ARTICLE

# Blockchain and Data Protection: A Compatible Couple?

Ceyhun Necati Pehlivan* & Inés Isidro Read**

In recent years, blockchain, the underlying technology behind cryptocurrencies, such as Bitcoin, has emerged as one of the technological innovations with the greatest potential to transform the economy and society. As various European institutions and data protection authorities point out, this type of technologies may, by its very nature, be unable to comply with the General Data Protection Regulation (GDPR). We believe that, in order to understand the risks that the use of blockchain technology could pose, it is necessary to examine the architecture and the specific characteristics of the technology in question, in particular the way in which personal data are stored or processed. Thus, the impact of blockchain on the privacy and personal data of data subjects generally requires analysis on a case-by-case basis. We conclude that, while there are some tensions to be resolved, blockchain technology is not per se incompatible with the GDPR.

**Keywords:** blockchain, DLT, data protection, privacy, GDPR, cryptocurrency

## 1 Introduction

In recent years, blockchain, the underlying technology behind cryptocurrencies, such as Bitcoin, has emerged as one of the technological innovations with the greatest potential to transform the economy and society.

While it is true that blockchain technology is still in the experimental phase, numerous examples have already been seen for its use in industry, such as the issuance of bonds in the banking sector and fintech,[1] the monitoring of poultry products in the stages of production, processing and distribution,[2] and the registration of bids in public tenders.[3]

### 1.1 *What Is Blockchain?*

Although the first publications date back to the early nineties, such as the article by Stuart Haber and W. Scott Stornetta on a chain of cryptographically secured blocks,[4] its boom came about as a result of the development of cryptocurrencies, especially when Bitcoin hit the headlines and grew significantly in value and popularity over the past few years.[5]

Blockchain is the technology based on which most cryptocurrencies work. While this is the most well-known use of blockchain, in fact the scope of such technology is much broader and ever-growing. Blockchain is ultimately a particular way of applying the technology known as distributed ledger technology (DLT).

## Notes

The opinions expressed in this paper are those of the authors and do not reflect the position of any institution.

* An Adjunct Professor at IE Law School. Editor-in-Chief of the Global Privacy Law Review. Managing Associate in Linklaters' Technology, Media and Telecoms (TMT) and Intellectual Property (IP) practice group in Madrid. Lawyer admitted to practice in Spain. Co-chair of the KnowledgeNet Chapter of the International Association of Privacy Professionals (IAPP) in Spain. Email: ceyhun.pehlivan@linklaters.com.

** A Junior Associate at Linklaters. She is admitted to the Bar in Spain. She has a Law and Business degree from the Complutense University of Madrid and holds an LL.M. degree from Carlos III University.

1 Banco Santander, *Santander Launches the First End-to-End Blockchain Bond* (12 Sept. 2019), https://www.santander.com/csgs/Satellite/CFWCSancomQP01/en_GB/Corporate/Press-room/2019/09/12/Santander-launches-the-first-end-to-end-blockchain-bond.html (accessed 15 Oct. 2019); Paulina Duran & Alun John, *World Bank Launches World-First Blockchain Bond*, Reuters (23 Aug. 2018), https://uk.reuters.com/article/uk-worldbank-cba-blockchain/world-bank-launches-world-first-blockchain-bond-idUKKCN1L80DZ (accessed 15 Oct. 2019).

2 Luzi-Ann Javier, *Yes, These Chickens Are on the Blockchain*, Bloomberg (9 Apr. 2018), https://www.bloomberg.com/news/features/2018-04-09/yes-these-chickens-are-on-the-blockchain (accessed 15 Oct. 2019).

3 Europa Press, *El Gobierno aragonés utilizará la tecnología blockchain para recibir las ofertas de sus contratos públicos* (The Government of Aragon will use blockchain technology to receive bids for its public contracts) (9 Jan. 2019), https://www.europapress.es/economia/red-empresas-00953/noticia-gobierno-aragones-utilizara-tecnologia-blockchain-recibir-ofertas-contratos-publicos-20190109191949.html (accessed 15 Oct. 2019).

4 Stuart Haber & W. Scott Stornetta, *How to Time-Stamp a Digital Document*, 3 J. Cryptology 99 (1991).

5 Satoshi Nakamoto, *Bitcoin: A Peer-to-Peer Electronic Cash System* (2008), https://bitcoin.org/bitcoin.pdf (accessed 15 Oct. 2019).

In simplified terms, DLT is a decentralized information recording system; a database that digitally stores data based on a consensus protocol in a replicated, shared and synchronized manner among multiple servers, locations and institutions, without any central authority.[6] In a blockchain network, this shared record has the distinctive feature of storing the information in a consecutive series of blocks linked to one another by means of a signature or cryptographic seal called a 'hash', which is the result of applying a mathematical algorithm to such digital content resulting in a unique concatenation of alphanumeric characters.

Seen from this broader perspective, the door is open to host an unlimited number of applications for blockchain technology that go beyond the mere recording and transfer of digital financial assets. It becomes a decentralized way of executing, formalizing and managing a variety of transactions, ranging from public record-keeping or production flows, to e-voting management, Smart Contracts, product tracking along the supply chain, or autonomous vehicles sharing driving data with each other.

### 1.2 *How Does Blockchain Interact with Data Protection Regulations?*

If considered as a database or registry, platforms based on the blockchain model are likely to contain personal data, as discussed further below. Their use could therefore come into interaction or, in some cases, even conflict with the EU General Data Protection Regulation 2016/679 (GDPR).[7]

The GDPR seeks to reduce the asymmetry of power between the organizations that process personal data as data controllers and the data subjects. This objective is intended to be achieved by, on the one hand, reducing the influence of the party that centralizes greater decision-making capacity and control over such data, (i.e. data controllers), and, on the other hand, strengthening the rights and freedoms of data subjects, for example, by greater transparency in the processing of their personal data or introducing additional rights.

The blockchain technology shares, in theory, this objective of giving data subjects more power, but instead of doing it by regulating the way of exercising that centralized control, it does so by attacking its very existence. Thus, these centralized figures with significant duties and responsibilities under the GDPR cease to exist in such decentralized scenarios.

This is the reason why there is currently a heated debate about the potential incompatibilities between certain uses of the blockchain technology and the provisions of the GDPR, a debate that has not yet been conclusively resolved by the data protection authorities or courts.

As various European data protection authorities point out, in order to understand the risks that the use of blockchain technology could pose, it is necessary to examine the architecture and the specific characteristics of the technology in question, in particular the way in which personal data are stored or processed. Thus, the impact of the blockchain on the privacy and personal data of data subjects generally requires analysis on a case-by-case basis.

Among the various questions that stir up controversy, we will analyse those that seem most relevant to us, as well as the possible solutions that could be offered considering the current state of blockchain technology.

## 2 *Public v. Private Blockchain*

In order to analyse a potential incompatibility of the application of this technology with the GDPR, it is necessary to differentiate the various types of blockchain networks considering their access requirements and the powers of the actors operating in them.

Blockchain technology relies on a set of devices called 'nodes' that keep synchronized copies of the same data. Blockchain is structured as a peer-to-peer (P2P) network architecture, meaning that the nodes participating in the network are equal to each other and are responsible for providing services to the network and updating records in a distributed manner. It is therefore a mesh network in which the participating nodes are directly interconnected. There is no central server, centralized service, authority or hierarchy within the network. Therefore, it is common for blockchain networks that changes in the registry are not updated simultaneously at all nodes and delays may occur until all registries and nodes are synchronized.

Although each participant in a DLT is a node, not all participating nodes in the network have the same role. The different functions performed by the nodes depend on each particular network, such as storing a copy of the blockchain, validating transactions or approving the final record of a transaction.

In general, we can distinguish between the following nodes:

(1) full or validation nodes, which validate each block in the chain according to the rules of the network. Also, these nodes maintain a copy of the complete

*Notes*

[6] UK Government Office for Science, *Distributed Ledger Technology: Beyond Block Chain* (2016), https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/492972/gs-16-1-distributed-ledger-technology.pdf (accessed 15 Oct. 2019).

[7] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 Apr. 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).

chain of blocks ensuring that blockchain remains immutable and decentralized;

(2) lightweight nodes, which emit transactions and receive information from third parties, keeping partial copies of the chain (i.e. the block headers) in synchronized form and distributing it across the network. These nodes do not verify the complete chain, but only the authenticity of the new transaction included in the block header; and

(3) mining nodes, which are responsible for 'mining' the new transactions that are introduced into the network through a predetermined algorithm known as a consensus mechanism and generating new blocks by adding to each of them the hash that corresponds to it computationally.

Depending on the existing requirements to be a node and the functions they perform, among other factors, we can distinguish between the following networks:

### 2.1 *Public or Open Blockchain Networks*

Public blockchain networks are those in which anyone can own and manage a participating node on their local machine and validate transactions on the network. The idea is to replace the traditional bilateral provider/customer relationship with a model based on the collective processing of data through a protocol shared among all users. The best-known examples of public blockchain networks are Bitcoin and Ethereum, both created through open source software that can be inspected, verified and downloaded by anyone who wants to be a node.

In such public networks, all the members can access and view the status of the transactions and information contained in the verified blocks of the chain without being an active participant in it or restriction. In this regard, a public blockchain network complies more accurately with the principles of decentralization and independence.

The validation nodes, in addition to verifying the transactions individually, close each block using complex mathematical algorithms that generate a numerical combination or 'nonce'. This nonce is then added to a hashed block in the chain (the so-called 'Proof of Work') and aims to reduce the possibilities of fraud. The search for a nonce is what is popularly known as 'mining', and when the solution is found, miners are offered cryptocurrency in exchange. In a public blockchain network it is easier to carry out the mining process to verify transactions and add them to the blockchain, since any user who meets the computational force requirements can apply the process.

Thus, the data contained in the blocks of the public network chain are publicly available at all nodes of the network, so that the data cannot, in principle, be modified or deleted, as discussed below. This is precisely the reason why this type of open blockchain is called an 'immutable record'. However, to the extent that such records contain personal data, this immutable architecture of blockchain public networks comes into conflict with the provisions of the GDPR. In particular, this immutability may clash with certain rights of the data subjects such as the right to be forgotten or right of erasure, or the right of rectification. We will discuss the nature of the data involved, as well as their anonymization and pseudonymization later.

### 2.2 *Private or Closed Blockchain Networks*

Private blockchains are owned by an individual or institution and can only be accessed with an invitation from the owner, developer or administrator. Similarly, when accepted, the user's role will be subject to the internal rules of the private network. Examples of private blockchains include Ripple (a protocol to facilitate international money transfers), Hyperledger (initiated by the Linux Foundation for global business transactions), and R3 (a consortium of international banks to develop banking solutions based on a private blockchain network).

In this type of blockchain network, there is a party in charge of managing the network, which therefore provides a centralized and hierarchical structure. This feature makes this type of blockchain more common in consortiums created by financial institutions or others, such as market regulators, to develop new ways of offering their services. There is therefore an entity that operates the network and has special powers as network administrator, such as deciding who can participate or mine the network. This introduces a certain degree of centralization, which goes, according to some, against the decentralized nature and philosophy of blockchain technology.

Nevertheless, the centralized structure of a private blockchain network could help its participants to comply with the GDPR, determining in a clear way the identity and responsibilities of the data controller, intervening in the network or even modifying the blocks of the chain when individuals exercise their rights.

### 2.3 *Hybrid Blockchain Networks*

In addition to the public and private networks mentioned above, there is a third form of blockchain which is the mixture of both networks, a hybrid blockchain.

In this type of blockchain it is usual that the participating nodes need to be invited, but that the transactions are made public. That is to say, the nodes participate in the operation and security of the blockchain, but unlike the private blockchain networks, the transactions are not private, but visible to users all over the world for the sake of greater transparency.

It is worth mentioning that there are all sorts of hybrid blockchains between these two main types of network, and the more public a network is the more challenging it becomes from a data protection standpoint, especially regarding the identification of its actors and their

responsibilities under the GDPR. The network governance system will be especially critical to ensure adequate risk management and, ultimately, to maintain the stability and security of the system.

## 3 Possible challenges of blockchain technology in relation to the GDPR

### 3.1 *Identification of the Data Controller and the Data Processor*

In the same way as the previous European Data Protection Directive[8] did, the GDPR, applicable since May 2018, includes two main actors in the framework of the processing of personal data: the data controller and the data processor.

The data controller is defined as the one who determines the purposes and means of such processing,[9] whereas the data processor is the one who processes the data on behalf of and following the documented instructions of the data controller.[10] Their identification is required for the fulfilment and even enforcement of the provisions of the GDPR, since these two actors bear different obligations to ensure the protection of the personal data of the individuals. Similarly, data subjects need to be able to identify these actors so as to be able to exercise their rights[11] and claim compensation from the controller or processors for the damage suffered.[12]

But, who are the data controller or processor in a blockchain network?

The answer to this question is controversial. The legal need to identify the data controllers and processors in a blockchain network may collide with the decentralized essence of the blockchain technology, where the participants, such as the owner of the network, participating nodes, technology developers, and the data subjects themselves, are considered equal, though they may have different roles in the manner in which they support the network. Therefore, depending on the type of blockchain, it becomes highly difficult to determine who holds the status of data controller and processor on the network.

#### 3.1.1 *Data Controller*

As discussed above, it might be relatively easier to identify these actors in private networks, where there is usually a centralized structure and hierarchical governance scheme. Since most of these cases are collective initiatives including several entities, the members of the network could be considered as joint controllers in accordance with Article 26 of the GDPR, as further examined in the next chapter. As the European data protection authorities point out, it is strongly recommended to identify clearly and at an early stage the responsibilities of each participant involved in the project.

However, the difficulties of identification increase as the degree of openness or publicity of the network increases. This is due to the decentralization of the control of information and a certain level of anonymity offered by the public networks that predominate in the market.

This may result in additional problems such as territoriality or applicable law, making it difficult to control that personal data included in a public network is not transferred or processed in territories that, according to the GDPR's international data transfer restrictions, do not ensure an adequate level of protection and safeguards equivalent to those provided in the European Union.

In general, the French data protection authority *Commission Nationale de l'Informatique et des Libertés* (CNIL) argues that all participating users with the capacity to record transactions in the blockchain and distribute the data on the network for validation by the miners, can be considered as data controllers. This is because they would, through their decisions, define the purposes and means of the processing, such as the format in which the data are recorded, or the actual use of the blockchain technology.[13]

In particular, the CNIL consider that such users would be data controllers when they are:

- natural persons who process data in the course of a professional or commercial activity, thus excluding any exclusively personal or domestic activity in accordance with the material scope of the GDPR[14]; or,
- legal entities.

**Notes**

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

[9] Article 4(7) of the GDPR.

[10] Article 4(8) of the GDPR.

[11] Chapter III of the GDPR.

[12] Article 82 of the GDPR.

[13] Commission Nationale Informatique & Libertés (CNIL), *Solutions for a Responsible Use of the Blockchain in the Context of Personal Data*, https://www.cnil.fr/sites/default/files/atoms/files/blockchain.pdf (accessed 15 Oct. 2019).

[14] According to Art. 2(2)(c) of the GDPR, '[t]his Regulation does not apply to the processing of personal data: … by a natural person in the course of a purely personal or household activity'. Furthermore, Recital 18 clarifies that the GDPR does not apply to the processing of personal data by a natural person in the course of purely personal or household activity and thus with no connection to a professional or commercial activity. It also indicates that personal or household activities may include correspondence and the holding of addresses, or social networking and online activity undertaken within the context of such activities. However, the GDPR applies to controllers or processors who provide the means for processing personal data for such personal or household activities.

To illustrate this, the CNIL quotes the example of a notary public who, in his/her capacity as data controller, records the property deed of a client on a blockchain network.

Regarding the nodes, the opinion of the French authorities does not provide a detailed analysis. As mentioned above, both lightweight and full nodes would be responsible for validating user blocks and transactions and for updating the chain, either in part or in full. They also keep a copy of the transaction and distribute it to other nodes.[15] In other words, these nodes are the ones that 'decide' on the creation of the blocks, as well as the processing of the data. Therefore, those nodes that participate in the validation and creation of the blocks should be considered, in our opinion, as data controllers.

However, the identification of the controllers does not necessarily solve the problem that lies under the GDPR, and, in certain cases, makes it even worse. Once the controller has been identified, there is still the question of with whom the data subjects can actually exercise their rights or whom the authorities may hold accountable.

In a private or hybrid blockchain network the criteria would be somewhat easier to apply since there is an identified entity or group of entities that operate the network and they would be considered as data controllers against whom the data subjects and data protection authorities may take action.

However, in a public and fully decentralized blockchain network, the picture is different and the answer to the question is far from obvious. From the data subjects' perspective, the data would be hosted on thousands of anonymous user computers spread all over the world. In this regard, the application or enforcement of the GDPR is considerably more difficult.

### 3.1.2 *Joint Data Controllers*

As mentioned above, it is common for different entities or consortiums to jointly participate in private or hybrid blockchain networks. These network members could be considered as joint controllers in accordance with Article 26 of the GDPR.

Although the courts have not yet ruled on this matter in the field of blockchain technology, the precedents set by the Court of Justice of the European Union (CJEU) seem to be inclined to follow a broad interpretation of the concept of joint data controllers. This is demonstrated by its recent ruling of 5 June 2018, which responds to the request made by the Administrative Chamber of the German Supreme Court *Bundesverwaltungsgericht*[16] regarding the role played by the administrators of fan pages of the social network Facebook under the GDPR. The CJEU finds that each and every one of the administrators of these pages should be regarded as joint data controllers together with Facebook within the EU. Its decision was based on the fact that, merely by creating these pages, these administrators make an active and voluntary contribution to Facebook's collection of personal data through cookies relating to visits to these pages. That is to say, from their mere existence it can be inferred that there is a configuration action, by means of which an administrator takes part in the determination of the purposes and means of processing the personal data of the visitors to its fan page. In addition, the administrators also receive anonymous statistical and demographic data on visitors to the fan pages, and thereby they determine the way in which personal data are processed, even though they do not have access to such data. The broad application of this concept leads to the inference that the nodes in blockchain networks also take action and decide on the purposes of the processing, similar to the conclusions of the CNIL.

From other rulings, such as that handed down in the case of the Jehovah's Witnesses Community[17] in relation to the collection and processing of personal data in the course of door-to-door preaching by its members, we can conclude that it is not necessary for the controller to give written guidelines, orders or instructions about the processing to be considered as such. It is enough to organize, coordinate and encourage the data processing activities jointly with others.

This ruling may also provide an answer to the question of whether the nodes make personal use of the network, since the CJEU found that such use cannot be regarded as being personal or domestic if the purpose is 'to make the data collected accessible to an unrestricted number of people or where that activity extends, even partially, to a public space and is accordingly directed outwards from the private setting of the person processing the data in that manner'.[18]

However, as a reassuring element, the CJEU clarified in the above rulings that in both cases the responsibility of these joint controllers is not equivalent to that of Facebook or the Jehovah's Witness organization, but that 'those operators may be involved at different stages of that processing of personal data and to different degrees, so that the level of responsibility of each of them must be assessed with regard to all the relevant circumstances of the particular case'.

**Notes**

[15] Mario Martini und Quirin Weinzierl, *Die Blockchain-Technologie und das Recht auf Vergessenwerden*, 17 Neue Zeitschrift für Verwaltungsrecht 1251 (2017).

[16] *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v. Wirtschaftsakademie Schleswig-Holstein*, Case C-210/16 (CJEU 2018).

[17] *Tietosuojavaltuutettu v. Jehovan todistajat (Jehova's Witnesses) – uskonnollinen yhdyskunta*, Case C-25/17 (CJEU 2018).

[18] *Bodil Lindqvist v. Åklagarkammaren i Jönköping*, C-101/01 (CJEU 2003); *Tietosuojavaltuutettu v. Satakunnan Markkinapörssi and Satamedia*, C-73/07 (CJEU 2008), and *František Ryneš v. Úřad pro ochranu osobních údajů*, C-212/13 (CJEU 2014).

It is worth mentioning that, as a consequence of the shared responsibility of the nodes jointly participating in a network, it is necessary to identify in a clear and transparent way the responsibilities corresponding to each of them for the fulfilment of the obligations imposed by the GDPR. This is especially important with respect to the exercise of data subjects' rights and the obligations of the joint controllers to provide information to them.

### 3.1.3 *Data Processor*

In order to determine whether there is a data processor with respect to a DLT-based personal data processing activity, it is necessary to consider each case separately and to analyse the data flows involved in the blockchain network.

In general, however, we may state that mere intermediaries or external service providers, such as companies offering software solutions or providing blockchain infrastructure services, or Blockchain-as-a-Service (BaaS) similar to software, platform, or infrastructure as a service (SaaS, PaaS, or IaaS, respectively) for the purpose of establishing or maintaining nodes connected to the blockchain on behalf of third-party controllers, should be deemed to be data processors.

Also, according to the opinion of the CNIL,[19] software and smart contracts developers, whose activity is limited to creating tools without determining how they should be used or by whom, should be considered as processors, since they are only responsible for processing the information on behalf of the controller.

Similarly, the CNIL considers that the mining nodes would be data processors in certain cases. Indeed, in our opinion, the mining nodes cannot be considered as controllers, since they do not determine the purposes of the transactions introduced in the blockchain network, but merely seek a valid hash with their computational power to generate a new block and communicate it to other nodes so that they can verify it and update the chain of blocks. As compensation for their efforts, miners are awarded new coins created with each new block, and transaction fees from all the transactions included in the block.[20] The miners are responsible for generating new blocks, but they are not responsible for maintaining the blockchain. However, there is some debate about this, since the miners do add additional information to the register, keep a copy of it in their equipment, and exercise some control over the means of the processing, such as choosing the version of the consensus protocol.[21]

Consequently, data processors are required to comply with the provisions of Article 28 of the GDPR, including, among others, to execute a data processing agreement with their data controllers. The truth is that, as admitted by the CNIL itself, this obligation becomes practically impossible in public blockchain networks where thousands of participants operate on the network without necessarily knowing each other.

## 3.2 *Data Protection Principles*

Article 5 of the GDPR sets out a number of principles to be taken into account in the collection and processing of personal data which are essential for their protection. We will discuss below some of these principles that we consider to be of special interest for the use of blockchain technology, whose operation and purpose could call into question the ability to comply with some of them.

### 3.2.1 *Lawfulness of the Processing*

The first principle that could be affected is that of lawfulness, fairness and transparency in relation to the data subject. In particular, it may not be obvious on which lawful basis the processing of data takes place in this type of networks; especially in public ones.

If consent is to be relied upon,[22] it must first and foremost be free, specific, informed and unequivocal, which implies a statement or clear affirmative action on the part of the data subject.[23] Some authors have suggested that, for example, in the case of Bitcoin it can be argued that data subjects implicitly consent to the processing of their data, in particular, the Bitcoin address when creating a Bitcoin account or wallet.[24] We however consider that such consent would not be valid as it is neither an affirmative action nor specific in accordance with the GDPR.[25] It is also worth mentioning that the data

**Notes**

[19] Commission Nationale Informatique & Libertés (CNIL), *supra* n. 13.

[20] Arvind Narayanan et al., *Bitcoin and Cryptocurrency Technologies: A Comprehensive Introduction* (Princeton University Press 2016).

[21] European Parliament, Panel for the Future of Science and Technology, *Blockchain and the General Data Protection Regulation – Can Distributed Ledgers Be Squared with European Data Protection Law?* 46 (July 2019), https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634445/EPRS_STU(2019)634445_EN.pdf (accessed 15 Oct. 2019).

[22] Article 6(1)(a) of the GDPR.

[23] Article 4(11) of the GDPR.

[24] Jean Bacon et al., *Blockchain Demystified*, Queen Mary School of Law Legal Studies Research Paper No. 268/2017, 46 (2017); referred to in European Parliament, Panel for the Future of Science and Technology, *supra* n. 21, at 61.

[25] Recital 32 of the GDPR clarifies that consent should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her, such as by a written statement, including by electronic means, or an oral statement. Furthermore, it states that this could include ticking a box when visiting an internet website, choosing technical settings for information society services or another

subjects may withdraw their consent at any time,[26] in which case the data must be deleted as there is no other purpose justifying the retention of such data.[27] Therefore, this requirement may be particularly difficult to implement in a blockchain network. In light of the above and the immutable character of the blockchain, we consider that consent would not be the most appropriate legitimate basis for the processing of data in the blockchain technology.

On the other hand, the need for processing for the performance of a contract as a lawful basis[28] may be relevant in the case of service providers such as banks or insurers with respect to their customers. However, in order to ensure the possibility of relying on this lawful base, it should at least be possible to identify the terms and parties thereto, which is no easy task in public blockchain networks taking into account the automation and anonymity of most of the actors in these platforms.

Likewise, the need for processing for compliance with a legal obligation to which the controller is subject[29] must be taken into account as a lawful basis by those controllers who are subject to *Know Your Client* obligations or anti-money laundering and counter terrorist finance compliance in the context of cryptocurrency transactions.

Finally, the application of a legitimate interest[30] as a lawful basis for the use of blockchain would be limited in practice. In any case, it should be mentioned that the processing of data for a legitimate interest will require a specific assessment, taking into account the reasonable expectations that the data subjects would have at the time and in the context of the collection of their personal data, in order for the processing to take place on that basis.[31] With regard to public blockchain networks, bearing in mind that most users are not aware that public keys are personal data in accordance with the GDPR and that transactions may reveal information about them, we note that it may be difficult to argue that data subjects could reasonably expect their data to be processed for the purposes of a legitimate interest.[32]

### 3.2.2 *Data Governance*

The greatest of the dilemmas in both public and private networks is found in relation to the clash with the principles of storage limitation, data minimization and accuracy.

On the one hand, personal data should be kept in a form that permits identification of data subjects for no longer than is necessary for the purposes of the processing.[33] In addition, the data should be adequate, relevant and limited to what is necessary in relation to the purposes for which they have been collected[34] and should be updated or rectified where appropriate.[35] This makes it unlawful to retain data for an indefinite period of time and implies that the controller must be able to modify the information when it is incorrect or to delete the information or part of it when it is no longer necessary or when requested by the data subjects, with some legal exceptions.

These principles could collide with the way in which some blockchain networks store data. The cornerstone of this technology is that it is designed in such a way that data, once included in the blockchain, cannot be altered or deleted. As an example of the technical difficulty involved in rectifying data in these networks, we have seen cases in which it was necessary to apply the technique known as 'Hard Fork'[36] or hard bifurcation to annul erroneous transactions. This technique consists essentially of splitting the chain from the point in which the error was found and rewriting all the information in a new parallel chain.

However, there are also advanced projects that work on the creation of private blockchain networks editable by

**Notes**

statement or conduct which clearly indicates in this context the data subject's acceptance of the proposed processing of his or her personal data. Silence, pre-ticked boxes or inactivity should not therefore constitute consent.

[26] Article 7(3) of the GDPR.

[27] This does not mean that the data controller may substitute another legal basis for consent, as set out in the Guidelines of the Art. 29 Working Party on Consent under Regulation 2016/679, WP259 and rev.01, at 25.

[28] Article 6(1)(b) of the GDPR.

[29] Article 6(1)(c) of the GDPR.

[30] Article 6(1)(f) of the GDPR.

[31] Recital 47 of the GDPR.

[32] European Parliament, Panel for the Future of Science and Technology, *supra* n. 21, at 64.

[33] Article 5(1)(e) of the GDPR.

[34] Article 5(1)(c) of the GDPR.

[35] Article 5(1)(d) of the GDPR.

[36] By way of example, we could mention the Hard Fork Ethereum case in 2016, https://blog.ethereum.org/2016/07/20/hard-fork-completed/ (accessed 15 Oct. 2019). Also, as stated in the European Parliament's 2017 report: 'In June 2016 an attack exploited weaknesses in the [decentralized autonomous organizations (DAO)]'s code, siphoning almost one third of its assets and sparking a controversy in the community about what to do next. The options were to freeze funds in the account (a "soft fork"), to hack the system and restore the original balance (a "hard fork"), or to do nothing at all. On one hand, since the attacker(s) exploited a weakness in the code, it could be argued that they did not breach the contract and that modifying The DAO's blockchain would undermine public confidence in its principle of immutability. On the other hand, the attack clearly went against the spirit of the contract, may have contravened contract law and could discourage actual and potential participants in the community'. European Parliament Research Service, Scientific Foresight Unit (STOA), *How Blockchain Technology Could Change Our Lives* 21 (Feb. 2017), https://op.europa.eu/en/publication-detail/-/publication/9964fbfd-6141-11e7-8dc1-01aa75ed71a1 (accessed 15 Oct. 2019).

the administrator or controller under certain circumstances and in a traceable way. Although this may call into question the reliability and security of this technology if the information is no longer unalterable, such technological advances may enable controllers to fulfil their data governance obligations.

It could also be asked whether the processing of data using blockchain technology can collide with the principle of data minimization within the meaning of Article 5 of the GDPR. Indeed, the blockchain technology is based on a distributed database in which tens of thousands of nodes participating in the network keep a copy of it. In fact, as we discussed earlier, full nodes keep a copy of the entire chain of blocks to ensure that blockchain remains immutable and decentralized.

Would this massive distribution or reproduction of data on the network be contrary to the principle of minimization? In our opinion, not necessarily. First, Recital 39 of the GDPR clarifies that personal data should only be processed if the purpose of the processing could not reasonably be achieved by other means. In relation to this, the distribution of data is essential for the functioning of the blockchain network because of its very nature. However, technical security measures such as encryption and the use of the hash function, as discussed below, help to mitigate risks to data subjects by reducing the likelihood of re-identification, thereby minimizing the data that are replicated and processed on the network. Moreover, in order to correctly interpret the principle of data minimization, both the content and nature of the data must be taken into account. Although, in terms of quantity, data are replicated in multiple nodes, the quality of the data processed would be, in our opinion, the determining element for the principle of minimization.[37] Bearing in mind that the reproduction of data on the blockchain network would be limited to those that are necessary for the purpose of the processing for which they are collected, in our opinion, the principle of minimization should not necessarily be affected.

### 3.3 *Rights of the Data Subjects*

The data protection principles outlined above are put in practice through a series of rights for data subjects and obligations for data controllers in accordance with the GDPR. Consequently, the difficulties encountered in the application of the data protection principles, as discussed above, are equally present in the exercise of the rights of data subjects in the blockchain networks.

First, the GDPR requires the controller, at the time when personal data are obtained, to provide the data subjects with certain information on the processing, such as the identity and contact details of the controller, where it takes place or with whom such information is shared.[38] Furthermore, the GDPR sets out the right of access by data subjects,[39] which allows data subjects to require the data controller to confirm whether it processes personal data concerning them and, if so, to grant access to the personal data and provide a copy thereof.[40] In the context of a blockchain, fulfilling this information obligation and exercising a right of access may prove problematic, due to the difficulties associated with the identification of the data controller with whom the data subject may exercise her right, and controllers' ability to meet such data subject access requests.

We nevertheless consider that it is possible to overcome this problem if the data controller is clearly identified and assumes its obligations, as it will then depend solely on the format chosen to record the information and its accessibility by data subjects.[41] Thus, it is not a right incompatible with the blockchain technology per se, since the difficulties encountered are not any different from those that any other data controllers face in other sectors and areas.

Second, difficulties may arise in exercising rights to rectification, where data subjects are entitled to request the correction of inaccurate or incomplete data,[42] as well as their right to erasure or to be forgotten[43] when their data are no longer necessary, or they decide to withdraw their consent, among other circumstances. As mentioned previously, one of the main characteristics of the blockchain technology is that the data, once recorded in the chain, can no longer be altered. In view of the precedents, it seems not to be technically possible to erase or modify that information without altering every other block in the chain and thus the functioning of the decentralized trust on which this technology relies, especially in public networks.

Nevertheless, there is a possibility of overcoming this situation. If the data recorded in the chain are only accessible through a key or if they contain only the status and existence of certain information that is stored outside

**Notes**

[37] European Parliament, Panel for the Future of Science and Technology, *supra* n. 21, at 68.

[38] Articles 13 and 14 of the GDPR.

[39] Article 15 of the GDPR.

[40] Article 15(3) of the GDPR.

[41] Commission Nationale Informatique & Libertés (CNIL), *supra* n. 13.

[42] Article 16 of the GDPR.

[43] Article 17 of the GDPR.

the chain, the controller can make the data practically inaccessible or unusable by destroying the key or the underlying information stored outside the chain. Once the information corresponding to the hash in blockchain has been removed, the hash would become a random number without any link to the personal data which is already deleted.

As a privacy-by-design measure, it could also be possible to preconfigure a Smart Contract in its programming to restrict access to these data under certain circumstances and make them invisible to others, while guaranteeing the right to restriction of processing and erasure. For all this to be possible, it would be necessary not to record data in a legible format without encryption or anonymization technique, as detailed in the following section. Although these are midway solutions, and an analysis on a case by case basis is in any event required, these techniques could make it possible to resemble the effects of data erasure by minimizing the risks to confidentiality, as admitted by the CNIL itself. After all, in our opinion, the GDPR provides certain flexibility in performing the 'erasure', and in the event that the data have been made public, the data controller, taking account of available technology and the cost of implementation, shall take reasonable steps, including technical measures to inform controllers which are processing the personal data that the data subject has requested the erasure by such controllers of any links to, or copy or replication of, those personal data.[44]

Finally, the right not to be subject to a decision based solely on automated processing[45] seeks to protect data subjects against indiscriminate profiling, which may have legal ramifications or similarly significantly affect those individuals. In fact, automated decisions are the basis for the use of Smart Contracts, one of the best-known features of blockchain. To grant the right to request a human intervention, though possible, is to call into question again the reliability of this technology and trust in it. In this regard, the possibility of introducing the option of a subsequent human intervention irrespective of what has already been recorded in the chain, could be considered in order to comply with the provisions of the GDPR.

### 3.4 *Anonymization and Pseudonymization of Personal Data*

The material scope of the GDPR covers the processing of personal data, thus excluding anonymous data. That is to say, the GDPR does not apply to the data that may not be linked, directly or indirectly, to any identified or identifiable natural person.[46]

In order to achieve that purpose, the irreversibility of the anonymization process must be guaranteed. If it is technically possible to retrieve the underlying information or to link the data to a person through patterns, comparison with other databases or to another context (e.g. information contained in other sources), these data are still considered to be (pseudonymized) personal data, and not anonymous data. The provisions contained in the GDPR will therefore apply to such data.

In its Opinion 05/2014 on anonymization techniques,[47] the Article 29 Data Protection Working Party (currently the European Data Protection Board) previously carried out an in-depth analysis of the anonymization processes and adopted a conservative approach. The Working Party concluded that an anonymized dataset can still present residual risk to data subjects, and even when it is no longer possible to precisely retrieve the personal data of a data subject, it may remain possible to collect information about that individual with the help of other sources of information that are available (publicly or not).[48]

Most blockchains use asymmetric cryptography, also called public-key cryptography,[49] which uses a pair of complementary keys called public and private keys. Due to the mathematical relationship between both keys, a cryptogram generated by one of the keys can only be decrypted by the other key. To illustrate the difference between them, we can think of the public key as a bank account number, and the private key as a user's PIN code that must be kept secret. Although a public key cannot be attributed to a data subject without using additional identifiable information, it serves to identify a user. In other words, a public key is pseudonymized personal data,[50] as the natural person is still likely to be identified indirectly in accordance with the GDPR.[51]

*Notes*

[44] Article 17(2) of the GDPR.

[45] Article 22 of the GDPR.

[46] It should be however noted that anonymization of personal data itself constitutes a processing of personal data, which should therefore satisfy the requirement of the purpose limitation principle under the GDPR by having regard to the legal grounds and circumstances of the further processing.

[47] Article 29 Data Protection Working Party, *Opinion 05/2014 on Anonymisation Techniques* (WP 216 10 Apr. 2014).

[48] *Ibid.*, at 23.

[49] Bacon et al., *supra* n. 24, at 62.

[50] Article 4(5) of the GDPR.

[51] Recital 30 of the GDPR states that 'natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This 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'.

Moreover, blockchain networks may include additional personal information. The CNIL considers that the personal data, other than the public keys, would consist of those data contained in the transactions,[52] such as the property deeds registered in the blockchain, including personal data.

As mentioned, the hash function is the common technique on which the blockchain is based in order to compact and pseudonymize the data in the network. The authorities agree that a correct use of these algorithms can contribute significantly to the confidentiality of information, since the hash function is in principle a one-way technique that cannot be reversed (as opposed to encryption) to recover the underlying data. In other words, the transactions are 'transparent', but they are not directly connected to persons or organizations, by protecting the identity and confidentiality of the parties.[53]

However, there is currently a debate as to whether such data pseudonymized by a robust hash function would be considered as personal data within the meaning of the GDPR. The European data protection authorities point out that there are reasons to be cautious. In its Opinion 05/2014, the Working Party concluded that pseudonymization consists of replacing a unique attribute in a record with another, and when used alone will not result in an anonymous dataset.[54]

While risks exist, there are also solutions to mitigate them. The data protection authorities agree that pseudonymization techniques can provide privacy guarantees and may be used to generate efficient anonymization processes, if their application is engineered appropriately. Pseudonymiztion helps reduce the linkability of a dataset with the original identity of a data subject, which therefore ensures a certain level of data protection and reduces the risk for data subjects.[55] The authorities recommend in this respect to use the most advanced pseudonymization techniques, including the use additional characters to obfuscate the data [56] and several layers of encryption[57] to reduce the likelihood of deriving the input value.

In summary, the nature of each pseudonymization technique and DLT should be analysed on a case-by-case basis. It must be determined whether a reasonable effort proportional to the degree of sensitivity of the data has been applied in its pseudonymization process, so that it is technically and economically unfeasible to attack and access the pseudonymized data requiring an excessive or practically impossible effort. This analysis should be done periodically throughout the life cycle of the information, as the power and sophistication of reverse engineering techniques increases.

## 4 Conclusion

We can conclude that, while there are some tensions to be resolved, blockchain technology is not *per se* incompatible with the GDPR. In any event, analysis is needed on a case-by-case basis and a moderate approach adopted that safeguards data protection principles and the rights of data subjects.

Among other precautions, it is required to identify clearly the (joint) data controllers and the applicable lawful basis for processing, respect data protection principles, adopt the privacy-by-design approach, and implement the necessary technical measures to protect the confidentiality of personal data, including state-of-the-art anonymization techniques.

The biggest challenge is undoubtedly to make the use of this technology technically compatible with data subjects exercising their rights, such as the right to rectification and erasure. To this end, it will be necessary to study and apply new techniques to manage the blockchain, but also apply the GDPR in such a way as to enable technological advancements while fully respecting the fundamental rights and freedoms of individuals.

In conclusion, blockchain technology does not pose a threat to the protection of personal data; nor should the GDPR be understood as an obstacle to innovation. The two realities must coexist in balance and harmony.

### Notes

[52] Commission Nationale Informatique & Libertés (CNIL), *supra* n. 13.

[53] European Parliament Research Service, Scientific Foresight Unit (STOA), *supra* n. 36, at 18.

[54] Article 29 Data Protection Working Party, *supra* n. 47, at 20.

[55] Article 32(1)(a) of the GDPR.

[56] This is known as salt and pepper techniques, which consist of adding additional information to the data.

[57] Most networks use what is known as the Merkle tree technique, which is the concatenation structure of blocks of data in multiple layers or levels.