 2018, Reports of the European Society for Socially Embedded Technologies Privacy by BlockChain Design, ISSN 2510-259, p. 1.

25 of the GDPR, in order to allow natural and legal persons to take the fully advantages of blockchain technology, without undermining data protection rules and the rights and freedoms of the data subjects.[148] The legal literature has conceptualized a few solutions and strategies that could be adopted to achieve the GDPR's objectives and, more generally, to ensure compliance with the regulation. Without prejudice to other solutions, we will now provide a brief overview of the two main solutions we believe to be more feasible to be used on blockchain-based applications to achieve and comply with the GDPR's objectives.

5.1 Using blockchain technology to improve data subjects' control over personal data

By laying down rules relating to the protection of natural persons with regard to the processing of their personal data, the GDPR recognizes that control over personal data is a crucial element of data protection.[149] As examined above, the accuracy principle and the data subjects' right of access, as well as the right to rectification and erasure, provide data subjects with different means to exercise control over their personal data. Recital 7 of the GDPR expressly foresees that '*natural persons should have control of their own personal data'*. The control over personal data implies that, at least, two main elements can be observed: first, data subjects shall have the possibility to monitor how their personal data is processed; and second, data subjects should have the opportunity to decide who should have access to their personal data.[150] However, it could be difficult to ensure the data subject has such control over her data, as in most cases, the data subject simply relies on the data controller or processor to process personal data in a lawful manner.[151]

In this context, blockchain technology, if properly designed to such ends, could be used to enable data subjects to exercise such control over their data. Indeed, there are a few use-cases that provide us with an idea on how blockchain-based applications can be used in this regard. Besides the Estonian experience on the health field, in which the patients can manage

148 *See* Lyons, T., Courcelas, L., and Timsit, K. (2018), supra note 51, p. 29. *See* also Hildebrandt, M. and Tielemans, L. (2013) "*Data Protection by Design and Technology Neutral Law*" Computer Law & Security Review 19, p. 516.
149 Article 1 (1) of the GDPR.
150 *See* Finck, Michèle (2019), supra note 19, p. 92.
151 *Ibid*.

the access authorizations to their health data through a blockchain-based application,[152] there are other blockchain-based applications such as Patientory and MedRec that give data subjects considerable control over their personal data.[153-154] In this regard, Faber *et al*. have conceived and proposed an interesting solution that, once incorporated on blockchain-based applications' design, would allow data subjects to better control access to their personal data.[155] The idea of a Blockchain-based Personal Data and Identity Management System (hereafter, BPDIMS) is to '*provide a holistic, personal data management tool to the user, meaning that the user of the system can expect full transparency and control over his personal data*'.[156] For such purpose, the BPDIMS's design contain three blockchain layers: a smart contract layer, an access layer, and a hash storage layer. The smart contract layer is used to store conditions for data exchanges between user and service providers or purchasers.[157] The access layer contains a '*a tool to ensure privacy*', through which it connects an offline storage with the blockchain, allowing the data subjects ' *to control and own their personal data, while service providers are guests with delegated permissions*".[158] Finally, the hash storage layer, which is used to store hashes of data that are created when *'personal data of the user is verified by certain trusted authorities like government organisations who could verify the user's personal details*'.[159]

Although the solution conceived by Faber *et al*. is not universal and it could not be used in all personal data processing cases, more importantly, it shows that it is possible to design blockchain-based applications in a manner that is compatible the GDPR, considering that the proposed BPDIMS also relies on the solution we present next.

---

152 *See* Priisalu, J. and Ottis, R. (2017) "*Personal control of privacy and data: Estonian experience*", 4 Health and Technology 441, *apud* Finck, Michèle (2019), supra note 136, p. 92.
153 *See* https://patientory.com/technology/, [accessed on the 29th of August of 2019].
154 MedRec can be described as a *"combination of a social need with a technological enabler: a system that prioritizes patient agency, giving a transparent and accessible view of medical history*". As it explained on MedRec's website, "*Smart contracts act as an intelligent representation that links patients and providers to the addresses of existing medical records. Medrec does not 'store' the record directly; rather encodes metadata that allows records to be accessed securely by patients, unifying access to data across disparate providers. The metadata contains information about ownership, permission and the integrity of the data being requested*". *See* https://medrec.media.mit.edu/ [accessed on the 29th of August of 2019].
155 *See* Faber, B., Michelet, G., Weidmann, N., Mukkamala, R. R., and Vatrapu, R. (2019), "BPDIMS: A Blockchain-based Personal Data and Identity Management System", *in* Proceedings of the 52nd Hawaii International Conference on System Sciences, pp. 6859-6860.
156 *Ibid*.
157 *Ibid*.
158 *Ibid*.
159 *Ibid*.

5.2 The off-chain repository solution

Storing personal data into the blockchain arises numerous tensions with the GDPR. Even in cases where encryption techniques and hash functions are used to store personal data into the chain, these methods only pseudonymize personal data, which triggers the GDPR's applicability. In this context, it seems that blockchain technology and the GDPR stand in conflict with one another. The apparent antagonistic relationship between blockchain technology and the GDPR has been highlighted numerous times,[160] although further studies and experimentation indicate this technology may be suitable to achieve some of the GDPR's objectives and, when properly designed, it can comply with the data protection requirements.[161]

In this regard, the use of an off-chain repository provides us with a feasible solution that allows blockchain-based applications to operate in line with the GDPR's requirements. This approach suggests that personal data should be stored in an off-chain repository, while blockchain-based applications only store hashed links to the data residing on the off-chain repository (hashed data pointers). Such approach guarantees simultaneously that fragmented data becomes less attractive for hacking, while accessibility to the data in the database is not compromised.[162] Moreover, security measures can be adopted as the data subject's personal data could be stored in the off-chain repository in an '*encrypted form using symmetric encryption keys that are owned by the respective user who owns the data'*.[163]

The main advantages of such architecture are that, by storing hashed data pointers to data stored on the off-chain repository, data breaches can be detected more easily as any alteration made to the data stored on the off-chain repository can be spotted and, more importantly, since personal data is kept off the blockchain database, it allows data subjects to exercise their rights, namely, the right of access, the right to rectification, and the right to erasure.[164]

---

160 Jan Philip Albrecht, a member of the European Parliament, has reportedly stated that "*Certain technologies will not be compatible with the GDPR if they don't provide for [the exercising of data subjects' rights] based on their architectural design. This does not mean that blockchain technology, in general, has to adapt to the GDPR, it just means that it probably can't be used for the processing of personal data*". *See* Cate, Fred H.; Kuner, Christopher; Lynskey, Orla; Millard, Christopher; Ni Loideain, Nora; and Svantesson, Dan Jerker B., (2018), supra note 18, p. 103.

161 *See* Finck, Michèle (2019), supra note 19, p. 91.

162 *See* Faber, B., Michelet, G., Weidmann, N., Mukkamala, R. R., and Vatrapu, R. (2019), supra note 155, p. 6860.

163 *Ibid*.

164 Although hashed data pointers may persist on the blockchain database, this approach guarantees that personal data (including all instances of a private key for the encrypted data) can be deleted. *See* Bacon, J.; Michels, J.; Millard, C.; Singh, J. (2017), supra note 21, p. 48.

Many other solutions and technical approaches have been proposed as both areas of our study, blockchain technology and data protection, are being subject to further examination through interdisciplinary research.[165] Nevertheless, the solutions provided above show that blockchain-based applications can be designed to provide data subjects with more control over their data, by creating new forms of data management and sharing. The solutions exemplified above also prove that it is possible to use blockchain-based applications to achieve the GDPR's objectives.

165 For instance, Accenture has recently registered a patent for an editable blockchain, in which blocks of data can be edited, rewritten or removed without break the chain. For more details, *see* Edgar, Laura, (2018), supra note 97, pp. 49-52.

6 CONCLUSION

This study has examined the application of the General Data Protection Regulation to blockchain-based applications. As observed, a universal definition of blockchain technology is hard to achieve, since different blockchain-based applications can present different structures and operate in different environments. Nonetheless, in order to create an immutable, tamper-evident record of transactions (or any other dataset) between parties, blockchain technology relies on two main components: hash functions and a public key infrastructure. While the former is used to guarantee data integrity by creating an immutable and tamper-evident record of transactions or any other dataset, the latter is used for identity authentication's purposes, with private keys being used to encrypt data and create digital signatures.

Blockchain-based applications can use this technology to create platforms with different features. In this study, we observed that blockchain-based applications can be qualified as public, private or consortium blockchains, depending on the roles of the users, nodes and miners; on permissions regarding the ledger of transactions' visibility and accessibility; and on the control over the blockchain-based application's underlying software.

Public blockchain-based applications are designed as open and permissionless platforms, where anybody can participate, either as a user, a node or a miner. Since these types of platforms are meant to operate in trustless environments, they rely on resource-intensive consensus protocols, such as the *Proof-of-Work*. Public blockchain-based applications tend to offer high transparency, strong data integrity, and high resilience. As observed, the users of this type of applications are identified by their public key or by their address, which makes it difficult to identify the user's real-world identity. However, in case users obtain their private and public key pair in an online wallet service, their real-world identity can be found easily, since many of these services can request a proof of the identity of their clients to comply with Know Your Client and Anti-Money Laundering laws.

On the contrary, private and consortium blockchain-based applications operate in a more centralized manner, where only a trusted third party or a few selected groups of participants can join the network and act as a user, a node or a miner. Because a certain level of trust can be found among the network in which these applications operate, private and consortium blockchains are designed as closed and permissioned platforms. Considering that only a few selected groups of participants can join these platforms, there is no need to implement a

resource-intensive consensus protocol, which enables these types of platforms to process a large number of transactions in an efficient way. Nevertheless, when compared with public blockchain-based applications, the users of private and consortium blockchain-based applications are easily identified since the participants of such platforms are selected and not everyone can join the network.

In regards to data protection, it was observed that there are numerous legal tensions between blockchain-based applications and the GDPR. At first sight, it appears that this is an antagonistic relationship, since the GDPR relies on the implicit assumption that data is controlled or processed by identifiable actors (data controllers or processors), in a centralized manner, while blockchain-based applications operate in a decentralized manner, with multiple actors and participants within a distributed network. Consequently, this study identifies and examines the three main categories of legal tensions that may occur, which relates to the determination of data controllers and processors; the exercise of the rights and freedoms of data subjects; and the transfers of personal to a third country or to an international organization.

This study has shown that the qualification of users, nodes and miners as (joint) controllers or processors has not reached a consensus within the legal literature. Although a case-by-case analysis is needed, taking into consideration the Article 29 Working Party criterion of the '*factual influence*' over the purposes and means of personal data processing, our study examined each blockchain-based applications type's participants, in order to qualify them into the categories established under the GDPR. Our findings suggest that users should be regarded as data controllers in a general manner, while nodes and miners tend to be qualified as data processors.

In relation to the data subjects' rights and freedoms set under the GDPR, it has been shown that, on the public blockchain-based applications, data subjects are enable to exercise their rights, namely the right of access (article 15), the right to rectification (article 16), and the right to erasure (article 17). On the contrary, on private and consortium blockchain-based applications it is easier for data subjects to exercise their rights, although the immutable feature of blockchain-based applications could pose serious challenges for data controllers to comply with such requests.

Finally, this study observed that, on public blockchain-based applications, personal data could be transferred to a third country or an international organization without relying on any legal basis the GDPR provides. By contrast, the operators and participants of private and consortium blockchain-based applications are in a better position to control the location of

nodes and miners and, for such reason, these types of applications can comply with the transfer rules set under the articles 44 to 49 of the GDPR.

The study has also highlighted possible solutions for blockchain-based applications to comply with GDPR's requirements and to help achieve its objectives. If properly designed, blockchain-based application could be used to improve data subjects' control over their personal data, an objective that GDPR clearly establishes under its recital 7. The off-chain repository solution also shows that it is possible to take full advantage of blockchain technology without jeopardizing the data subjects' rights and freedoms.

Although this study's findings allow a first approach to these matters to be taken by legal professionals, researchers and students, further interdisciplinary research on the blockchain-based applications' technical structure, design and governance is still needed in order to achieve compliance with GDPR's requirements. On the other hand, an interdisciplinary approach is also needed at the policy-making's level. Legislators and regulators should find new ways to cooperate with tech developers and entrepreneurs in order to effectively regulate the personal data processing within blockchain-based applications, mitigating the legal uncertainties related to their use and allowing the development of the European technological market, while protecting the rights and freedoms of data subjects.

7.BIBLIOGRAPHY

Books

Bashir, I., (2017), "Mastering blockchain", Packt Publishing Ltd

Nian, L., Chuen, D. (2015), "Introduction to Bitcoin", Handbook of Digital Currency, Chapter 1.

Reid, F. and Harrigan, M., (2013), "An analysis of anonymity in the bitcoin system", Security and privacy in social networks, Springer, New York

Case Law

ECJ Cases C-293/12 and C-594/12, Digital Rights Ireland, EU:C:2014:238, 8 April 2014

ECJ C-101/01, Bodil Lindqvist, EU:C:2003:596, 6 November 2003

ECJ Case C-131/12, Google Spain, ECLI:EU:C:2014:317, 13 May 2014

ECJ Case C-230/14, Weltimmo, EU:C:2015:639, 1 October 2015

Guidelines

Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN

Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN

Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN

Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN

Commission Nationale Informatique et Libertés, 'Premiers Éléments d'analyse de la CNIL: Blockchain' (September 2018)

Journals

Bacon, Jean and Michels, Johan David and Millard, Christopher and Singh, Jatinder, (2017) “Blockchain Demystified”, Queen Mary School of Law Legal Studies Research Paper No. 268/2017

Berberich, Matthias and Steiner, Malgorzata (2016), “Blockchain Technology and the GDPR - How to Reconcile Privacy and Distributed Ledgers”, 2 Eur. Data Prot. L. Rev

Biryukov, A., Khovratovich, D. and Pustogarov, I., (2014),” Deanonymization of clients in Bitcoin P2P network” in Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security

Bonneau, J., Miller, A., Clark, J., Narayanan, A., A. Kroll, J., and Felten, E., (2015) “SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies”, IEEE Symposium on Security and Privacy

C., Henry, (2017) “Blockchain: Disrupting Data Protection?”, Privacy Law and Business International Report, November 2017; University of Hong Kong Faculty of Law Research Paper No. 2017/041

Cate, Fred H.; Kuner, Christopher; Lynskey, Orla; Millard, Christopher; Ni Loideain, Nora; and Svantesson, Dan Jerker B., (2018), "Blockchain versus Data Protection", International Data Privacy Law, Volume 8, Issue 2

Crosby, M., Nachiappan, Pattanayak, P., Verma, S., and Kalyanaraman, V., (2016), “Blockchain Technology: Beyond Bitcoin”, Applied Innovation Review, Issue No 2

D. Chaum (1983) “Blind Signatures for Untraceable Payments, Advances in Cryptology”, Proceedings of the Springer-Verlag Crypto'82, Vol. 3

Faber, B., Michelet, G., Weidmann, N., Mukkamala, R. R., and Vatrapu, R. (2019), “BPDIMS: A Blockchain-based Personal Data and Identity Management System”, in Proceedings of the 52nd Hawaii International Conference on System Sciences

Fabiano, N. (2018), “Blockchain and Data Protection: The Value of Personal Data”, J. Systemics, Cybernetics & Informatics

Finck, Michèle (2019), “Blockchain and the General Data Protection Regulation: Can distributed ledgers be squared with European data protection law?”, Panel for the Future of Science and Technology, European Parliamentary Research Service, Scientific Foresight Unit

Finck, Michèle, (2017), “Blockchains and Data Protection in the European Union”, Max Planck Institute for Innovation & Competition Research Paper No. 18-01.

Giannopoulou, Alexandra and Ferrari, Valeria, (2016), “Distributed Data Protection and Liability on Blockchains”, in Internet Science: 5th International Conference proceedings, Vol. 2. Workshops; Amsterdam Law School Research Paper No. 2019-06; Institute for Information Law Research Paper No. 2019-03.

Hildebrandt, M. and Tielemans, L. (2013) "Data Protection by Design and Technology Neutral Law" Computer Law & Security Review 19

Ibáñez, Luis-Daniel, O'Hara, Kieron and Simperl, Elena (2018), “On Blockchains and the General Data Protection Regulation”, EU Blockchain Forum and Observatory

Lyons, T., Courcelas, L., and Timsit, K. (2018), “Blockchain and the GDPR”, European Union Blockchain Observatory and Forum

Maurer, B., Nelms, T. C., & Swartz, L. (2013), “When perhaps the real problem is money itself! the practical materiality of Bitcoin”, Social Semiotics, 23(2)

Priisalu, J. and Ottis, R. (2017) “Personal control of privacy and data: Estonian experience”, 4 Health and Technology 441

Sater, Stan (2017), “Blockchain and the European Union's General Data Protection Regulation: A Chance to Harmonize International Data Flows”, Social Science Research Network

Schwerin, Simon (2018) “Blockchain and Privacy Protection in the Case of the European General Data Protection Regulation (GDPR): A Delphi Study”, The Journal of The British Blockchain Association, Vol. 1, Issue 1

Truong, N., Sun, K., Lee, G., Guo, Y. (2019) “GDPR-Compliant Personal Data Management: A Blockchain-based Solution”, IEEE transaction on information forensics and security

Vetter, Greg R., (2004) "'Infectious' Open Source Software: Spreading Incentives or Promoting Resistance?", Rutgers Law Journal, University of Houston Law Center, No. 2004-A-11

Wirth, C. and Kolain, M. (2018), "Privacy by BlockChain Design: A Blockchain-enabled GDPR-compliant Approach for Handling Personal Data", Wolfgang Prinz and Peter Hoschka (eds) Proceedings of the 1st ERCIM Blockchain Workshop 2018, Reports of the European Society for Socially Embedded Technologies Privacy by BlockChain Design

Wright, Aaron and De Filippi, Primavera, (2015) "Decentralized Blockchain Technology and the Rise of Lex Cryptographia", Social Science Research Network

Zheng, Z., Xie, S., Dai, H.N., Chen, X. and Wang, H., (2017), "An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends". In 2017 IEEE International Congress on Big Data

Legislation

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

Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce')

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)

Thesis

Edgar, Laura, (2018), "*Blockchain and data protection: evaluating the legal compatibility of blockchain technology with the general data protection regulation*", Queen Mary University of London, Centre of Commercial Law Studies

Websites

Khatri, Yogita, (December 28, 2018) “Electrum Wallet Attack May Have Stolen As Much as 245 Bitcoin” in https://www.coindesk.com/electrum-wallet-attack-may-have-stolen-as-much-as-245-bitcoin

Whitepapers

Nakamoto, Satoshi, (2008) “*Bitcoin: A Peer-to-Peer Electronic Cash System*”, www.bitcoin.org