Blockchain and the General Data Protection Regulation

Can distributed ledgers be squared with European data protection law?

STUDY

Panel for the Future of Science and Technology

EPRS | European Parliamentary Research Service

Scientific Foresight Unit (STOA)

PE 634.445 – July 2019

EN

Blockchain and the General Data Protection Regulation

Can distributed ledgers be squared with European data protection law?

Blockchain is a much-discussed instrument that, according to some, promises to inaugurate a new era of data storage and code-execution, which could, in turn, stimulate new business models and markets. The precise impact of the technology is, of course, hard to anticipate with certainty, in particular as many remain sceptical of blockchain's potential impact. In recent times, there has been much discussion in policy circles, academia and the private sector regarding the tension between blockchain and the European Union's General Data Protection Regulation (GDPR). Indeed, many of the points of tension between blockchain and the GDPR are due to two overarching factors.

First, the GDPR is based on an underlying assumption that in relation to each personal data point there is at least one natural or legal person – the data controller – whom data subjects can address to enforce their rights under EU data protection law. These data controllers must comply with the GDPR's obligations. Blockchains, however, are distributed databases that often seek to achieve decentralisation by replacing a unitary actor with many different players. The lack of consensus as to how (joint-) controllership ought to be defined hampers the allocation of responsibility and accountability.

Second, the GDPR is based on the assumption that data can be modified or erased where necessary to comply with legal requirements, such as Articles 16 and 17 GDPR. Blockchains, however, render the unilateral modification of data purposefully onerous in order to ensure data integrity and to increase trust in the network. Furthermore, blockchains underline the challenges of adhering to the requirements of data minimisation and purpose limitation in the current form of the data economy.

This study examines the European data protection framework and applies it to blockchain technologies so as to document these tensions. It also highlights the fact that blockchain may help further some of the GDPR's objectives. Concrete policy options are developed on the basis of this analysis.

AUTHOR

This study was written by Dr Michèle Finck at the request of the Panel for the Future of Science and Technology (STOA) and managed by the Scientific Foresight Unit, within the Directorate-General for Parliamentary Research Services (EPRS) of the Secretariat of the European Parliament.

ADMINISTRATOR RESPONSIBLE

Mihalis Kritikos, Scientific Foresight Unit (STOA)

To contact the publisher, please e-mail stoa@ep.europa.eu

LINGUISTIC VERSION

Original: EN

Manuscript completed in July 2019.

PE 634.445
ISBN: 978-92-846-5044-6
doi: 10.2861/535
QA-02-19-516-EN-N

http://www.europarl.europa.eu/stoa (STOA website)
http://www.eprs.ep.parl.union.eu (intranet)
http://www.europarl.europa.eu/thinktank (internet)
http://epthinktank.eu (blog)

Executive summary

In recent years, there has been ample discussion of blockchain technologies (or distributed ledger technology – DLT[1]) and their potential for the European Union's digital single market. A recurring argument has been that this class of technologies may, by its very nature, be unable to comply with European data protection law, which in turn risks stifling its own development to the detriment of the European digital single market project. The present study analyses the relationship between blockchain and the GDPR, so as to highlight existing tensions and advance possible solutions. It looks into developments up until March 2019.

1. Blockchain technology

In essence, a blockchain is a shared and synchronised digital database that is maintained by a consensus algorithm and stored on multiple nodes (computers that store a local version of the database). Blockchains are designed to achieve resilience through replication, meaning that there are often many parties involved in the maintenance of these databases. Each node stores an integral copy of the database and can independently update the database. In such systems, data is collected, stored and processed in a decentralised manner. Furthermore, blockchains are append-only ledgers to which data can be added but removed only in extraordinary circumstances.

It is important to note that blockchains are a class of technology. Indeed, there is not one version of this technology. Rather, the term refers to many different forms of distributed database that present much variation in their technical and governance arrangements and complexity. This also implies, as will be amply stressed in the analysis below, that the compatibility between distributed ledgers and the GDPR can only be assessed on the basis of a detailed case-by-case analysis that accounts for the specific technical design and governance set-up of the relevant blockchain use case. As a result, this study finds that it cannot be concluded in a generalised fashion that blockchains are either all compatible or incompatible with European data protection law. Rather, each use of the technology must be examined on its own merits to reach such a conclusion. That said, it is easier to design private and permissioned blockchains in a manner that is compatible with EU data protection law than public and permissionless networks. This is because participants in permissioned networks are known to another, allowing for the definition, for example, of contractual relationships that enable an appropriate allocation of responsibility. Furthermore, these networks are, in contrast to public and permissionless networks, designed in a way that enables control over the network, such as to treat data in a compliant manner. Moreover, there is control over which actors have access to the relevant personal data, which is not the case with public and unpermissioned blockchains.

2. The European Union's General Data Protection Regulation

The European Union's General Data Protection Regulation (GDPR) became binding in May 2018. It is based on the 1995 Data Protection Directive. The GDPR's objective is essentially two-fold. On the one hand, it seeks to facilitate the free movement of personal data between the EU's various Member States. On the other hand, it establishes a framework of fundamental rights protection, based on the right to data protection in Article 8 of the Charter of Fundamental Rights. The legal framework creates a number of obligations resting on data controllers, which are the entities determining the means and purposes of data processing. It also allocates a number of rights to data subjects – the natural persons to whom personal data relates – that can be enforced via-à-vis data controllers.

[1] Various definitions of blockchain and distributed ledger technology exist, and some of these stress different technical features of these respective forms of data management. Given the nature of this study and the lack of definitional consensus the terms are used synonymously.

3. The tension between blockchain and the GDPR

In recent years, multiple points of tension between blockchain technologies and the GDPR have been identified. These are examined in detail below. Broadly, it can be argued that these tensions are due to two overarching factors.

First, the GDPR is based on the underlying assumption that in relation to each personal data point there is at least one natural or legal person – the data controller – whom data subjects can address to enforce their rights under EU data protection law. Blockchains, however, often seek to achieve decentralisation in replacing a unitary actor with many different players. This makes the allocation of responsibility and accountability burdensome, particularly in light of the uncertain contours of the notion of (joint)-controllership under the regulation. A further complicating factor in this respect is that in the light of recent case law developments, defining which entities qualify as (joint-) controllers can be fraught with a lack of legal certainty.

Second, the GDPR is based on the assumption that data can be modified or erased where necessary to comply with legal requirements such as Articles 16 and 17 GDPR. Blockchains, however, render such modifications of data purposefully onerous in order to ensure data integrity and to increase trust in the network. Again, the uncertainties pertaining to this area of data protection law are increased by the existing uncertainty in EU data protection law. For instance, it is presently unclear how the notion of 'erasure' in Article 17 GDPR ought to be interpreted.

It will be seen that these tensions play out in many domains. For example, there is an ongoing debate surrounding whether data typically stored on a distributed ledger, such as public keys and transactional data qualify as personal data for the purposes of the GDPR. Specifically, the question is whether personal data that has been encrypted or hashed still qualifies as personal data. Whereas it is often assumed that this is not the case, such data likely does qualify as personal data for GDPR purposes, meaning that European data protection law applies where such data is processed. More broadly, this analysis also highlights the difficulty in determining whether data that was once personal data can be sufficiently 'anonymised' to meet the GDPR threshold of anonymisation.

Another example of the tension between blockchain and the GDPR relates to the overarching principles of data minimisation and purpose limitation. Whereas the GDPR requires that personal data that is processed be kept to a minimum and only processed for purposes that have been specified in advance, these principles can be hard to apply to blockchain technologies. Distributed ledgers are append-only databases that continuously grow as new data is added. In addition, such data is replicated on many different computers. Both aspects are problematic from the perspective of the data minimisation principle. It is moreover unclear how the 'purpose' of personal data processing ought to be applied in the blockchain context, specifically whether this only includes the initial transaction or whether it also encompasses the continued processing of personal data (such as its storage and its usage for consensus) once it has been put on-chain.

It is the tension between the right to erasure (the 'right to be forgotten') and blockchains that has probably been discussed most in recent years. Indeed, blockchains are usually deliberately designed to render the (unilateral) modification of data difficult or impossible. This, of course, is hard to reconcile with the GDPR's requirements that personal data must be amended (under Article 16 GDPR) and erased (under Article 17 GDPR) in specific circumstances.

These and additional points of tension between the GDPR and blockchain are examined in detail below. This analysis leads to two overarching conclusions. First, that the very technical specificities and governance design of blockchain use cases can be hard to reconcile with the GDPR. Therefore, blockchain architects need to be aware of this from the outset and make sure that they design their respective use cases in a manner that allows compliance with European data protection law. Second, it will however also be stressed that the current lack of legal certainty as to how blockchains can be designed in a manner that is compliant with the regulation is not just due to the specific features of

this technology. Rather, examining this technology through the lens of the GDPR also highlights significant conceptual uncertainties in relation to the regulation that are of a relevance that significantly exceeds the specific blockchain context. Indeed, the analysis below will show that the lack of legal certainty pertaining to numerous concepts of the GDPR makes it hard to determine how the latter should apply both to this technology and to others.

In order to reach this conclusion, this report evaluates those aspects of European data protection law that have to date proven to be the most relevant in relation to blockchain. This includes the regulation's territorial and material scope, the definition of responsibility though a determination of which actors may qualify as data controllers, the application of the core principles of personal data processing to blockchains, the implementation of data subject rights in such networks, international data transfers and the possible need for data protection impact assessments.

Whereas much of the debate has focused on the tensions between blockchains and European data protection law, the former may also provide means to comply with the objectives of the latter.

4. Blockchain as a means to achieve GDPR objectives

It has been argued that blockchain technologies might be a suitable tool to achieve some of the GDPR's underlying objectives. Indeed, blockchain technologies are a data governance tool that could support alternative forms of data management and distribution and provide benefits compared with other contemporary solutions. Blockchains can be designed to enable data-sharing without the need for a central trusted intermediary, they offer transparency as to who has accessed data, and blockchain-based smart contracts can moreover automate the sharing of data, hence also reducing transaction costs. Furthermore, blockchains' crypto-economic incentive structures might have the potential to influence the current economics behind data-sharing.

These features may benefit the contemporary data economy more widely, such as where they serve to support data marketplaces by facilitating the inter-institutional sharing of data, which may in turn support the development of artificial intelligence in the European Union. These same features may, however, also be relied upon to support some of the GDPR's objectives, such as to provide data subjects with more control over the personal data that directly or indirectly relates to them. This rationale can also be observed on the basis of data subject rights, such as the right of access (Article 15 GDPR) or the right to data portability (Article 20 GDPR), that provide data subjects with control over what others do with their personal data and what they can do with that personal data themselves.

The analysis below surveys a number of ongoing pilot projects that seek to make this a reality. The ideas behind these projects might be helpful in ensuring compliance with the right to access to personal data that data subjects benefit from in accordance with Article 15 GDPR. Furthermore, DLT could support control over personal data in allowing them to monitor respect for the purpose limitation principle. In the same spirit, the technology could be used to help with the detection of data breaches and fraud.

5. Policy options

This study has highlighted that, on the one hand, there is a significant tension between the very nature of blockchain technologies and the overall structure of data protection law. It has also been stressed that the relationship between the technology and the legal framework cannot be determined in a general manner but must rather be determined on a case-by-case basis. On the other hand, it has also been highlighted that this class of technologies could offer distinct advantages that might help to achieve some of the GDPR's objectives. It is on the basis of the preceding analysis that this section develops concrete policy recommendations.

Policy option 1 – regulatory guidance

The key point highlighted in the first and main part of the present study is that there is currently a lack of legal certainty as to how various elements of European data protection law ought to be applied in the blockchain context. This is due to two overarching factors. First, it has been seen that, very often, the very technical structure of blockchain technology as well as its governance arrangements stand in contrast with the requirements of the GDPR. Second, an attempt to map the regulation to blockchain technologies reveals broader uncertainties regarding the interpretation and application of this legal framework.

Indeed, almost one year after the GDPR became binding and although the legal regime is largely based on the previous 1995 Data Protection Directive, it is evident that many pivotal concepts remain unclear. For instance, it has been seen above that central concepts such as that of anonymisation or that of (joint-) data controllers remain unsettled. Very often the interpretation of these concepts is moreover burdened by a lack of agreement on interpretation between the various supervisory authorities in the European Union.

Furthermore, this study has observed that blockchain technologies challenge core assumptions of European data protection law, such as that of data minimisation and purpose limitation. At the same time, however, this is a broader phenomenon as these principles are just as difficult to map to other expressions of the contemporary data economy such as big data analytics. Nonetheless, the study recommends that it has not become necessary to revise the GDPR. It will be seen that the regulation is an expression of principles-based regulation that was designed to be technologically-neutral and stand the test of time in a fast-moving data-economy. What is needed to increase legal certainty for those wanting to use blockchain technologies is regulatory guidance regarding how specific concepts ought to be applied where these mechanisms are used.

These elements illustrate that regulatory guidance could provide much legal certainty compared to the current status quo. This could take the form of various regulatory initiatives. On the one hand, supervisory authorities could coordinate action with the European Data Protection Board to draft specific guidance on the application of the GDPR to blockchain technologies. On the other, the updating of some of the opinions of the Article 29 Working Party that have not been endorsed by the EDPD, such as the one on anonymisation techniques, could be helpful to provide further legal certainty for the blockchain industry and beyond.

Such initiatives could achieve a dual objective. On the one hand, regulatory guidance could offer additional certainty to actors in the blockchain space who have long stressed that the difficulty of designing compliant blockchain use cases relates in part to the lack of legal certainty as to what exactly is required to design a compliant product. On the other hand, regulatory guidance on how the GDPR is applied to blockchains, and on specific elements of the GDPR that have generated uncertainties in their application more generally, such as anonymisation, could bring more certainty and transparency to the wider data economy, not just to the specific blockchain context.

Policy option 2 – support codes of conduct and certification mechanisms

As a technologically-neutral legal framework, the GDPR was designed in such a way as to enable its application to any technology. This design presents many advantages, not least being that it is supposed to stand the test of time and that it does not discriminate between particular technologies or use cases thereof. Indeed, as an example of principles-based regulation, the regulation devises a number of general overarching principles that must then be applied to the specificities of concrete personal data processing operations.

The technology-neutrality of the GDPR however also means that it can at times be difficult to apply it to specific cases of personal data processing, as evidenced by the analysis above. The regulation itself provides mechanisms designed to deal with this. Indeed, both certification mechanisms and codes of conduct are tools specifically mentioned by the GDPR that are aimed at helping to apply the GDPR's overarching principles to concrete contexts where personal data is processed.

Both certification mechanisms and codes of conduct exemplify a co-regulatory spirit whereby regulators and the private sector devise principles designed to ensure that the principles of European data protection law are upheld where personal data is processed. This has, for instance, been achieved in relation to cloud computing, where many of the difficult questions examined above have also arisen.

Policy option 3 – research funding

Regulatory guidance as well as codes of conduct and certification mechanisms could add much legal certainty regarding how the specific provisions of the GDPR ought to be applied in relation to blockchain technologies.

This, however, will not always be sufficient to enable the compliance of a specific distributed ledger use case with the GDPR. Indeed, as it has been amply underlined in the above analysis, in some cases there are technical limitations to compliance, such as for instance when it comes to the requirement to 'erase' data where a data subject exercises their rights under Article 17 GDPR, In other cases, the current governance design of blockchain use cases is not designed to enable compliance as it does not enable the coordination of multiple actors, who could be joint-controllers, to comply with specific legal requirements. Solutions could be found by means of interdisciplinary research, devising both technical and governance remedies and experiments with blockchain protocols that could be compliant by design.

Table of contents

1. Introduction

Blockchain technologies are a much-discussed instrument that, according to some, promises to inaugurate a new era of data storage and code-execution, which could in turn stimulate new business models and markets. The precise impact of the technology is, of course, hard to anticipate with certainty, and many remain deeply sceptical of blockchains' eventual impact. In recent times, many have voiced concerns that existing regulatory paradigms risk stifling the technology's future development and accordingly stand in the way of transforming the European Union into a global leader in blockchain technology and related developments at a time where there are already broader concerns regarding the EU's ability to keep up with the data-driven economy.

In particular the EU General Data Protection Regulation ('GDPR') is a much-discussed topic in this regard. Indeed, many points of tension between blockchain technologies and the GDPR can be identified. Broadly, it can be maintained that these are due to two overarching factors. First, the GDPR is based on the underlying assumption that in relation to each personal data point there is at least one natural or legal person – the data controller – that data subjects can address to enforce their rights under EU data protection law. Blockchains, however, often seek to achieve decentralisation in replacing a unitary actor with many different players. This makes the allocation of responsibility and accountability burdensome, particularly in light of the uncertain contours of the notion of (joint)-controllership under the Regulation. A further complicating factor in this respect is that in light of recent developments in the case law, defining which entities qualify as (joint-) controllers can be fraught with uncertainty. Second, the GDPR is based on the assumption that data can be modified or erased where necessary to comply with legal requirements such as Articles 16 and 17 GDPR. Blockchains, however, render such modifications of data purposefully onerous in order to ensure data integrity and increase trust in the network. Determining whether distributed ledger technology may nonetheless be able to comply with Article 17 GDPR is burdened by the uncertain definition of 'erasure' in Article 17 GDPR as will be seen in further detail below.

These factors have triggered a debate about whether the GDPR stands in the way of an innovative EU-based blockchain ecosystem. Indeed, some have argued that in order to facilitate innovation and to strengthen the Digital Single Market, a revision of the GDPR may be in order, or that blockchains should benefit from an altogether exemption of the EU data protection framework. Others have stressed the primacy of the legal framework and stated that if blockchains are unable to comply with EU data protection law then this means that they are probably an undesirable innovation considering their inability to comply with established public policy objectives.[2]

These debates have not gone unnoticed to the European Parliament. A recent European Parliament report by the Committee on International Trade highlighted the 'challenge posed by the relationship between blockchain and the implementation of the GDPR'.[3] A 2018 European Parliament resolution underlined that blockchain-based applications must be compatible with the GDPR, and that the Commission and European Data Protection Supervisor should provide further clarification on this matter.[4] Recently, the European Data Protection Board ('EDPB') indicated that blockchain may be one of the topics that it may examine in the context of its 2019/2020 work

[2] Meyer D (27 February 2018), *Blockchain technology is on a collision course with EU privacy law* <https://iapp.org/news/a/blockchain-technology-is-on-a-collision-course-with-eu-privacy-law/>.

[3] European Parliament Report on Blockchain: a Forward-Looking Trade Policy (AB-0407/2018) (27 November 2018), para 14, http://www.europarl.europa.eu/doceo/document/A-8-2018-0407_EN.html .

[4] Proposition de Résolution déposée à la suite de la question avec demande de réponse orale B8-0405/2018 (24 September 2018), para 33, http://www.europarl.europa.eu/doceo/document/B-8-2018-0397_FR.html

programme.[5] The present study seeks to contribute to these on-going reflections in providing a detailed analysis of the GDPR's application to blockchain technologies.

As a starting point, it must be noted that blockchains are in reality a class of technologies with disparate technical features and governance arrangements. This implies that it is not possible to assess the compatibility between 'the blockchain' and EU data protection law. The approach adopted in this study is accordingly to map various areas of the GDPR to the features generally shared by this class of technologies, and to draw attention to how nuances in blockchains' configuration may affect their ability to comply with related legal requirements. Indeed, the key takeaway from this study should be that it is impossible to state that blockchains are, as a whole, either completely compliant or incompliant with the GDPR. Rather, while numerous important points of tension need to be highlighted, ultimately each concrete use case needs to be examined on the basis of a detailed case-by-case analysis.

The second key element highlighted in this study is that whereas there certainly is a certain tension between many key features of blockchain technologies setup and some elements of European data protection law, many of the related uncertainties should not only be traced back to the specific features of DLT. Rather, examining this technology through the lens of the GDPR also highlights significant conceptual uncertainties in relation to the Regulation that are of a relevance that significantly exceeds the specific blockchain context. Indeed, the below analysis will highlight that the lack of legal certainty pertaining to numerous concepts of the GDPR makes it hard to determine how the latter should apply to this technology, but also others. This is, for instance, the case regarding the concept of anonymous data, the definition of the data controller, and the meaning of 'erasure' under Article 17 GDPR. A further clarification of these concepts would be important to create more legal certainty for those wishing to use DLT, but also beyond and thus also to strengthen the European data economy through increased legal certainty.

This study proceeds in three parts. Part One will provide a detailed examination of the application of European data protection to blockchain technologies. Part Two explores whether blockchains may be able to support GDPR compliance, in particular in relation to data governance as well as the prevention and detection of data breaches and fraud. Part Three subsequently seeks to identify a number of policy options available to the European Parliament that would ensure that innovation is not stifled and remains responsible. It will also be specifically assessed whether there is a need for a revision of existing supranational legislation to achieve that objective. This question will be answered negatively. Before moving on to these various elements, a cursory overview of blockchain technology, focusing on the most important elements of the technology from a data protection perspective, is in order.

[5] European Data Protection Board (12 February 2019) *EDPB Work Program 2019/2020* 3 <https://edpb.europa.eu/sites/edpb/files/files/file1/edpb-2019-02-12plen-2.1edpb_work_program_en.pdf>

1.1. Blockchain technology

Any overview of blockchain technology must commence with the observation that there is not one 'blockchain technology'.[6] Rather, blockchains (or Distributed Ledger Technology – 'DLT'[7]) are better seen as **a class of technologies** operating on a spectrum that present different technical and governance structures. This is of pivotal importance as these divergent characteristics ought to be taken into account when determining the compliance of a specific use case with the GDPR. As a consequence, the compliance of a specific use case of the technology and the law must ultimately be determined on a **case-by-case basis**. It should further be stressed that rather than being a completely novel technology, DLT is better understood as an inventive combination of existing mechanisms. Indeed, nearly all of its technical components originated in academic research from the 1980s and 1990s.[8]

In general, it can be said that a blockchain is a **shared and synchronised digital database** that is maintained by a consensus algorithm and stored on multiple nodes (the computers that store a local version of the distributed ledger). Blockchains can be imagined as a peer-to-peer network, with the nodes serving as the different peers.[9] Some blockchains count both full and lightweight nodes whereby only full nodes store an integral copy of the ledger. Other nodes may only store those parts of the ledger of relevance to them.

As its etymology reveals, a blockchain is often structured as a **chain of blocks**.[10] A single block groups together multiple transactions and is then added to the existing chain of blocks through a hashing process. A hash function (or 'hash') provides a unique fingerprint that represents information as a string of characters and numbers. It is a one-way cryptographic function, designed to be impossible to revert.[11] The blocks themselves are made up of different kinds of data, which includes a hash of all transactions contained in the block (its 'fingerprint'), a timestamp, and a hash of the previous block that creates the sequential chain of blocks.[12] As will be seen, some of this data qualifies as personal data for the purposes of the GDPR.

Because blocks are continuously added but never removed a blockchain can be qualified as an **append-only data structure**. Cryptographic hash-chaining makes the log tamper-evident, which increases transparency and accountability.[13] Indeed, because of the hash linking one block to another, changes in one block change the hash of that block, as well as of all subsequent blocks. It is because of DLT's append-only nature that the modification and erasure of data that is required by the GDPR under some circumstances cannot straightforwardly be implemented.

Blockchain networks achieve **resilience through replication**. The ledger's data is resilient as it is simultaneously stored on many nodes so that even if one or several nodes fail, the data goes unaffected. Such replication achieves that there is no central point of failure or attack at the

[6] The technology was first described – although not yet labelled as 'blockchain' in Nakamoto S (2009), *Bitcoin: A Peer-to-Peer Electronic Cash System* https://bitcoin.org/bitcoin.pdf. Satoshi Nakamoto is/are the pseudonymous inventor(s) of Bitcoin.

[7] Various definitions of blockchains and Distributed Ledger Technology exist, and some of these stress different technical features of these respective forms of data management. Given the nature of this study and the lack of definitional consensus I will use both terminologies as synonyms.

[8] Narayanan, A and Clark J (2017) 'Bitcoin's academic pedigree' 60 *Communications of the ACM* 36.

[9] A 'peer' of course does not have to be a private individual but can also be a corporation.

[10] It is worth noting that as the technology evolves this structure might eventually cede way to other forms of data-storage.

[11] Has functions are introduced in further detail below.

[12] Antonopoulos A (2017), Mastering Bitcoin, O'Reilly, xxiii.

[13] Felten E (26 February 2018) *Blockchain: What is it good for?* <https://freedom-to-tinker.com/2018/02/26/bloc>.

hardware level.[14] The replicated data stored in blocks is synchronised through a **consensus protocol**, which enables the distributed network to agree on the current state of the ledger in the absence of a centralised point of control. This protocol determines how new blocks are added to the existing ledger. Through this process, data is chronologically ordered in a manner that makes it difficult to alter data without altering subsequent blocks.

Blockchains are both a new technology for data storage as well as a novel variant of programmable platform that enables new applications such as smart contracts.[15] It is indeed crucial to note that a blockchain ecosystem is multilayered. First, blockchains themselves rely on the Internet and TCP/IP to operate. Second, distributed ledgers provide an **infrastructure for data management that either directly stores data or links to data**. They can serve as an accounting system shared between many actors that can be used by different entities to standardize and link data and 'enable credible accounting of digital events'.[16] DLT can accordingly **coordinate information between many stakeholders** such as to track and store evidence about transactions and participants in that network in a decentralised fashion.

While blockchains only ever store **data**, this data can be taken to represent anything we believe and agree it represents. Bitcoin is essentially data that is valuable because people have come to believe it is. Similarly, over time other forms of digital assets have emerged that are still nothing but raw data taken to represent a good, service or entitlement. Blockchain-based assets can purely have on-chain value (as in Bitcoin) or be the avatar of a real-world asset, whether a good (such as a token representing a bike), a service (such as a voucher for a haircut) or an entitlement (such as a legal right). Seen from this perspective, distributed ledgers have the potential to disrupt the online circulation of value.[17] A 2018 European Parliament study moreover anticipates that '[b]y 2035, tax reporting, e-identity databases, voting schemes, may run on blockchain or another form of Distributed Ledger Technology'.[18]

Blockchains provide thus at once a replicated database that is updated in a decentralised manner (which can be used independently to record transactions in cryptoassets or register information) but also an infrastructure for the **decentralised execution of software**. Examples include the so-called smart contracts or 'decentralised applications' (applications that reflect the decentralised structure of the underlying network).[19] These applications can take a wide variety of forms and serve a wide variety of use cases.[20] This multi-layered nature must be borne in mind whenever compliance of a given blockchain use case with the GDPR is assessed as there may for instance be different data controllers situated at these various layers.

It must be emphasised that there is a **large variety of blockchains**. There is indeed immense variance in blockchains' technical and functional configuration as well as their internal governance structures.[21] DLT is accordingly not a singular technology with a predefined set of characteristics

[14] This does not necessarily entail that there are no central points of attack or failure at the level of software governance.

[15] A smart contract essentially is self-executing software code. I examine smart contracts in further depth just below.

[16] Matzutt R et al (26 February 2018) *A Quantitative Analysis of the Impact of Arbitrary Blockchain Content on Bitcoin* https://fc18.ifca.ai/preproceedings/6.pdf 1.

[17] Cortese A (10 February 2016) *Blockchain Technology Ushers in "The Internet of Value"* https://newsroom.cisco.com/feature-content?articleId=1741667.

[18] European Parliament (November 2018) 'Global Trends to 2035 – Economy and Society' PE 627.126.

[19] This terminology reflects, on the one hand, that these are applications running on a decentralised infrastructure and that they can be managed in a decentralised fashion just as the infrastructure itself.

[20] In addition, there can also be intermediary layers such as decentralised application frameworks that implement their own protocols for the creation and maintenance of decentralised applications

[21] Blockchain governance refers to the process of maintaining the software.

but rather 'a class of technologies'.[22] There is pronounced diversity regarding software management, the visibility and identifiability of transactions on the ledger and the right to add new data to a ledger. Conventionally, DLT is often grouped in two categories of 'public and permissionless' and 'private and permissioned'.

In **public and permissionless** blockchains, anyone can entertain a node by downloading and running the relevant software – no permission is needed. In such an unpermissioned system, there are no identity restrictions for participation.[23] Transparency is moreover an important feature of these systems as anyone can download the entire ledger and view transaction data (which is why they are referred to as 'public' blockchains). For example, any interested party can create a Bitcoin or Ethereum (both are permissionless systems) account using public-private key cryptography without the need for prior permission from a gatekeeper. Permissionless blockchains rely on open source software that anyone can download to participate in the network. Blockexplorers are a form of a search engine that moreover make such blockchain data searchable to anyone. The public auditability of these ledgers enhances transparency but minimizes privacy.

Private and permissioned blockchains run on a private network such as intranet or a VPN and an administrator needs to grant permission to actors wanting to maintain a node. The key distinction between permissioned and unpermissioned blockchains is indeed that while one needs access permission to join the former, this is not necessary in respect of the latter. Whereas unpermissioned blockchains are often a general-purpose infrastructure, permissioned ledgers are frequently designed for a specific purpose. These systems are not open for anyone to join and see. Rather a single party or a consortium acts as the gatekeeper. Permissioned blockchains can be internal to a specific company or joint venture (which is why they are also often referred to as 'private' or 'enterprise' blockchains). While public and permissionless blockchains are pseudonymous networks, in permissioned systems parties' identity is usually known – at least to the gatekeeper granting permission to join the network.

Blockchains' tamper-evident nature constitutes a particularly challenging feature from a data protection perspective. It is often stated that distributed ledgers are 'immutable'. This is misleading as the data contained in such networks can indeed be manipulated in extraordinary circumstances.[24] Indeed, various participants can collude to change the current state of the ledger. While such efforts would be extremely burdensome and expensive, they are not impossible.[25] As per the Bitcoin White Paper there is an 'ongoing chain of hash-based proof-of-work, forming a record that cannot be changed *without redoing the proof-of-work*'.[26] Nonetheless, DLT is tamper-evident and making changes to a ledger can be extremely burdensome. Indeed, there are 'no technical means, short of undermining the integrity of the entire system, to unwind a transfer'.[27] Because blocks are linked through hashes, changing information on a blockchain is difficult and expensive. Making changes to blockchain data is thus extremely hard, and where it is done it is likely visible to all those having access to the ledger.

---

[22] Beck R, Müller-Bloch C and King J (2018) *Governance in the Blockchain Economy: A Framework and Research Agenda* https://www.researchgate.net/publication/323689461_Governance_in_the_Blockchain_Economy_A_Framework_and_Research_Agenda 3.

[23] This is true at least in theory as over time informal restrictions for participation in mining (of an economic nature) and software governance have emerged.

[24] Conte de Leon D et al (2017), 'Blockchain: Properties and Misconceptions' 11 *Asia Pacific Journal of Innovation and Entrepreneurship* 286, 290.

[25] Walch A (2017), 'The Path of the Blockchain Lexicon (and the Law)' 36 *Review of Banking and Financial Law* 713.

[26] Nakamoto S (2009), *Bitcoin: A Peer-to-Peer Electronic Cash System* (2009) https://bitcoin.org/bitcoin.pdf 1 (my own emphasis).

[27] Werbach K and Cornell N (2017), 'Contracts Ex Machina' 67 *Duke Law Journal* 313, 335.

Blockchains' tamper-proof nature is challenging from a legal perspective. As a general matter, this is likely to generate problems as DLT freezes facts (information entered can as a general rule not be changed) and the future (smart contracts' execution cannot be halted even where parties change their mind). Blockchains are thus set up in a manner that may burden compliance with the law for they are often not in a position to absorb changes required by law (such as a change in token ownership mandated by a court order). This is of course also problematic from a GDPR perspective as will be illustrated in further detail below.

Blockchains' nature as a **general-purpose technology** that can be used for both data storage and the execution of computer code explains that various actors are currently experimenting with this technology to achieve different objectives in manifold contexts. In the private sector, DLT has been experimented with to enable various forms of digital money[28]; mobile banking[29]; tracking goods in international trade[30]; manage software licenses;[31] power machine-to-machine electricity markets[32] and replace centralised sharing economy platforms[33] among many others. The public sector equally trials the technology. The European Union is currently exploring the option of a supranational blockchain infrastructure[34] while a UK report suggested using the technology to protect critical infrastructure against cyberattacks; for operational and budgetary transparency and traceability; and to reduce tax fraud.[35] Such variegated applications are possible because blockchains are simultaneously a programmable platform that enables new applications as well as a method for data storage (essentially an accounting system).

Despite avid experimentation and projections of the technology's disruptive nature, there are presently little concrete applications thereof and it is **difficult to predict whether, where and in what form blockchain technology will have practical future impact**. At this moment in time blockchains indeed remain immature as they suffer from 'severe technical and procedural limitations'.[36] These shortcomings include most prominently the lacking scalability that would be necessary for wide deployment. Blockchains are inefficient by design as every full node must process every transaction and maintain a copy of its entire state. While this process eliminates the single point of failure and presents security benefits, it lowers throughput and slows down transactions.[37] This problem is only likely to increase as distributed ledgers grow in size. Scalability forms an important concern in an append-only and thus ever-growing database where each new transaction causes the network to grow.

---

[28] Such as Bitcoin.

[29] https://www.bitpesa.co/

[30] https://www.everledger.io/

[31] Blocher W, Hoppen A and Hoppen P (2017) 'Softwarelizenzen auf der Blockchain' 33 *Computer und Recht* 337.

[32] Sikorski J, Haughton J and Kraft M (2017), 'Blockchain technology in the chemical industry: Machine-to-machine electricity market' 195 *Applied Energy* 234.

[33] Huckle S et al (2016), 'Internet of Things, Blockchain and Shared Economy Applications' 98 *Procedia Computer Science* 461.

[34] See further: https://ec.europa.eu/digital-single-market/en/news/study-opportunity-and-feasibility-eu-blockchain-infrastructure.

[35] Government Office for Science (2016) 'Distributed Ledger Technology: Beyond block chain. A Report by the UK Government Chief Scientific Adviser' https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/492972/gs-16-1-distributed-ledger-technology.pdf 14.

[36] Sillaber C and Waltl B (2017), 'Life Cycle of Smart Contracts in Blockchain Ecosystems' 41 *Datenschutz und Datensicherheit* 497.

[37] Kasireddy P (10 December 2017), *Fundamental Challenges with Public Blockchains* https://medium.com/@preethikasireddy/fundamental-challenges-with-public-blockchains-253c800e9428.

After having provided a cursory overview of the variety of form of DLT as well as general characteristics, the GDPR is now introduced to determine its application to various forms of blockchain technology.

1.2. Blockchains and the GDPR

This section first briefly introduces the General Data Protection ('GDPR') and subsequently provides an overview of its application to various variants of Distributed Ledger Technology.

In the European Union, the right to data protection enjoys the status of a **fundamental right**. Article 8 of the Charter of Fundamental Rights provides that everyone has the right to the protection of personal data concerning him or her.[38] As a consequence, personal data 'must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law' under Article 8(2) of the Charter. The Charter furthermore provides that everyone has a right to access personal data relating to them, including a right to have such data rectified.[39] Article 16 TFEU moreover states that the Parliament and the Council shall lay down rules relating to the protection of individuals with regard to the processing of personal data by Union institutions, bodies, offices and agencies, and by the Member States when carrying out activities that fall within the scope of Union law.[40]

The **General Data Protection Regulation**, as the successor of the 1995 Data Protection Directive, establishes a detailed legislative framework that harmonizes data protection across the European Union.[41] It pursues a dual objective. On the one hand, it seeks to **promote fundamental rights** through a high level of rights protection of natural persons. On the other hand, it pursues an **economic aim** in seeking to remove the obstacles to personal data flows between the various Member States to strengthen the Digital Single Market.[42] The GDPR also emphasizes that whereas data protection enjoys the status of a fundamental right it is not an absolute right but must rather be considered in relation to its function in society and be balanced against other fundamental rights in respect of the proportionality principle.[43]

Whereas the compatibility between blockchain technology and the GDPR can only ever be determined on a case-by-case basis that accounts for the respective technical and contextual factors (such as the governance framework), their general relationship is introduced below in view of drawing attention to the interaction of specific elements of the technology and the legal framework.

[38] Article 8(1) of the Charter of Fundamental Rights. Article 16(1) TFEU.

[39] Article 8(2) of the Charter of Fundamental Rights.

[40] Article 16(2) TFEU.

[41] 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, 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.

[42] Article 1(1) GDPR and Recital 10 GDPR.

[43] Recital 4 GDPR.

2. Applying European data protection law to blockchain

This section maps EU data protection law applies to blockchains. Whereas it is important to bear in mind that the compatibility of a specific use case with specific elements of the GDPR always needs to be determined on the basis of a case-by-case analysis, there is room of general observations regarding the interplay between blockchains and the GDPR. First, it is necessary to define the legal framework's territorial scope of application to determine under which circumstances the use of DLT will be subject to EU law.

2.1. Territorial scope

The analysis must commence with an overview of the circumstances under which the GDPR applies to blockchains. This exercise will underline that although the GDPR is an instrument of EU law, its effects do not stop at the European Union's borders.

Article 3 GDPR provides that the GDPR applies to the processing of personal data whenever certain requirements are met. First, where personal data processing occurs 'in the context of the activities of an **establishment of a controller or a processor in the Union**, regardless of whether the processing takes place in the Union or not'.[44] This implies that where a natural or legal person that qualifies as the data controller or data processor under the GDPR is established in the EU and processes personal data (through blockchains or other means), the European data protection framework applies to such processing.[45]

The European Court of Justice (hereafter also referred to as 'the ECJ' or 'the Court') has confirmed that establishment is a question of fact that ought to be determined on the basis of factors such as 'the degree of stability of the arrangements and the effective exercise of activities' which must be 'interpreted in the light of the specific nature of the economic activities and the provision of services concerned'.[46] Indeed, the concept of establishment 'extends to any real and effective activity – even a minimal one – exercised through stable arrangements'.[47] To assess whether a controller or processor is established in the EU it ought to be determined whether the establishment is an 'effective and real exercise of activity through stable arrangements'.[48] This underlines that a **functional approach** ought to trump formal analysis. The GDPR applies even where the actual processing of personal data is not carried 'by' the establishment concerned itself, but only 'in the context of the activities' of the establishment'.[49] In *Google Spain*, the Court indeed embraced a broad take on this concept in deciding that even though Google's office in Spain only engaged in the sale of advertising, this activity was 'inextricably linked' to the activities of the search engine as the latter would not be profitable without the former.[50]

Even where the establishment criterion does not trigger the GDPR's application other factors may still do so. Indeed, the Regulation also applies where the personal data relates to **data subjects that are based in the EU** even where the data controller and data processor are not established in the Union where one of two conditions are met.[51] First, where personal data processing occurs in the

[44] Article 3(1) GDPR.

[45] See further below for the definitions of data controller and data processor under the GDPR and the question of which actors in a blockchain network are likely to qualify as such.

[46] Case C-230/14 *Weltimmo* [2015] EU:C:2015:639, para 28.

[47] Ibid, para 31.

[48] Recital 22 GDPR.

[49] Case C-131/12 *Google Spain* [2014] EU:C:2014:317, para 52.

[50] Ibid.

[51] Article 3(2) GDPR.

context of the 'offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union'.[52] Thus where data controllers or processors established in a third country offer goods or services to data subjects based in the EU, the GDPR applies whether the data subject provides payment or not.[53]

A further scenario capable of triggering the application of EU data protection law is where personal data processing occurs in the context of the **monitoring of behaviour** as far as this behaviour takes place within the Union.[54] As a consequence the GDPR applies where a data controller or processor not established in the EU monitors the behaviour of an individual based in the Union. A further, less common, scenario where the GDPR applies in the absence of a controller or processor's establishment in the European Union is where the processing occurs in a place where Member State law applies by virtue of **public international law**.[55]

This underlines that the GDPR doubtlessly has a **broad territorial scope.** As a consequence, there are manifold instances where personal data processing through blockchains will fall within the ambit of the GDPR's board territorial scope. This is given where the natural or legal person in charge of the specific use case is established in the EU or where a company or a public administration that ordinarily operate out of the EU rely on blockchains to process personal data. Yet, even where this is not the case, personal data processing based on DLT will oftentimes be subject to European data protection requirements, such as where a natural or legal person offers goods or services to data subjects in the EU. This could, for instance, be the case where operators of a blockchain make available their infrastructure (which can be interpreted to constitute a 'service') to individuals in the Union.[56] Where someone based outside of the EU uses blockchain to process personal data in the context of monitoring the behavior of EU-based individuals the Regulation equally applies.

To determine which Data Protection Authority ('DPA') has competence in relation to a specific processing activity rely on DLT, **Article 56 GDPR** provides that it is that 'of the main establishment'.[57] Pursuant to Article 56(2) GDPR, Member States may however derogate from this general rule and determine that their 'supervisory authority shall be competent to handle a complaint lodged with it or a possible infringement of this Regulation, if the subject matter relates only to an establishment in its Member State or substantially affects data subjects only in its Member State'.[58] In relation to private and/or permissioned blockchains, the competent DPA is thus likely that of the main establishment of the data controller, which will usually be the legal person that operates or has contracted access to a specific DLT infrastructure. For public and permissionless projects it can be difficult to determine 'the main establishment' in light of the absence of a single legal entity governing such projects.[59] Existing case law suggests that in such circumstances, a functional approach ought to be adopted to determine where relevant activity for the specific processing in question was carried out.[60]

---

[52] Article 3(2)(a) GDPR.

[53] This would for instance be the case of services that the data subject receives without the need for monetary compensation but where they make behavioural data available to the controller or processor.

[54] Article 3(2)(b) GDPR.

[55] Article 3(3) GDPR.

[56] On blockchainas-as-a-service, see further Singh J and Michels J (2017), Blockchain As a Service: Providers and Trust *Queen Mary School of Law Legal Studies Research Paper* No. 269/17, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3091223.

[57] Article 56 (1) GDPR. Article 4(16) GDPR specifies that for a controller with multiple EU establishments, this should be its 'place of central administration' or, in the absence thereof the place where its 'main processing activities' take place.

[58] Article 56(2) GDPR.

[59] This point is examined in further detail in the section dealing with controllership below.

[60] Case C-131/12 *Google Spain* [2014] EU:C:2014:317.

The above analysis underlined that the GDPR benefits from a broad territorial scope. Whereas the GDPR's application ought to be assessed in relation to each specific project on a case-by-case basis, it is apparent from the above that oftentimes blockchains that are used to process personal data and have some link to the European Union are subject to GDPR requirements. The Regulation however only applies to the processing of personal data is processed, a concept that is introduced below.

2.2. Material scope

Pursuant to **Article 2(1) GDPR**, the Regulation applies 'to the processing of personal data wholly or partly by automated means and to the processing other than by automated means of personal data which form part of a filing system or are intended to form part of a filing system'.[61]

The GDPR accordingly applies to any personal data processing that occurs entirely or in part by automated means as well as personal data processing that is not automated but forms part of, or is intended to form part of, a filing system.[62] Blockchain-enabled data processing qualifies as data processing 'through automated means'. Existing case law moreover underlines that Article 2(1) GDPR's reference to 'the processing of personal data' ought to be defined broadly to secure the full and complete protection of data subjects.

2.2.1. The meaning of 'processing'

Personal data processing is defined as '**any operation or set of operations which is performed on personal data or sets of personal data**'.[63] Any handling of personal data essentially qualifies as processing – a notion that ought to be interpreted broadly under EU data protecting law. Processing includes the collection and recording of personal data but also its simple storage.[64]

In respect of blockchains, this very broad understanding of what counts as data processing implies that the initial addition of personal data to a distributed ledger, its continued storage and any further processing (such as for any form of data analysis but also to reach consensus on the current state of the network) constitutes personal data processing under Article 4(2) GDPR. Indeed, the European Court of Justice affirmed that personal data processing includes 'any operation or set of operations' performed on personal data.[65]

Processing operations are also subject to EU law where they do not necessarily fall within the economic activities connected with the economic freedoms in EU law. Indeed, in the early *Bodil Lindqvist* case, the loading of information on a webpage by a private person that had no nexus to an economic activity was found to be within the scope of the GDPR. At the time, the ECJ stressed that the Data Protection Directive was based on what is now Article 114 TFEU and that recourse to that legal basis for EU secondary legislation does not presuppose 'the existence of an actual link with free movement between Member States in every situation'.[66] Indeed, 'a contrary interpretation could

[61] Article 2(1) GDPR.

[62] In *Jehovan Todistajat*, the ECJ provided a broad interpretation of the terminology of the 'filing system' covers 'a set of personal data collected in the course of door-to-door preaching, consisting of the names and addresses and other information concerning the persons contacted, if those data are structured according to specific criteria which, in practice, enable them to be easily retrieved for subsequent use. In order for such a set of data to fall within that concept, it is not necessary that they include data sheets, specific lists or other search methods'). Case C-25/17 *Jehovan Todistajat* [2018] EU:C:2018:551, para 62.

[63] Article 4(2) GDPR.

[64] Ibid.

[65] Case C-101/01 *Bodil Lindqvist* [2003] EU:C:2003:596, para 25.

[66] Ibid, para 40.

make the limits' of EU data protection law 'particularly unsure and uncertain, which would be contrary to its essential objective of approximating the laws, regulations and administrative provisions of the Member States in order to eliminate obstacles to the functioning of the internal market deriving precisely from disparities between national legislations'.[67] It follows that any processing of personal data (relying on DLT or any other technology) will be subject to European data protection law, and this even where there is no link to the European Treaties' economic freedoms.

There is, however, one important exception to the GDPR's broad material scope. Where personal data processing constitutes a purely private affair it is shielded from the application of the EU data protection regime.

2.2.2. The 'household exemption'

According to **Article 2(2)(c) GDPR**, the Regulation does not apply to the processing of personal data by a natural person that occurs 'in the course of purely personal or household activity'.[68] Accordingly, where the processing of personal data is a purely personal matter, EU law does not intervene. The difficulty resides in drawing a line between what is purely personal and what is not.

Recital 18 clarifies that personal data processing ought to be considered personal or household activity (which is referred to jointly as 'household activity' below) where it has '**no connection to a professional or commercial activity**'. The same recital also lists a number of examples of such activities, including private correspondence and the holding of addresses, but also social networking and 'online activity undertaken within the context of such activities'.[69]

This raises the question of whether some blockchain use cases could fall under the household exemption, as a consequence of which they would be shielded from the GDPR's scope of application. The ***Commission Nationale de l'Informatique et des Libertés*** ('CNIL'), the French Data Protection Authority, as a matter of fact announced in its 2018 guidance on blockchains that where natural persons add personal data to a blockchain in circumstances that bear no link to a commercial or professional activity, these natural persons ought not to be considered data controllers by virtue of the application of the household exemption.[70] The CNIL provided the example of a physical person that buys or sells Bitcoin for their own account as an example of the household exemption's application.[71]

There is, however, reason to doubt whether this reasoning really holds in light of the Court's case law on this matter. The ECJ has emphasised time and time again that the notion of household activity has to be **interpreted strictly**. In its landmark ruling in *Bodil Lindqvist*, it held that the household exception must be interpreted as 'relating only to activities which are carried out in the course of private or family life of individuals, *which is clearly not the case with the processing of personal data consisting in publication on the internet so that those data are made accessible to an indefinite number of people*'.[72]

The Court has thus added an additional criterion to that of Article 2(2)(c) GDPR. Whereas the legislative text only looks at the **nature of the activity** (private or commercial/professional), the ECJ has added a second test relating to the **scope of dissemination of personal data**. It is worth noting

[67] Ibid. See also Joined Cases C-465/00, C-138/01 and C-139/01 *Österreichischer Rundfunk* [2003] EU:C:2003:294, para 41.

[68] Article 2(2)(c) GDPR.

[69] Recital 18 GDPR.

[70] Commission Nationale Informatique et Libertés (September 2018) *Premiers Éléments d'analyse de la CNIL : Blockchain* https://www.cnil.fr/sites/default/files/atoms/files/la_blockchain.pdf 3.

[71] Ibid.

[72] Case C-101/01 *Bodil Lindqvist* [2003] EU:C:2003:596, para 46.

that during the drafting of the GDPR there was a suggestion to clarify that the household exemption only applies 'where it can be reasonably expected that it will be only accessed by a limited number of persons'.[73] Whereas this did not make it into the final text, the preamble's reference to social media networks (where this is generally the case) and the exclusion of commercial or professional activity might be understood as a reaffirmation of this early suggestion in the legislative text.

Accordingly, the household exemption cannot be applied to circumstances where activity is 'carried out in the course of private or family life of individuals' but is at the same time '**made accessible to an indefinite number of people**'.[74] In *Bodil Lindqvist*, the personal data in question had been made available 'to an indefinite number of people' through its publication on the Internet. Where personal data is made available through a public and permissionless blockchain, it is, however, also made accessible to an indefinite number of people. Indeed, anyone can download the software and store a copy of the entire database on their computer. Tools such as Blockexplorers (which can be compared to a browser for the blockchain that enables anyone to monitor blockchain transactions) moreover make information on a public and permissionless blockchain available to even those that do not download the software.[75]

This conclusion seems all the more warranted considering that both subsequent case law as well as regulatory guidance have underlined the importance of a restrictive interpretation of the household exemption. Whereas the GDPR's preamble refers to social networking as an area shielded from its application, the Article 29 Working Party considers that the household exemption only applies where social networking is of a purely personal nature (as opposed to usages of social media for commercial uses such as the promotion of a small business). For this to be the case, users must 'operate within a purely personal sphere, contacting people as part of the management of their personal, family or household affairs'.[76]

The Working Party also stressed the importance of the **scope of dissemination** regarding the application of the household exemption. In social networking, access to postings made by users is typically constrained to a limited number of self-selected contacts. Where a user however acquires 'a high number of third party contacts, some of whom he may not actually know' this could be 'an indication that the household exemption does not apply and therefore that the user would be considered a data controller'.[77] In social networking, this is the case 'when access to a profile is provided to all members' of the network of where 'data is indexable by search engines, access goes beyond the personal or household sphere'.[78]

In more **recent case law**, the ECJ also confirmed its approach in *Bodil Lindqvist*. In 2014, it recalled in *Ryneš* that in order to ensure a high level of protection of data subjects, the household exemption must be 'narrowly construed' – which it considered to be mandated by the word 'purely' in Article 2(2)(c) GDPR.[79] In *Satamedia*, the Court had already affirmed that Article 2(2)(c) GDPR 'must be interpreted as relating only to activities which are carried out in the course of private or family life of individuals'.[80] As a consequence, the exception 'clearly does not apply' to activities the purpose of which is to 'make the data collected accessible to an unrestricted number of people'.[81] The need

[73] See further Edwards L (2018) Law, Policy and the Internet, Oxford: Hart Publishing.

[74] Case C-101/01 *Bodil Lindqvist* [2003] EU:C:2003:596, para 46.

[75] Instead of many, see https://blockexplorer.com/.

[76] Article 29 Working Party, Opinion 5/2009 on online social networking (WP 163) 01189/09/EN, 3.

[77] Ibid, 6.

[78] Ibid, 7.

[79] Case C-212/13 *Ryneš* [2014] EU:C:2014:2428, para 30.

[80] Case C-73/07 *Satamedia* [2008] EU:C:2008:727, para 44, referring to Case C-101/01 *Bodil Lindqvist* [2003] EU:C:2003:596, para 47.

[81] Case C-73/07 *Satamedia* [2008] EU:C:2008:727, para 44.

to restrictively interpret the household exemption has again been affirmed by the Court in early 2019. In *Buivids* it held in relation to a video recording that had been posted on YouTube that 'permitting access to personal data to an indefinite number of people, the processing of personal data at issue in the main proceedings does not come within the context of purely personal or household activities'.[82]

It accordingly appears **questionable whether the household exemption can at all apply to personal data processing through blockchains**. First, reliance on private and/or permissioned databases in general occurs in a context that is commercial or professional and as a consequence falls short of the test set out in Article 2(2)(c) GDPR regarding the nature of the activity (even though the scope of dissemination is controlled where a permissioned blockchain is used). Second, a public and permissionless blockchain may be used for purely private purposes, yet by definition the scope of dissemination of such data cannot be controlled by the data subject.

It is worth noting that even where the household exemption applies, related personal data processing does not entirely fall outside the scope of the GDPR. As per Recital 18, the GDPR applies 'to controllers or processors which provide the means for processing personal data for such personal or household activities'.[83] This entails that where the household exemption applies but there is a joint-controller or a processor, then the GDPR applies to the personal data processing undertaken by the latter.[84] Next, the concept of personal data under the GDPR and its application to DLT is introduced to further determine the scope of the Regulation.

[82] Case C-345/17, *Sergejs Buivids* [2019] EU:C:2019:122, para 43.

[83] Recital 18 GDPR.

[84] Below various actors that qualify as (joint-)controllers or processors in blockchain contexts will be introduced.

3. The definition of personal data

The definition of personal data determines the GDPR's scope of application and is accordingly of paramount importance. The Regulation only applies to data that is considered 'personal' in nature. Notwithstanding, '[w]hat constitutes personal data is one of the central causes of doubt' in the current data protection regime.[85] The difficulty of determining what counts as personal data is anchored in various factors. First, continuing technical developments make it ever easier to identify individuals on the basis of data that may not be personal on its face. Second, the GDPR's broad definition of personal data encompasses ever more data points. Third, much uncertainty pertains to the notions of pseudonymisation and anonymisation in the GDPR; and finally, despite the GDPR's harmonising aim considerable divergences remain in national law and policy that have added confusion to this area of the law.

The Regulation adopts a **binary perspective between personal data and non-personal data** and subjects only the former to its scope of application.[86] Pursuant to Recital 26 GDPR, the Regulation indeed does not apply to anonymous data. In contrast with this binary legal perspective, reality operates on a spectrum between data that is clearly personal, data that is clearly anonymous (an uncontroversial example should be that of climatic data from outer space that does not reveal information about those that collected it) and anything in between.[87]

Today, much economic value is derived from data that is not personal on its face but can be rendered personal if sufficient effort is put in place. The current battlefield in defining personal data relates to 'data which when collected and processed has the *potential* to have an impact on the personal privacy of particular users, perhaps including their economic and emotional wellbeing, from data which definitely does *not* have such potential. Data which originally related to a living person but now claims to be 'anonymised' in some sense – perhaps merely by the substitutions of an identifier for a name – can still be very useful for businesses and very intrusive to personal privacy'.[88] Beyond, there is an ongoing debate as to whether personal data can be manipulated to become anonymous that is of much relevance in contexts where encryption and hashing are used, as is the case for DLT.[89] This section traces the uncertain contours of personal and anonymous data respectively to determine what data that is frequently used in relation to blockchains may qualify as personal data.

Article 4(1) GDPR defines personal data as follows:

> any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors

[85] Edwards L (2018), *Law, Policy and the Internet*, Oxford: Hart Publishing, 84.

[86] Some might object that 'pseudonymous data' was introduced as third category by the GDPR. Below it will be seen that pseudonymization is more adequately seen as a method of data processing rather than a separate category of data in EU data protection law.

[87] Note however, Purtova N (2018) 'The law of everything. Broad concept of personal data and future of EU data protection law' 10 *Law, Innovation and Technology* 40.

[88] Edwards L (2018), *Law, Policy and the Internet*, Oxford: Hart Publishing, 85

[89] See further Finck M and Pallas F, 'Anonymisation Techniques and the GDPR' (draft on file with author).

> specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person[90]

Article 4 (1) GDPR underlines that **personal data is data that directly or indirectly relates to an identified or identifiable natural person**. The reference to an 'identifiable' person underlines that the data subject does not need to be already identified for data to qualify as personal data. The mere possibility of identification is sufficient.[91] The Article 29 Working Party has issued guidance on how the four constituent elements of the test in Article 4 (1) GDPR – 'any information', 'relating to', 'an identified or identifiable' and 'natural person' – ought to be interpreted.[92]

Information is to be construed broadly, and includes both objective information (such as a name or the presence of a given substance in one's blood) but also subjective analysis such as information, opinions and assessments.[93] Note, however, that the ECJ has clarified in the meantime that whereas information contained in the application for a residence permit and data contained in legal analysis qualify as personal data, related legal analysis does not.[94] Information qualified as personal data can include information that is unrelated to one's private life, underlining the distinction between the concepts of data protection and privacy.[95] Personal data can also take any form, whether it is alphabetical or numerical data, videos and pictures.[96] The Court has indeed confirmed that 'the image of a person recorded by a camera' constitutes personal data.[97]

Second, data can be considered to be **'relating to'** a data subject 'when it is *about* that individual'.[98] This obviously includes information that is in an individual's file but can also include vehicle data that reveals information about a given data subject such as a driver or passenger.[99] An individual is considered to be 'identified' or 'identifiable' where it can be 'distinguished' from others.[100] This does not require that the individual's name can be found. According to the Court, identifying individuals 'by name or by other means, for instance by giving their telephone number or information regarding their working conditions, and hobbies, constitutes the processing of personal data'.[101] Personal data is accordingly 'information, by reason of its content, purpose or effect, is linked to a particular person'.[102]

Personal data relates to an **identified or identifiable natural person**. Where data obviously relates to a natural person, as is the case regarding the data subject's full name, the conclusion that such data is personal data appears uncontroversial.[103] Article 4(1) GDPR however also provides the examples of location data or an identifier, as personal data. This underlines that data, such as health

---

[90] Article 4(1) GDPR (my own emphasis).

[91] Below, the required standard of identifiability is examined in detail.

[92] Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, 6.

[93] Ibid.

[94] Joined Cases C-141/12 and C-372/12 *YS v Minister voor Immigratie* [2014] EU:C:2014:2081.

[95] Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, 7.

[96] Ibid.

[97] Case C-345/17, *Sergejs Buivids* EU:C:2019:122, para 31.

[98] Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, 9 (emphasis in original).

[99] Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, 10.

[100] Ibid, 12.

[101] Ibid, 14 and Case C-101/01 *Bodil Lindqvist* [2003] EU:C:2003:596, para 27.

[102] Case C-434/16 *Nowak* [2017] EU:C:2017:994, para 35.

[103] In *Bodil Lindqvist*, the Court held that the term personal data 'undoubtedly covers the name of a person in conjunction with his telephone coordinates or information about his working conditions or hobbies'. Case C-101/01 *Bodil Lindqvist* [2003] EU:C:2003:596, para 24.

data, that does not relate to an identified but identifiable natural person still falls within this scope. Indeed, the concept of personal data ought to be interpreted broadly – as has by now been amply confirmed in relevant case law.

In *Nowak*, the ECJ concluded that examinations from further education institutions are personal data. It explained that the expression 'any information' reflects 'the aim of the EU legislature to assign a wide scope to that concept, which is not restricted to information that is sensitive or private, but potentially encompasses all kinds of information, not only objective but also subjective, in the form of opinions and assessments, provided that it 'relates' to the data subject'.[104] As written answers reflect a candidate's knowledge and competence in a given field and contain his handwriting, they qualified as personal data.[105] The examiner's written comments were considered to be personal data of both the candidate and the examiner.[106]

In *Digital Rights Ireland* the ECJ held that metadata (such as location data or IP addresses) which only allows for the indirect identification of the data subject can also be personal data as it 'may allow very precise conclusions to be drawn concerning the private lives of the persons whose data has been retained, such as the habits of everyday life, permanent or temporary places of residence, daily or other movements, the activities carried out, the social relationships of those persons and social environments frequented by them'.[107]

The broad definition of personal data has led some to observe that data protection law has become the 'law of everything' as in the near future all data may be personal data and thus subject to GDPR requirements.[108] This is so as 'technology is rapidly moving towards perfect identifiability of information; datafication and advances in data analytics make everything (contain) information; and in increasingly 'smart' environments any information is likely to relate to a person in purpose or effect'.[109] The Article 29 Working Party has also warned that 'anonymisation is increasingly difficult to achieve with the advance of modern computer technology and the unbiquitous availability of information'.[110]

Finally, personal data is only data which relates to a **natural person**. As a fundamental rights framework, the GDPR accordingly does not apply to legal persons.[111] Similarly, the Regulation does not apply to data relating to the deceased.[112] This does not however mean that data relating to a deceased person is not personal data of a related data subject, such as a family member.

3.1. Drawing the line between personal and non-personal data

Drawing the line between personal and non-personal data is fraught with uncertainty due to the broad scope of personal data and the technical possibility to infer information about data subjects from datapoints that are ostensibly unrelated to them. This is not only due to the Court's expansive interpretative stance but also to the difficulty of determining whether data that has been

[104] Case C-434/16 *Nowak* [2017] EU:C:2017:994, para 34.

[105] Ibid, para 37.

[106] Case C-434/16 *Nowak* [2017] EU:C:2017:994, para 44.

[107] Cases C-293/12 and C-594/12 *Digital Rights Ireland* [2014] EU:C:2014:238, para 27.

[108] Purtova N (2018) 'The law of everything. Broad concept of personal data and future of EU data protection law' 10 *Law, Innovation and Technology* 40.

[109] Ibid.

[110] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 31.

[111] See further van der Sloot B (2015), 'Do Privacy and Data Protection Rules Apply to Legal Persons and Should They? A Proposal for a Two-Tiered System' 31 *Computer Law and Security Review*.

[112] Recital 27 GDPR.

manipulated to prevent identification can actually be considered as anonymous data for GDPR purposes.[113] In particular, the meaning of pseudonymisation in the Regulation has created uncertainty. This convoluted area of the law is first introduced in a general fashion to set out key principles before it is mapped to blockchains further below.

Article 4(5) GDPR introduces pseudonymisation as the

> processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person[114]

The concept of pseudonymisation is one of the novelties of the GDPR compared to the 1995 Data Protection Directive. At this stage, there is an ongoing debate regarding the implications of Article 4(5) GDPR for EU data protection law. In particular, it is being discussed whether the provision gives rise to the third category of data (in addition to personal and anonymous data) and if so, whether pseudonymous data qualifies as personal data or whether it can meet the anonymisation threshold.

A literal interpretation of this provision however reveals that Article 4(5) GDPR deals with **a method, not an outcome of data processing**.[115] It defines pseudonymisation as the 'processing' of personal data in such a way that data can only be attributed to a data subject with the help of additional information. No precise methods are prescribed, in line with the Regulation's technologically-neutral spirit. This underlines that **pseudonymised data remains personal data**, in line with the Article 29 Working Party's finding that 'pseudonymisation is not a method of anonymisation. It merely reduces the linkability of a dataset with the original identity of a data subject, and is accordingly a useful security measure'.[116] Thus pseudonymous data is still 'explicitly and importantly, personal data, but its processing is seen as presenting less risk to data subjects, and as such is given certain privileges designed to incentivise its use'.[117]

The GDPR indeed explicitly encourages pseudonymisation as a **risk-management measure**. Pseudonymisation can be taken as evidence of compliance with the controller's security obligation under Article 5(f) GDPR and that the data protection by design and by default requirements under Article 25 GDPR have been given due consideration. Recital 28 GDPR further provides that '[t]he application of pseudonymisation to personal data can reduce the risks to the data subjects concerned and help controllers and processors to meet their data-protection obligations'.[118] According to **Recital 29 GDPR**:

> [i]n order to create incentives to apply pseudonymisation when processing personal data, measures of pseudonymisation should, whilst allowing general analysis, be possible within the same controller when that controller has taken technical and organisational measures necessary to ensure, for the processing

[113] Anonymous data is data that has been modified so that it no longer relates to an identified or identifiable natural person. Where anonymisation was effective, the GDPR does not apply.

[114] Article 4(5) GDPR.

[115] See also Mourby M et al (2018), 'Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK' 34 *Computer Law & Security Review* 222, 223.

[116] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 3.

[117] Edwards L (2018) Law, Policy and the Internet, Oxford: Hart Publishing, 88.

[118] Recital 28 GDPR.

> concerned, that this Regulation is implemented, and that additional information for attributing the personal data to a specific data subject is kept separately. The controller processing the personal data should indicate the authorised persons within the same controller[119]

It is crucial to remember that, as per Recital 30, data subjects 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'.[120] Whereas such identifiers are of a pseudonymous character, they may nonetheless enable the indirect identification of a data subject as they 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'.[121] Below, it will be seen that the **public keys** that function as identifiers in blockchains can be qualified as such an identifier and that as such qualify as personal data.

It should be stressed that even though pseudonymised data may fall short of qualifying as anonymised data, it may fall under **Article 11 GDPR**, pursuant to which the controller is not obliged to maintain, acquire or process additional information to identify the data subject in order to comply with the Regulation.[122] In such scenarios, the controller does not need to comply with the data subject rights in Articles 15 to 20 GDPR unless the data subject provides additional information enabling their identification for the purposes of exercising their GDPR rights.[123]

There is thus ample recognition in the text of the GDPR that pseudonimisation is a valuable risk-minimisation approach, but that at the same time it should not be seen as an anonymisation technique. It is in this context important to understand that the *legal* concept of pseudonymisation does not overlap with the common-sense understanding thereof. From a legal perspective, pseudonymous data is always personal data. This raises the question, however, of whether pseudonymisation measures in the computer science understanding of the term can produce anonymous data.[124] Some Data Protection Authorities have considered that pseudonymisation can indeed lead to the generation of anonymous data.[125] The below section examines whether it is possible to transform personal data into anonymous data.

3.1.1. Transforming personal data into anonymous data

There is currently ample uncertainty as to when the line between personal and non-personal data is crossed in practice. The principle that should be used to determine whether data is personal data or not is that of the **reasonable likelihood** of identification, which is enshrined in **Recital 26 GDPR** according to which:

[119] Recital 29 GDPR.

[120] Recital 30 GDPR.

[121] Ibid.

[122] Article 11(1) GDPR.

[123] Article 11(2) GDPR.

[124] Zuiderveen Borgesius F (2016), 'Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation' 32 *Computer Law & Security Review* 256, 258.

[125] Information Commissioner's Office (November 2012), 'Anonymisation: managing data protection risk code of practice' https://ico.org.uk/media/1061/anonymisation-code.pdf 21 ('This does not mean, though, that effective anonymization through pseudonymization becomes impossible').

> [t]he principles of data protection should apply to any information concerning an identified or identifiable natural person. Personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information should be considered to be information on an identifiable natural person. To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments. The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes[126]

Recital 26 GDPR first recalls that pseudonymous data qualifies as personal data in line with Article 4(5) GDPR. Thereafter, it formulates the test that ought to be employed to determine whether data is personal data or not, namely whether the controller or another person are able to identify the data subject in using all the '**means reasonably likely to be used**'.[127] Where personal data is no longer likely to be reasonably 'attributed to a natural person by the use of additional information', it is no longer personal data.[128]

The GDPR is thus clear that, at least as a matter of principle, **it is possible to manipulate personal data in a manner removing the reasonable likelihood of identifying a data subject** through such data. Recital 26 GDPR as a matter of fact explicitly envisages that there can be scenarios where personal data has been 'rendered **anonymous** in such a manner that the data subject is not or no longer identifiable'.[129] Where such an attempt proves successful, personal data has been transformed into anonymous data which evades the Regulation's scope of application.

Essentially, Recital 26 GDPR thus imposes a **risk-based approach** to determine whether data qualifies as personal data. Where there is a reasonable risk of identification, data ought to be treated as personal data and is hence subject to the GDPR. Where the risk is merely negligent (that is to say that identification is not likely through reliance on all the means reasonably likely to be used), it can be treated as anonymous data, even though identification cannot be excluded with absolute certainty.

The relevant criterion to determine whether data is personal data is that of **identifiability**.[130] The GDPR's preamble furthermore provides a list of elements to be taken into account to determine the likelihood of identifiability through all the means reasonably likely to be used. These include 'all objective factors, such as the costs of and the amount of time required for identification, taking into

[126] Emphasis added.
[127] Recital 26 GDPR.
[128] Emphasis added.
[129] Recital 26 GDPR (my own emphasis).
[130] Recital 26 GDPR.

consideration the available technology at the time of the processing and technological developments'.[131]

Over time, **national supervisory authorities and courts** have found that data that was once personal had crossed this threshold to become anonymous data. For example, the UK High Court held in 2011 that data on certain abortions that had been turned into statistical information was anonymous data that could be publicly released.[132] Similarly, the UK Information Commissioner's Office (the British Data Protection Authority, hereafter also referred to as 'ICO') embraced a relativist understanding of Recital 26 GDPR, stressing that the relevant criterion is not that of the possibility of identification but rather of 'the identification or likely identification' of a data subject .[133] This risk-based approach acknowledges that 'the risk of re-identification through data linkage is essentially unpredictable because it can never be assessed with certainty what data is already available or what data may be released in the future'.[134]

Whereas some thus favour a risk-based approach, the Article 29 Working Party leaned towards a **zero-risk approach**. It noted in its 2014 guidelines on anonymisation and pseudonymisation techniques that 'anonymisation results from processing personal data in order to *irreversibly* prevent identification'.[135] Indeed, in its guidance on the matter, the Working Party appears to at once apply the risk-based test inherent in the legislation, whereas at the same time adding its own – stricter – test. This has been the source of much confusion, which is examined in further detail below. It will be seen that these guidelines diverge from the test that is set out in Recital 26 GDPR. These guidelines are examined here as they represent the only available guidance at supranational that is available at this stage. It is, however, worth noting that these guidelines were not part of the Article 29 Working Party's opinions that were endorsed by the EDPB when it took office in 2018.[136] There is accordingly considerable uncertainty regarding the appropriate elements of the GDPR's identifiability test, which are now examined in turn.

3.1.2. The uncertain standard of identifiability

Risk must evidently be assessed on a case-by-case basis as '[n]o one method of identifying an individual is considered 'reasonably likely' to identify individuals in all cases, each set of data must be considered in its own unique set of circumstances'.[137] This raises the question of what standards ought to be adopted to assess the risk of identification in a given scenario.

The Article 29 Working Party announced in its 2014 guidelines on anonymisation and pseudonymisation techniques that 'anonymisation results from processing personal data in order to ***irreversibly* prevent identification**'.[138] This is in line with earlier guidance according to which anonymised data is data 'that previously referred to an identifiable person, but where that

[131] Ibid

[132] See *R (on the application of the Department of Health) v Information Commissioner* [2011] EWHC 1430 (Admin).

[133] Information Commissioner's Office (November 2012) *Anonymisation: managing data protection risk code of practice* https://ico.org.uk/media/1061/anonymisation-code.pdf 16.

[134] Ibid.

[135] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 3 (my own emphasis).

[136] This list is available online: https://edpb.europa.eu/node/89.

[137] Mourby M (2018) et al, 'Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK' 34 *Computer Law & Security Review* 222, 228.

[138] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 3 (my own emphasis).

identification is *no longer possible*'.[139] This in turn has been interpreted to mean that 'the outcome of anonymisation as a technique applied to personal data should be, in the current state of technology, ***as permanent as erasure***, i.e. making it impossible to process personal data'.[140] To the Article 29 Working Party, a simple risk-based approach is accordingly insufficient – it deems that the risk of identification must be zero. At the same time, its guidance also stresses that a residual risk of identification is not a problem if no one is 'reasonably likely' to exploit it.[141] The relevant question to be asked is thus 'whether identification has become 'reasonably' impossible' – as opposed to absolutely impossible.[142] Notwithstanding, this approach has been criticised as 'idealistic and impractical'.[143] In any event, this is an area where there is much confusion regarding the correct application of the law. The irreversible impossibility of identification amounts to a high threshold, especially if one considers that the assessment of data's character ought to be dynamic, accounting not just for present but also future technical developments.

3.2. The criteria of identifiability

According to the Article 29 Working Party, **three different criteria** ought to be considered to determine whether de-identification is 'irreversible' or 'as permanent as erasure' namely whether (i) it is still possible to single out an individual; (ii) it is still possible to link records relating to an individual, and (iii) whether information concerning an individual can still be inferred.[144] Where the answer to these three questions is negative, data can be considered to be anonymous.

Singling out refers to 'the possibility to isolate some or all records which identify an individual in the dataset'.[145] An example would be a dataset containing medical information which enables identification of a specific data subject, for example through a combination of medical information (such as the presence of a rare disease) and additional demographic factors (such as their date of birth). It is worth noting that a reference to singling out has in the meantime been introduced into the text of the GDPR in the form of Recital 26 GDPR.

Linkability denotes the risk generated where at least two data sets contain information about the same data subject. If in such circumstances an 'attacker can establish (e.g. by means of correlation analysis) that two records are assigned to a same group of individuals but cannot single out individuals in this group', then the used technique only provides resistance against singling out but not against linkability.[146] Assessing linkability can be burdened with difficulty as it is hard to establish what other information capable of triggering identification through linkage is available to a controller now or may be in the future.

Finally, **inference** may still be possible even where singling out and linkability are not. Inference has been defined by the Working Party as 'the possibility to deduce, with significant probability, the value of an attribute from the values of a set of other attributes'.[147] For example, where a dataset

[139] Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, 21 (my own emphasis).

[140] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 6 (my own emphasis).

[141] Ibid, 7.

[142] Ibid, 8.

[143] Stalla-Bourdillon, S and Knight, A (2017) 'Anonymous data v. personal data—a false debate: an EU perspective on anonymization, pseudonymization and personal data' *Wisconsin International Law Journal*, 34 (2), 284-322.

[144] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 3.

[145] Ibid, 11.

[146] Ibid, 11.

[147] Ibid, 12.

refers not to Angela Merkel but rather to a female German chancellor in the early 2000s, her identity would nonetheless be possible to reasonably infer.

Transforming personal data in a manner that excludes singling out, linkability and inference in a reasonable manner is difficult. This is confirmed by the Working Party's analysis of the most commonly used 'anonymisation' methods, which lead it to conclude that each of them leaves a residual risk of identification so that, if at all, only a combination of different approaches can de-personalise data.[148] Research has as a matter of fact amply confirmed the difficulties in achieving anonymisation, such as where an 'anonymised' profile can still be used to single out a specific individual.[149] The increasing abundance of data moreover facilitates the de-anonymisation of given data points through the combination of various datasets.[150] It is accordingly **often easy to identify data subjects on the basis of purportedly anonymous data**.[151] Some computer scientists have even warned that the de-identification of personal data is an 'unattainable goal'.[152]

Data Protection Authorities and courts elsewhere have provided somewhat different interpretations. The United Kingdom's Information Commissioner Office focuses on a 'risk-based' approach to identification.[153] This has been subject to criticism.[154] The British DPA has suggested that the test adopted to carry out risk assessment of re-identification should be the '**motivated intruder**' test whereby companies should determine whether an intruder could achieve re-identification if motivated to attempt this.[155] The motivated intruder is assumed to be 'reasonably competent' and with access to resources such as the internet, libraries or all public documents but should not be assumed to have specialist knowledge such as hacking skills or to have access to 'specialist equipment'.[156] In a December 2018 decision, the Austrian Data Protection Authority moreover affirmed that there is no need for anonymisation to be irreversible – at least in instances where anoymisation is used to trigger the 'erasure' of data under Article 17 GDPR.[157] It is, however, unclear whether supervisory authorities across the EU would adhere to this stance. Beyond this lack of legal certainty, technical developments also burden the implementation of the risk-based approach. Establishing the risk of re-identification can for example be difficult 'where complex statistical methods may be used to match various pieces of anonymised data'.[158] Indeed 'the

[148] Ibid, 3.

[149] Instead of many, see Miller A (2014) 'What do we Worry about when we worry about Price Discrimination? The Law and Ethics of Using Personal Information for Pricing' 19 *Journal of Technology Law & Policy* 41.

[150] Ohm P (2010) 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' 57 *UCLA Law Review* 1701, Veale M (2018) et al, 'When data protection by design and data subject rights clash' 8 *International Data Privacy Law* 105, 113.

[151] Sweeney L (2000) 'Simple Demographics Often Identify People Uniquely' *Data Privacy Working Paper* 3 https://dataprivacylab.org/projects/identifiability/paper1.pdf.

[152] Narayanan A and Shmatikov V (2010) 'Myths and Fallacies of Personally Identifiable Information' 53 *Communications of the ACM* 24-26, 26.

[153] Information Commissioner's Office (November 2012) *Anonymisation: managing data protection risk code of practice* https://ico.org.uk/media/1061/anonymisation-code.pdf.

[154] Anderson R (2012) *The Foundation for Information Policy Research: Written evidence to the Information Commissioner on The Draft Anonymisation Code of Practice*, https://www.cl.cam.ac.uk/~rja14/Papers/fipr-ico-anoncop-2012.pdf.

[155] Information Commissioner's Office (November 2012), *Anonymisation: managing data protection risk code of practice* (November 2012), https://ico.org.uk/media/1061/anonymisation-code.pdf 22.

[156] Information Commissioner's Office (November 2012), 'Anonymisation: managing data protection risk code of practice' 21-22 https://ico.org.uk/media/1061/anonymisation-code.pdf 21-22.

[157] Austrian Data Protection Authority (05 December 2018) Decision DSB-D123.270/0009-DSB/2018 https://www.ris.bka.gv.at/Dokumente/Dsk/DSBT_20181205_DSB_D123_270_0009_DSB_2018_00/DSBT_20181205_DSB_D123_270_0009_DSB_2018_00.html.

[158] Information Commissioner's Office (November 2012) Anonymisation: managing data protection risk code of practice https://ico.org.uk/media/1061/anonymisation-code.pdf 21.

possibility of linking several anonymised datasets to the same individual can be a precursor to identification'.[159]

Pseudonymous data on a blockchain can, in principle, be related to an identified or identifiable natural personal through singling out, inference or linkability. To provide an example of the latter, we may imagine a situation whereby two individuals, A and B, have coffee together, and A sees that B purchases her coffee through a cryptocurrency that is based on a public and permissionless blockchain. As this transaction is recorded in the public ledger together with information regarding the amount paid and a timestamp, it can be possible for A (or possibly a third observer such as the cashier) to find this transaction on a blockchain and accordingly gain knowledge of B's pseudonymous public key. Depending on the relevant set-up of this use case and specifically whether a new key is used for each transaction, it may also be possible to trace back all transactions that B has ever made using this cryptocurrency.

An objective or subjective approach?

It is furthermore unclear whether an **objective or subjective approach** needs to be adopted to evaluate the risk of identification. Recital 26 GDPR foresees that, in light of the current state of the art, a 'reasonable' investment of time and financial resources should be considered to determine whether a specified natural person can be identified on the basis of the underlying information. There is, however, an argument to be made that what is a 'reasonable' depends heavily on context. Whereas a case-by-case basis is in any event required, it is not obvious from the text of the GDPR itself what standard of reasonableness ought to be applied, specifically whether this is an objective or subjective criterion.

The dimension of time

Recital 26 GDPR requires that the 'means' taken into account are not just those that are available in this moment in time, but also '**technological developments**'. It is, however, far from obvious what timescale ought to be considered. Recital 26 GDPR does not reveal whether it ought to be interpreted as merely requiring a consideration of technical developments that are ongoing (such as a new technique that has been rolled out across many sectors of the economy but not yet to the specific data controller or processor), or whether developments currently just explored in research should also be given consideration. To provide a concrete example, it is not at all clear whether the still uncertain prospect of quantum computing should be taken into account when determining whether a certain encryption technique used with respect to blockchains could turn personal data into anonymous data.[160]

The Article 29 Working Party has issued guidance on this matter. It indicated that one

> should consider the state of the art in technology at the time of the processing and the possibilities for development during the period for which the data will be processed. Identification may not be possible today with all the means likely reasonably to be used today. If the data are intended to be stored for one month, identification may not be anticipated to be possible during the 'lifetime' of the information, and they should not be considered as personal data. However, if

[159] Ibid.

[160] The Economist (20 October 2018) *Quantum computers will break the encryption that protects the internet* https://www.economist.com/science-and-technology/2018/10/20/quantum-computers-will-break-the-encryption-that-protects-the-internet.

> they are intended to be kept for 10 years, the controller should consider that possibility of identification that may occur also within the ninth year of their lifetime, and which may make them personal data at that moment. The system should be able to adapt to these developments as they happen, and to incorporate the appropriate technical and organisational measures in due course[161]

Blockchains are append-only ledgers from which data cannot easily be deleted once it has been added. There may be blockchain use-cases which only require the ledger to be used for a specified period of time, such as a fiscal year. In this circumstance, technical developments should be evaluated for that time period only. Yet, other blockchain use cases are built on the assumption that the infrastructure will serve as a perpetual record of transactions, meaning that the envisaged time period of usage is indefinite. It is, however, impossible to envisage developments in data processing and analysis until the end of time as arguably anything then becomes possible. The argument may thus be made that where data is added to a blockchain that is designed to be used for a time frame that exceeds reasonable analysis, any data ought to be considered personal data as it cannot be reasonably assumed that identification remains unlikely in the future.

Personal data to whom?

It is at present unclear **from whose perspective the likelihood of identifiability ought to be assessed**. The formulation of Recital 26 GDPR as well as existing case law on the matter are unclear whether identifiability should be assessed only from the perspective of the data controller (a relative approach) or any third party that may be able to identify a data subject (an absolute approach).

The leading case on this matter is *Breyer*. Mr Breyer had accessed several websites of the German federal government that stored information regarding access operations in logfiles.[162] This included the visitor's dynamic IP address, which is an IP address that changes with every new connection to the internet to prevent the linkage through publicly available files between a specific computer and the network used by the ISP. The Court had already decided in *Scarlet Extended* that static IP addresses are personal data as they allow users to be precisely identified.[163] The Court noted the differences between static and dynamic IP addresses as in the former case, the collection and identification of IP addresses was carried out by the ISP, whereas in the case at issue the collection and identification of the IP address was carried out by an online media services provider, which 'registers IP addresses of the users of a website that it makes accessible to the public, without having the additional data necessary in order to identify those users'.[164]

The Court found that while a dynamic IP address is data relating to an 'identifiable' natural person 'where the additional data necessary in order to identify the user of a website that the services provider makes accessible to the public are held by that users' internet service provider'.[165] The dynamic IP address accordingly qualified as personal data even though the data to identify Mr Breyer was not held by German authorities but by the ISP.[166]

[161] Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, 15 (emphasis added).

[162] Case C-582/14 *Breyer* [2016] EU:C:2016:779.

[163] Case C-70/10 *Scarlet Extended* [2011] EU:C:2011:771.

[164] Case C-582/14 *Breyer* [2016] EU:C:2016:779, para 35.

[165] Ibid, para 39.

[166] Ibid, para 49.

In isolation, this would imply that the nature of data ought not just to be evaluated from the perspective of the data controller (German authorities) but also from the perspective of third parties (the ISP). Indeed, the Court stated clearly that '**there is no requirement that all the information enabling the identification of the data subject must be in the hands of one person**'.[167] However, that finding may have been warranted by the specific facts at issue. The Court stressed that whereas it is in principle prohibited under German law for the ISP to transmit such data to website operators, the government has the power to compel the ISP to do so in the event of a cyberattack. As a consequence, it had the means likely reasonably to be used to identify the data subject.[168]

This would indicate that the perspective from which identifiability ought to be assessed is that of the initial data controller. In *Breyer*, Advocate General Campos Sa nchez-Bordona warned that if the contrary perspective were adopted, it would never be possible to rule out with absolute certainty 'that there is no third party in possession of additional data which may be combined with that information and are, therefore, capable of revealing a person's identity'.[169]

It has, however, also been pointed out that even though Articles 2 and 4(1) GDPR are both kept in a passive voice, the wording of Recital 26 GDPR ('by the controller or by another person') could be taken to suggest that third parties' ability to identify a data subject also ought to also be considered.[170] The GDPR is a **fundamental rights framework** and the ECJ has time and time again emphasised the need to provide an interpretation thereof capable of ensuring the complete and effective protection of data subjects. From this perspective, it indeed matters little from whose perspective data qualifies as personal data – anyone should protect the data subject's rights under the Regulation.

As a consequence, there is currently 'a very significant grey area, where a data controller may believe a dataset is anonymised, but a motivated third party will still be able to identify at least some of the individuals from the information released'.[171] Research has moreover pointed out that where a data controller implements strategies that makes it unlikely or at least difficult to re-identify data, it may be 'far from trivial for an adversary to, given that adversaries likely have a high tolerance for inaccuracy and access to many additional, possibly illegal, databases to triangulate individuals with'.[172] On the other hand, adopting an absolutist approach could effectively rule out the existence of anonymous data as ultimately there will always be parties able to combine a dataset with additional information that may re-identify it.

It is worth noting that Article 4(5) GDPR's requirement that where pseudonymisation occurs, the additional information that could enable identification 'is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person'.[173] It however appears that such precautionary measures only apply to the original data controller that pseudonymised the dataset, not necessarily those that may subsequently handle it. These are points of broader relevance also beyond the specific blockchain context.

---

[167] Ibid, para 31.

[168] Ibid, para 47.

[169] Opinion of AG Campos Sa nchez-Bordona in Case C-582/14 *Breyer* [2016] EU:C:2016:779, EU:C:2016:339, para 65.

[170] Buocz T et al (2019), 'Bitcoin and the GDPR: Allocating Responsibility in Distributed Networks' *Computer Law & Security Review* 1, 9.

[171] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 31.

[172] Veale M et al (2018), 'When data protection by design and data subject rights clash' 8 *International Data Privacy Law* 105, 107.

[173] Article 4(5) GDPR.

The purposes of data use

Finally, when determining the nature of personal data, it is crucial to evaluate the 'purpose pursued by the data controller in the data processing'.[174] Indeed, 'to argue that individuals are not identifiable, where the purpose of processing is precisely to identify them, would be a sheer contradiction in terms'.[175] This must also be remembered whenever a data processing operation that involves the use of blockchain is tested for its compatibility with the GDPR. Indeed, where certain data that is used serves precisely to identify an individual, it cannot be concluded that such data is not personal data. For example, the French CNIL held that the accumulation of data held by Google that enables it to individualise persons is personal data as 'the sole objective pursued by the company is to gather a maximum of details about individualised persons in an effort to boost the value of their profiles for advertising purposes'.[176] Thus, where the public key serves precisely to **identify a natural person**, the conclusion that it qualifies personal data appears unavoidable.

It is hence plain that there is currently much uncertainty regarding the dividing line between personal and non-personal data under the GDPR. This affects the development of blockchain use cases but is also a broader issue. It is for this reason that this study recommends the adoption of regulatory guidance on this matter, as will be seen further below.

After having introduced the general uncertainties regarding the taxonomy of personal, pseudonymous and anonymous data, these concepts are now applied to two categories of data that is frequently processed through DLT. First, the so-called public keys that serve as users' identifiers on such networks are introduced, and second, transactional data will be examined.

3.3. Public keys as personal data

In the blockchain context, public keys serve as the kind of **identifiers** mentioned in Recital 30 GDPR. Blockchains rely on a two-step verification process with asymmetric encryption. Every user has a public key (a string of letters and numbers representing the user), best thought of as an account number that is shared with others to enable transactions. In addition, each user holds a private key (also a string of letters and numbers), which is best thought of as a password that must never be shared with others. Both keys have a mathematical relationship by virtue of which the private key can decrypt data that has been encrypted through the public key.

Public keys thus hide the identity of the individual unless they are linked to additional identifiers. This is course only the case where the public key relates to a **natural person**. There are DLT use cases where public keys do not relate to natural persons. For example, where financial institutions are using a blockchain to settle end-of-day inter-bank payments for their own accounts public keys would relate to these institutions and not natural persons, meaning that they would not qualify as personal data that is subject to the GDPR.[177]

A public key is data that 'can no longer be attributed to a specific data subject' unless it is matched with 'additional information' such as a name, an address or other identifying information, and thus **pseudonymous data** according to Article 4(5) GDPR.[178] Indeed, there are many analogies between

[174] Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, 16.

[175] Ibid.

[176] Commission Nationale de l'Informatique et des Libertés (8 January 2015), *Délibération No. 2013-420 of the Sanctions Committee of CNIL, imposing a financial penality against Google Inc* www.cnil.fr/fileadmin/documents/en/D2013-420_Google_Inc_ENG.pdf.

[177] Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' 25 *Richmond Journal of Law and Technology* 1, 62.

[178] Article 4(5) GDPR.

public keys and other pseudonymous strings of letters and number such as unique identifiers in cookies, which have been said to qualify as personal data.[179]

As per the Article 29 Working Party, pseudonymisation is 'the process of disguising identities' which is precisely what public keys do – but not in an irreversible manner.[180] **Practice reveals that public keys can enable the identification of a specified natural person**. There have been instances where data subjects have been linked to public keys through the voluntary disclosure of their public key to receive funds; through illicit means, or where additional information is gathered in accordance with regulatory requirements, such as where cryptoasset exchanges perform Know Your Customer and Anti-Money Laundering duties.[181] Wallet services or exchanges may indeed need to store parties' real-world identities in order to comply with Anti-Money Laundering requirements while counter parties may do so, too for their own commercial purposes.[182] The **combination of such records with the public key** could thus reveal the real-world identity that lies hidden behind a blockchain address.

Beyond, **public keys may also reveal a pattern of transactions** with publicly known addresses that could 'be used to single out an individual user' such as through transaction graph analysis.[183] On the Bitcoin blockchain encrypted data has been proven capable of revealing a user and transaction nexus that allows for transactions to be traced back to users.[184] Academic research has also confirmed that public keys can be traced back to IP addresses, aiding identification.[185] Where a user transmits a transaction to the network, they usually connect directly to the network and reveal their IP address.[186] Law enforcement agencies across the world have moreover identified individuals through their public keys through forensic chain analysis techniques to identify suspected criminals on the basis of their public keys, and a range of professional service providers performing related services have emerged.[187]

In light of the above it is little surprising that commentators have noted that public keys may constitute personal data under the GDPR. Berberich and Steiner have stressed that '[e]ven if personal information only entails reference ID numbers, such identifiers are typically unique to a specific person. While in all such cases additional information may be necessary to attribute information to the data subject, such information would be merely pseudonymised and count as personal information'.[188] The French Data Protection Authority has equally stressed that public keys likely constitute personal data under the GDPR.[189] The same conclusion has been reached by the

---

[179] Zuiderveen Borgesius F (2016), 'Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation' 32 *Computer Law & Security Review* 256, 260.

[180] Article 29 Working Party, Opinion 04/2007 on the concept of personal data (WP 136) 01248/07/EN, 18.

[181] Philipps Erb K (20 March 2017), *IRS Tries Again To Make Coinbase Turn Over Customer Account Data* https://www.forbes.com/sites/kellyphillipserb/2017/03/20/irs-tries-again-to-make-coinbase-turn-over-customer-account-data/#1841d9e5175e.

[182] Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' 25 *Richmond Journal of Law and Technology* 1, 61.

[183] Ibid, 62.

[184] Reid F and Harrigan M (2018) 'An Analysis of Anonymity in the Bitcoin System' https://arxiv.org/abs/1107.4524

[185] Biryukov A et al (2014), 'Deanonymisation of Clients in Bitcoin P2P Network' https://arxiv.org/abs/1405.7418.

[186] Ibid.

[187] See, by way of example: https://www.chainalysis.com/.

[188] Berberich M and Steiner M (2016) 'Blockchain technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?' 2 *European Data Protection Law Review* 422.

[189] Commission Nationale de l'Informatique et des Libertés (06 November 2018) *Blockchain and the GDPR: Solutions for a responsible use of the blockchain in the context of personal data* https://www.cnil.fr/en/blockchain-and-gdpr-solutions-responsible-use-blockchain-context-personal-data

report of the European Union's Blockchain Observatory and Forum, which has stressed the linkability risk.[190]

Whereas there is a need for a careful case-by-case analysis in each instance, it is evident from the above that public keys directly or indirectly relating to an identified or identifiable natural person qualify as personal data under the EU. Singling out, linkability and even inference can enable to link public keys to an identified or identifiable natural person, and this on public and permissionless and private and permissioned blockchains alike. What is more, as per the Working Party's guidance, it seems that where a public key explicitly serves to identify a data subject, its classification as personal data is always a given.

In any event, entities using distributed ledgers should seek to rely on measures that purposefully make it unlikely that the public key can be related to an identified or identifiable natural person (such as technical and organisational measures that make it create hard barriers between the blockchain and other databases that may contain additional information to enable linkage). The use of one-time public keys also appears as a good practice in this respect. This may be easier to do on private and permissioned blockchains than public and permissionless ledgers due to existing governance mechanisms and institutional structures allowing for such a design.

3.4. Transactional data as personal data

'Transactional data' is the terminology used to refer to **other categories of data that may be used on blockchains but which are not public keys**. This is data about the transaction as such. According to the French Data Protection Authority, this denotes data that is 'contained 'within' a transaction (e.g.: diploma, property deed)'.[191] For example, transactional personal data could be a name, address or a date of birth that is contained in the payload of a given transaction.

To determine whether transactional data meets the GDPR's definition of personal data a case-by-case analysis ought to be undertaken. In some circumstances, transactional data will clearly not qualify as personal data. For example, where blockchains serve as a data infrastructure used to share climatic sensor data from outer space between participants, this may not be personal data. Furthermore, a cryptoasset transferred from A to B unlikely qualifies as personal data unless where it is combined with additional information that specified the product or service that was purchased, which could lead to identification.[192] In other circumstances, such data will however qualify as **personal data**. This could be the case where a group of banks use DLT to share Know Your Customer data.[193] Indeed, the French Data Protection Authority has rightly underlined that where 'such data concerns natural persons, possibly other than the participants, who may be directly or indirectly identified, such data is considered personal data'.[194]

In assessing whether transactional data qualifies as personal data it ought to be borne in mind that under EU data protection law, a **broad definition of the concept of personal data** ought to be

[190] Report of the European Blockchain Observatory and Forum (16 October 2018) *Blockchain and the GDPR* 20 https://www.eublockchainforum.eu/reports.

[191] Commission Nationale de l'Informatique et des Libertés (06 November 2018) *Blockchain and the GDPR: Solutions for a responsible use of the blockchain in the context of personal data* https://www.cnil.fr/en/blockchain-and-gdpr-solutions-responsible-use-blockchain-context-personal-data

[192] Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' 25 *Richmond Journal of Law and Technology* 1, 62.

[193] Ibid.

[194] Commission Nationale de l'Informatique et des Libertés (06 November 2018) *Blockchain and the GDPR: Solutions for a responsible use of the blockchain in the context of personal data* https://www.cnil.fr/en/blockchain-and-gdpr-solutions-responsible-use-blockchain-context-personal-data

embraced in order to safeguard the full and complete protection of data subjects in line with what has been observed above. Transactional data indeed constitutes personal data where it directly or indirectly relates to an identified or identifiable natural person. As distributed ledgers are often used for the tracking of assets (essentially as an accounting mechanism) it is worth highlighting that the United Kingdom's Data Protection Authority has considered that when applying its motivated intruder test (examined above) to financial data, it should be recognised that financial data is particularly appealing for attackers, meaning that intruders should be considered to be particularly motivated in this context.[195] In any event, it is evident that transactional data can be personal data.

Both public keys and transactional data can be used in plain text, in encrypted form, or hashed when put on the blockchain. Where personal data is used in plain text, it undoubtedly remains personal data and accordingly no specific examination of that scenario is necessary here. Below, it is examined whether encryption or hashing are methods capable of transforming personal data in anonymous data. Indeed, while in technical circles there is oftentimes a presumption that such processes anonymize data, this conclusion is not given under the GDPR.

3.4.1. Encryption

Where data is encrypted, the holder of the key can still re-identify each data subject through decryption given that the personal data is still present in the dataset that has been encrypted.[196] As a consequence, **encrypted data remains personal data** – at least for the holder of the key able to identify such data. The Article 29 Working Party indeed clarified in its opinion on cloud computing that although encryption 'may significantly contribute to the confidentiality of personal data if implemented correctly' it does not 'render personal data irreversibly anonymous'.[197]

Commentators have suggested that 'sufficiently well-encrypted data, where the provider has no access to the key, should not be 'personal data', and similarly with sufficiently anonymised data'.[198] This implies that a distinction may have to be operated between those that have **access to the private key** and those that have not. Whether this is the case should be clarified by further regulatory guidance on this matter.

3.4.2. Hash functions

A cryptographic hash is a **mathematical function that is fed an input value that is transformed into an output value of fixed length**. In order to understand hash functions, it is imperative to note that the same input always yields the same output (meaning that they are deterministic). It is moreover not possible to deduce the hash input from the hash output.

Hash functions are often used to strip personal identifiers (such as a name or client number) and replace them with a pseudonym that is difficult to reverse.[199] To illustrate, when I run my own name through the common SHA256 hashing algorithm, this gives me '0F0D284D20C3198C5769E7B19CA37EF5061BEB9FA9BD7C021B4177F06BC54F66' – an identifier that at first sight reveals nothing about myself. Yet, that does not necessarily turn the hash into anonymous data. Even though it is impossible to run that function backward (to derive the input

---

[195] Information Commissioner's Office (November 2012), 'Anonymisation: managing data protection risk code of practice' https://ico.org.uk/media/1061/anonymisation-code.pdf 23.

[196] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 20.

[197] Article 29 Working Party, Opinion 05/2012 on Cloud Computing (WP 196) 01037/12/EN.

[198] Kuan Hon W et al (2011), 'Who is Responsible for 'Personal Data' in Cloud Computing? The Cloud of Unknowing, Part 2' *Queen Mary School of Law Legal Studies Research Paper* No. 77, 18.

[199] Nisbet K (30 April 2018), *The False Allure of Hashing for Anonymization* https://gravitational.com/blog/hashing-for-anonymization/.

from the output) anyone knowing that my name is contained in a dataset may simply enter my name into SHA256 or other commonly used cryptographic hash functions to see what hash is revealed (as the same input always yields the same output).[200] Furthermore, linkability between this dataset and additional information always remains an issue to be carefully determined on a case-by-case basis.

The ease of relating a hash to a data subject should not be underestimated. It has recently been suggested that hashing all existing email addresses globally – around 5 billion – would take about ten miliseconds and cost less than one hundredth of a U.S. dollar.[201] Where an email address is known (such as where it has been revealed through a data breach or was purchased as part of a marketing mailing list), it can be hashed and compared against 'anonymous' email addresses.[202] As running an email address through the same hashing algorithm will always yield the same result, outputs can be guessed from known inputs. Thus, for hashing to be non-invertible, the number of possible inputs must be sufficiently large and unpredictable to prevent the option of trying all possible combinations. In light of the increasing power and decreasing cost of computing, this is hard to achieve. This led Edward Felten to argue that 'hashing is vastly overrated as an 'anonymisation' technique'.[203] He showed that it is in fact fairly easy to establish someone's identity on the basis of hash functions that have been derived from social security numbers by simply having a computer guess all possible social security numbers for one country (about one billion in this example from the United States), hash them, and see which one matches the allegedly anonymous string of letters and numbers that was generated. Whereas this sounds burdensome, Felten famously stated that doing all of this takes a computer less time then it takes the reader to make a cup of coffee.[204]

Whether hashed data always remains personal data for the purposes of the GDPR is a matter of ongoing debate. It should be evident from the above that **the mere use of a hash function will not automatically transform personal data into anonymous data**. The Article 29 Working Party warned that 'pseudonymisation as a process that 'consists of replacing one attribute (typically a unique attribute) in a record by another. The natural person is therefore still likely to be identified; accordingly, pseudonymisation when used alone will not result in an anonymous database'.[205] Hashing will often generate pseudonymous, not anonymous data. Whereas the reversal risk inherent to encryption does not apply to hashing, there is nonetheless a risk that 'if the range of input values the hash function are known they can be replayed through the hash function in order to derive the correct value for a particular record'.[206] As a consequence, the Working Party has warned that whereas hash functions can reduce 'the linkability of a dataset with the original identity of a data subject; as such, it is a useful security measure but not a method of anonymisation'.[207]

There are, however, hash functions with stronger privacy guarantees that may resist the '**means reasonably likely to be used**' test under Recital 26 GDPR. Hashing indeed operates on a spectrum and some of the techniques can go a long way to 'de-personalise' personal data. This has led some

[200] Ibid.

[201] Acar G (9 April 2018), *Four cents to deanonymize: Companies reverse hashed email addresses* https://freedom-to-tinker.com/2018/04/09/four-cents-to-deanonymize-companies-reverse-hashed-email-addresses/

[202] Ibid.

[203] Felten E (22 April 2012), *Does Hashing Make Data "Anonymous"* https://www.ftc.gov/news-events/blogs/techftc/2012/04/does-hashing-make-data-anonymous.

[204] Ibid.

[205] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 20.

[206] Ibid.

[207] Ibid.

authors to state that 'strongly encrypted 'personal data' should already be considered 'anonymous' in the hands of a provider without key access'.[208]

There is also uncertainty as to whether the use of salted and peppered hashes could make the identification of the data subject reasonably unlikely. A **salted hash** can reduce 'the likelihood of deriving the input value' nonetheless, the Working Party unequivocally stressed that they are incapable of producing anonymous data given that 'calculating the original attribute value hidden behind the result of a salted hash function may still be possible within reasonable means'.[209] In contrast to salted hashes, **peppered hashes** rely on a secret key (the 'pepper') as an additional input. [210] The Working Party agrees that peppered hashes offer stronger guarantees as while the 'data controller can replay the function on the attribute using the secret key' it is much more difficult 'for an attacker to replay the function without knowing the key as the number of possibilities to be tested is sufficiently large to be impractical'.[211] Yet, whereas the Working Party envisaged the option of peppered hashes, it falls short of clearly indicating whether these are capable of anonymising data for GDPR purposes. This must be accordingly determined on a case-by-case basis, taking account of the GDPR's test 'all means reasonably likely to be used' test.

The Working party issued **specific criteria** to be taken into account to determine whether identifiability is possible on the basis of the means reasonably likely to be used: (i) singling out, (ii) linking, and (iii) inferences. Even with peppered hashes, **singling out** may remain possible even where inference and likage may not as 'the individual is still identified by a unique attribute which is the result of the pseudonymisation function'.[212] Similarly, **linkability** will often be an issue where links between records relating to the same data subject can be established as even where divergent pseudonymised attributes are used for the same data subject, 'linkability may still be possible by means of other attributes'.[213] This could be the example of biometric data or addresses stored under a pseudonym. Finally, **inferences** also remain a reasonable option where a same dataset or different databases use the same attribute for the same data subject.[214]

It follows that, unless such mechanisms are combined with additional privacy guarantees, peppered hashes remain personal data as there remains a risk of linkability as data subjects may still be identified through indirect identifiers including other information in the dataset or from other sources.[215] Indeed, the Working Party considered that the use of **deterministic encryption or keyed-hash functions with deletion of the key**, can reduce the reasonable likelihood of identification.[216]

Whereas a case-by-case analysis is necessary to determine whether specific data constitutes personal data, the above analysis highlighted that encryption and hash functions do not automatically turn personal data into anonymous data rather, it is necessary to evaluate the precise status of each data item under the test set out in Recital 26 GDPR. It is worth noting that in the specific blockchain context, there are numerous technical developments that seek to offer stronger anonymity guarantees. Before moving on to that discussion, another good practice should be introduced, namely that of the off-chain storage of personal data.

---

[208] Millard C (2013), Cloud Computing Law Oxford: Oxford University Press 178.

[209] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 20.

[210] Ibid.

[211] Ibid.

[212] Ibid, 21.

[213] Ibid.

[214] Ibid.

[215] Millard C (2013) Cloud Computing Law Oxford: Oxford University Press 178.

[216] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 21.

3.4.3. Off-chain data storage

Depending on the specific blockchain use case, it may not be necessary to store all transactional data on the bblockchain itself. Rather, such data could be stored in another, off-chain database and merely linked to the distributed ledger through a hash, a process which would have a number of advantages from a data protection perspective.

Where data is found to classify as personal data it should, where possible, be kept off-chain and merely linked to the ledger through a hash-pointer.[217] Whereas this does not change its nature as personal data, it makes it easier to comply with GDPR requirements. Commentators have indeed recommended that the DLT transaction itself would 'only contain information needed to access the personal data in the separate database. In this manner, it would be possible to confine personal data to the off-chain storage and avoid storing such data on the blockchain'.[218] It is important to note that whereas it will often be possible to store transactional data off-chain, this is not the case for public keys.

Off-chain storage would enable the rectification and erasure of personal data stored off-chain in appropriate databases in light with Articles 16 and 17 GDPR. An open question in this regard is, however, that of the status of the remaining hash. Indeed, the data in off-chain storage will be linked to the database through a hash, and where the off-chain data is erased, that hash will remain on the ledger. To determine whether this hash remains personal data, the means reasonably likely to provoke identification need to be examined. However, this is an era where confusion reigns as many have expressed confusion as to how this ought to be determined. This study recommends that regulatory guidance should be issued on this specific point.

3.5. Ongoing technical developments

Blockchain technologies are a group of technologies that not only assume different characteristics but also remains immature in the sense that further developments are needed to render them useful to the envisaged purpose of use. For example, oftentimes the lack of scalability is an important limit to broad roll-out, as are lacking governance structures to coordinate action and responsibility among multiple actors. Those working on related projects should, and sometimes already are, working on technical solutions to facilitate GDPR compliance, in line with the data protection by design and data protection by default requirements under Article 25 GDPR. Some of these developments could have the potential to anonymize public keys or transactional data. This section provides a non-exhaustive overview of some of these solutions. It is important to highlight that each of these solutions comes with **important trade-offs** that vary depending on context and cannot be examined generically here. These various techniques are briefly introduced here. It will be suggested in the policy recommendation section that these are topics that may be addressed in regulatory guidance, and also that further interdisciplinary research on these matters could evaluate the possibility of making blockchains compliant-by-design through such mechanisms.

3.5.1. Zero knowledge proofs

Zero-knowledge proofs can be used to provide a binary true/false answer without providing access to the underlying data.[219] This allows someone to provide proof of a statement (such as: A is at least

[217] Berberich M and Steiner M (2016), 'Blockchain technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?' 2 *European Data Protection Law Review* 422, 425 and Finck M (2018), 'Blockchains and Data Protection in the European Union' 4 *European Data Protection Law Review* 17.

[218] Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' *Richmond Journal of Law and Technology* 1, 63.

[219] See further, Wu H and Feng Wang F (2014), A Survey of Noninteractive Zero Knowledge Proof System and Its Applications, *The Scientific World Journal*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4032740/.

18 years old or B has at least EUR 1000 in her account) without providing access to the underlying data. For example, the Zcash cryptocurrency relies on this process to ensure that even though transactions are published on a public blockchain its details (including the amount as well as its source and destination) remain hidden.[220] The ledger merely reveals whether a transaction has occurred, not which public key was used or what value (if any) was transferred. [221] The Zerocoin project explores zero knowledge proofs to fix the anonymity deficit of Bitcoin.[222] Where zero knowledge proofs are used, the blockchain indeed only shows that a transaction has happened, not which public key (as sender) transferred what amount to the recipient.[223] It has moreover been pointed ought that zero knowledge proofs and homomorphic encryption have the potential to solve the conflict between data minimisation and the verifiability of data between many parties.[224] A European Parliament report indeed appears to consider zk-SNARKs as a means to comply with the data protection by design requirement.[225]

3.5.2. Stealth addresses

The Bitcoin White Paper recommends that 'a new key pair should be used for each transaction to keep them from being linked to a common owner', while conceding that this is merely a security rather than anonymisation technique as '[s]ome linking is still unavoidable with multi-input transactions, which necessarily reveal that their inputs were owned by the same owner. The risk is that if the owner of a key is revealed, linking could reveal all other transactions that belonged to the same owner'.

Stealth addresses can be used to generate a one-time transaction that relies on hashed one-time keys. For example, the cryptocurrency Monero hides the recipient of a cryptocurrency transaction by generating a new dedicated address and a 'secret key'.[226] The use of one-time accounts for transactions requires that every transaction must completely empty at least one accounts and create one or multiple new accounts. [227] This so-called 'merge avoidance'[228] can be deployed on the Bitcoin blockchain but some consider that even where this is done that system 'has proven to be highly porous and heuristic, with nothing even close to approaching high guarantees' of privacy protection.[229]

3.5.3. Homomorphic encryption

Homomorphic encryption is an advanced method of encryption that enables the computation of cyphertexts. It allows for encrypted data to be subjected to computation, generate an encrypted result that, which decrypted produces the same results than if the computation had been done on

[220] https://z.cash/technology/zksnarks.html.

[221] https://z.cash/technology/zksnarks.html

[222] http://zerocoin.org/

[223] Martini M and Weinzierl Q (2017), 'Die Blockchain-Technologie und das Recht auf Vergessenwerden' 17 *Neue Zeitschrift für Verwaltungsrecht* 1251, 1256.

[224] Böhme R and Pesch P (2017), 'Technische Grundlagen und datenschutzrechtliche Fragen der Blockchain Technologie' 41 *Datenschutz und Datensicherheit* 473, 481.

[225] European Parliament (27 November 2018) Report on Blockchain: a Forward-Looking Trade Policy (AB-0407/2018) para 21.

[226] https://getmonero.org/resources/moneropedia/stealthaddress.html

[227] Buterin V (15 January 2016), *Privacy on the Blockchain* https://blog.ethereum.org/2016/01/15/privacy-on-the-blockchain/

[228] Hearn M (11 December 2013), *Merge Avoidance?* <https://medium.com/@octskyward/merge-avoidance-7f95a386692f.

[229] Buterin V (15 January 2016), *Privacy on the Blockchain* https://blog.ethereum.org/2016/01/15/privacy-on-the-blockchain/.

unencrypted data.[230] It has been argued that this could allow for the use of merely encrypted on-chain data.[231] Whereas, given the regulatory stance on encryption it is doubtful whether this would cross the GDPR's anonymisation threshold, the solution could serve as one element in a broader anonymisation toolbox.

3.5.4. State channels and ring signatures

Other options that are currently being deployed involve state channels for two-party smart contracts that only share information with outside parties in the event of a dispute.[232] Ring signatures on the other hand hide transactions within other transactions by tying a single transaction to multiple private keys even though only one of them initiated the transaction.[233] The signature proves that 'the signer has a private key corresponding to one of a specific set of public keys, without revealing which one'.[234] Whether any of the above solutions can be considered to anonymise public keys remains to be seen. There is presently no legal certainty for developers whishing to handle public keys in a GDPR-compliant manner.

3.5.5. The addition of noise

Another possible solution consists in adding 'noise' to the data.[235] Here, several transactions are grouped together so that from the outside it is impossible to discern the identity of the respective senders and recipients of a transaction. Algorithms similar to this model have already been defined for the Bitcoin[236] and Ethereum blockchains[237]. What is promising about this privacy technique is that the Article 29 Working Party has already recognised that, provided that the necessary safeguards are complied with, the addition of noise may be an acceptable anonymisation technique.[238] For this to be the case, it should be combined with additional privacy mechanisms 'such as the removal of obvious attributes and quasi-identifiers'.[239]

3.5.6. Chameleon hashes and an editable blockchain

Some actors have created 'editable' blockchains using chameleon hash functions to edit, remove or rewrite certain data, such as to accommodate regulatory requirements.[240] Depending on the specific design of these solutions, they could facilitate GDPR compliance. It has, however, been stressed, that as soon as a blockchain becomes editable, the initial argument of using this solution as opposed to other forms of (distributed) databases may be defeated. This will, however, in part depend on the surrounding governance arrangements.

---

[230] Brakerski Z and Gentry C and Vaikuntanathan V (11 August 2011), Fully Homomorphic Encryption without Bootstrapping, https://eprint.iacr.org/2011/277.

[231] Report of the European Blockchain Observatory and Forum, 'Blockchain and the GDPR' (16 October 2018), 23, https://www.eublockchainforum.eu/reports.

[232] Buterin V (15 January 2016), *Privacy on the Blockchain* https://blog.ethereum.org/2016/01/15/privacy-on-the-blockchain/

[233] https://getmonero.org/resources/moneropedia/ringsignatures.html

[234] Ibid.

[235] This has been explored by the MIT ENIGMA project and uses modified distributed hashables to store secret-shared data in combination with an external block chain for identity and access control.

[236] https://sx.dyne.org/anontx/.

[237] https://gist.github.com/gavofyork/dee1f3b727f691b381dc

[238] Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (WP 216) 0829/14/EN, 12-13 (discussing the technique in general, not specifically with respect to blockchains).

[239] Ibid, 12.

[240] See further https://www.accenture.com/us-en/insight-editing-uneditable-blockchain.

3.5.7. Storage limitations

Especially early blockchain projects were designed for the indefinite storage of ledger data to facilitate data integrity and auditability. The idea was that each transaction dating back to the first block (the 'genesis block') would remain on the ledger for as long as it was used. This obviously stands in tension with key data protection requirements such as that of data minimisation and storage limitation. It also renders the anonymisation of data harder, considering that linkage becomes much easier. Various forms of storage limitations could provide an at least partial solution to this issue.

3.5.8. Pruning

Pruning could be a solution to the problems the indefinite storage of data creates from the perspective of key data protection principles such as that of data minimisation and storage limitation. Pruning enables for data to be removed from blockchains when it is no longer needed or wanted.[241] As all solutions, pruning however comes with considerable trade-offs, and it has also been stressed that although the size of the archival node can be reduced through pruning, all the information necessary to recreate the older state is still saved on each node, meaning that is it unlikely that this could qualify as an anonymisation measure from a GDPR perspective.[242]

Beyond, other techniques are currently being explored. The EU Blockchain Observatory and Forum has for instance stressed the potential of **secure multi-party computation** as a further tool that may be explored in this context.[243] Another commonly used obfuscation technique is that of the **third-party indirection service** where a third party aggregated many blockchain transactions to post them to the ledger with their own public key.[244] Whereas these various techniques offer interesting approaches to anonymisation, this is an area that requires further clarity and development. It is with this in mind that further research funding for such methods is suggested below, in addition to a suggestion that the European Data Protection Board update the Article 29 Working Party's guidance on anonymisation and pseudonymisation to provide more clarity in this area. Beyond, it is also suggested below that the European Data Protection Board adopt specific guidelines on blockchain technologies, which should also include information regarding how anonymisation may be achieved in the specific blockchain context. Interdisciplinary research could explore whether it is possible to design protocols that are compliant-by-design. It is, however, also important to be aware of the tension between anonymity in data protection law and other areas of regulation.

3.6. Tension with other policy objectives

The above solutions may be desirable from a data protection perspective as they can go a long way towards anonymisation. This offers higher protection to data subjects, and is appealing to data controllers as it may bring their data processing operations altogether outside the scope of European data protection law.

It must, however, be emphasised that whereas such anonymity solutions are doubtlessly desirable from a pure data protection perspective, the resulting **anonymity can be problematic when**

---

[241] Report of the European Blockchain Observatory and Forum (16 October 2018),, 'Blockchain and the GDPR' 31, https://www.eublockchainforum.eu/reports.

[242] Martinez J (10 September 2018), *Dispelling Myths: How a Pruned Ethereum Node Can Fully Verify the Blockchain* https://medium.com/coinmonks/how-a-pruned-ethereum-node-can-fully-verify-the-blockchain-bbe9f29663ed.

[243] Ibid.

[244] Report of the European Blockchain Observatory and Forum (16 October 2018),, 'Blockchain and the GDPR' 20, https://www.eublockchainforum.eu/reports.

examined through the lens of other policy requirements, such as that of tax evasion or anti-terrorism legislation. The Finance Committee of the French *Assemblée Nationale* indeed suggested banning anonymous cryptocurrencies which rely on tools such as zero knowledge proofs as they facilitate fraudulent and illegal activity such as money laundering and terrorist financing.[245] Similar concerns regarding privacy-protection cryptocurrencies have for instance also been highlighted in Japan.[246]

Whether these techniques can be qualified as achieving the anonymisation threshold remains to be seen. As a consequence, there will be circumstances where public keys and transactional data qualify as personal data. It is hence opportune to now turn to examine the consequences of this state of affairs. Before moving on to determine the rights and obligations arising for various parties where blockchain data qualifies as personal data under the EU data protection regime, I first examine the question of the entities responsible for complying with these obligations, namely data controllers and, depending on the specific context, also data processors.

[245] Assemblée Nationale (30 janvier 2019) Rapport d'Information pat la Commission des Finances, de l'Economie Générale et du Contrôle Budgétaire relative aux monnaies virtuelles, http://www.assemblee-nationale.fr/15/pdf/rap-info/i1624.pdf.

[246] Suberg W (30 April 2018), *Japanese Regulatory Discussed Restricting Trade of Privacy-Focused Altcoins*, Report Says https://cointelegraph.com/news/japanese-regulators-discussed-restricting-trade-of-privacy-focused-altcoins-report-says.

4. Responsibility for GDPR compliance: the data controller

The data controller is **the entity responsible for complying with obligations arising under the GDPR**. The data controller can be a natural or legal person or any other body.[247] The correct identification of controllership in relation to each personal data processing operation is an important exercise as it enables the identification of the person or entity that the data subject is to address to enforce their rights under the Regulation. Indeed, in the words of the Article 29 Working Party, the first and foremost role of the controller is 'to determine who shall be responsible for compliance with data protection rules, and how data subjects can exercise the rights in practice'.[248]

The GDPR is built on the principle that **responsibility and accountability rest with the controller**, who is charged with the practical effectiveness of European data protection law. The controller must **implement appropriate measures**, both of a technical and organisational nature, to be able to demonstrate that its data processing occurs in line with GDPR requirements.[249] Where it is proportionate in relation to the processing activities, the latter shall include the implementation of appropriate data protection policies and compliance with the data protection by design and by default requirements.[250] The controller (or its representative) is moreover obliged to maintain a **record of processing activities** under its responsibility that provides information about the purposes of processing[251], the categories of data subjects and personal data[252], the categories of recipients to whom personal data is disclosed[253], information about personal data transfers[254], and also the envisaged time limits for erasure as well as information about technical and organisational security measures.[255] Beyond, there is an obligation that, at the moment of personal data collection, the controller **provide the data subject with information**, including regarding its own identity and contact details.[256] This highlights that the controller is the entity that is situated at the centre of EU data protection law, charged with the implementation of data protection safeguards ab initio, but also as the central point of contact for data subjects that wish to enforce their rights.

It is important to stress that the relevant data controller must be pinpointed in relation to **each personal data processing operation**, underlining the need for a case-by-case analysis accounting for all relevant technical and contextual factors. The concept of controllership is furthermore **autonomous** as it ought to be interpreted solely on the basis of EU data protection law, and **functional** as 'it is intended to allocate responsibilities where the factual influence ins, and thus based on a factual rather than formal analysis'.[257] Thus the formal identification of a controller in a contract or in terms of conditions is not decisive and can be overturned by a subsequent court decision that determines controllership on the basis of fact rather than form.

---

[247] Although the A29WP has cautioned that in case of doubt 'preference should be given to consider as controller the company or body as such' (such as where the question is whether the controller is a company or its employee). Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 15.

[248] Mahieu R et al (2018) *Responsibility for Data Protection in a Networked World. On the question of the controller, "effective and complete protection" and its application of data access rights in Europe* 12 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3256743.

[249] Article 24 (1) GDPR.

[250] Article 24 (2) GDPR and Article 25(1) GDPR.

[251] Article 30(1)(b) GDPR.

[252] Article 30(1)(c) GDPR.

[253] Article 30(1)(d) GDPR.

[254] Article 30(1)(e) GDPR.

[255] Article 30(1)(g) GDPR.

[256] Article 13(1)(a) GDPR.

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

It has also become evident that the concept of the controller ought to be given a **wide interpretation**. In *Google Spain*, the ECJ stressed the need 'to ensure, through a broad definition of the concept of 'controller', effective and complete protection of data subjects'.[258] As a consequence, the operator of the Google search engine was qualified as a data controller even though it did not 'exercise control over the personal data published on the web pages of third parties'.[259] *Google Spain* continues to have a lasting influence on this area of the law, not only because it set a firm precedent for the broad interpretation of the notion of controllership but also due to the justification that was used. The Court continues to rely on the criterion of '**effective and complete protection**' to justify broad interpretations of various concepts, including that of (joint) controllership as will be seen below.

4.1. The GDPR's definition of the data controller

The text of the GDPR itself contains a specific test designed to determine the identity of the controller in relation to each personal data processing operation. Article 4(7) GDPR indeed provides that the data controller is the person or entity that determines the purposes and means of personal data processing.

Article 4(7) GDPR defines the data controller as:

> the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law[260]

To determine the identity of the data controller in relation to a specific personal data processing operation it is thus necessary to enquire who determines the purposes and means of processing. According to the Article 29 Working Party, 'determining the purposes and means amounts to determining respectively **the 'why' and the 'how'** of certain processing activities'.[261] This underlines that controllership is a functional concept 'intended to allocate responsibilities where the factual influence is'.[262]

In its opinion on SWIFT, the Article 29 Working Party found in 2006 that even though SWIFT had presented itself as a mere data processor, it was in fact a data controller.[263] Indeed, in this specific case, the factual influence test had revealed that SWIFT had 'taken on specific responsibilities which go beyond the set of instructions and duties incumbent on a processor and cannot be considered compatible with its claim to be just a 'processor'' as it in fact determined the purposes and means of processing.[264] This illustrates that the designation of a given entity as the controller (such as in terms and conditions) who does not actually exercise control over the modalities of processing is void.[265] In order to determine controllership, it is accordingly necessary to operate a factual analysis that considers where influence over the means and purposes of personal data processing lies.

---

[258] Case C-131/12 *Google Spain* [2014] EU:C:2014:317, para 34.

[259] Ibid.

[260] Article 4 (7) GDPR.

[261] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 13 (my own emphasis).

[262] Ibid, 9.

[263] https://www.dataprotection.ro/servlet/ViewDocument?id=234

[264] https://www.dataprotection.ro/servlet/ViewDocument?id=234, page 11.

[265] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 9.

The determination of the **means** of processing includes 'both technical and organisational questions'.[266] Where an entity decides to rely on a blockchain as opposed to another form of (decentralised) database, it has made a decision regarding the means of personal data processing, creating a strong indication that it qualifies as the data controller. This would mean that a consortium that relies on a blockchain to manage its accounts, or an insurance company choosing blockchain for the automated payment of its clients are likely data controllers (as they also determine the purposes for which they will need this technology) in relation to the personal data processed through such systems, and accordingly liable to comply with related obligations arising under the GDPR.

Article 4(7) GDPR appears to indicate that the 'means' and the 'purposes' of personal data processing are two factors of equal importance in determining controllership. Over time, case law and regulatory guidance have however underlined the **primacy of the purposes criterion**, that is to say of the 'why' of personal data processing – the motivation of a party to process personal data. The Article 29 Working Party considers that, 'while determining the purposes of the processing would in any case trigger the qualification as controller, determining the means would imply control only when the determination concerns the essential elements of the means'.[267] Indeed, there can be scenarios where one entity alone determines the purposes of processing, while 'the technical and organisational means are determined exclusively by the data processor'.[268] The primacy of the purposes criterion is confirmed by the recent case law on joint-controllership that is discussed further below. Whereas the mere determination of the purpose of processing can lead to being qualified as a data controller, simple determination of the means does not appear to do so.

It is worth noting that according to the Article 29 Working Party, the effective identification of controllership is decisive, 'even if the designation appears to be unlawful or the processing of data is exercised in an unlawful way'.[269] This implies that even where personal data processing has occurred in a context that is per se unlawful, which in the blockchain context could be an Initial Coin Offering ('ICO') that is subsequently declared to be in violation of the applicable legal framework, the entity exercising effective control over the personal data processing in that context nonetheless remains the data controller for GDPR purposes and liable to comply with related obligations.[270]

Article 4(7) GDPR deals with the identification of 'the' controller. However, it is becoming increasingly evident that in personal data processing operations more often than not there are multiple joint-controllers responsible for GDPR compliance. This is also often the case in relation to the operation of a polycentric database such as a blockchain where many parties may contribute to the overall determination of the purposes and means of data processing.

4.2. Joint controllers

Article 26 GDPR foresees that the purposes and means of data processing can be jointly determined by more than one controller. It reads as follows:

> Where two or more controllers jointly determine the purposes and means of processing, they shall be joint controllers. They shall in a transparent manner determine their respective responsibilities for compliance with the obligations under this Regulation, in particular as regards the exercising of the rights of the data subject and their respective duties to provide the information referred to in

[266] Ibid, 14.

[267] Ibid (my own emphasis).

[268] Ibid.

[269] Ibid, 9.

[270] Initial Coin Offerings are a method of fundraising enabled by DLT.

> Articles 13 and 14, by means of an arrangement between them unless, and in so far as, the respective responsibilities of the controllers are determined by Union or Member State law to which the controllers are subject. The arrangement may designate a contact point for data subjects.

The arrangement referred to in paragraph 1 shall duly reflect the respective roles and relationships of the joint controllers vis-à-vis the data subjects. The essence of the arrangement shall be made available to the data subject.

Irrespective of the terms of the arrangement referred to in paragraph 1, the data subject may exercise his or her rights under this Regulation in respect of and against each of the controllers[271]

In assessing joint control, 'a substantive and functional approach' should be adopted.[272] A line of recent case law has provided further indications as to how joint controllership ought to be assessed.

In *Wirtschaftsakademie Schleswig-Holstein* a private educational institution had used Facebook to create a so-called fan page.[273] When users visited that page, a cookie was placed on their computer without them being notified about this by Facebook or the school. As a consequence, the local data protection authority ordered the school to deactivate the fan page. In its judgment, the Court emphasised the importance of embracing a **broad interpretation of joint controllership to ensure the effective and complete protection of data subjects**.[274] It held that the educational institution qualified as a joint controller considering that it subscribed to the conditions of use of the page including its cookie-policy.[275] Before enquiring about the consequences of this ruling for blockchains, it is first necessary to examine the Court's decision in further detail and to also consider subsequent case law on joint-controllership.

In his opinion, Advocate General Bot stated that merely by having recourse to Facebook for the publication of its offers, 'a fan page administrator is subscribing to the principle that the personal data of visitors to his page will be processed for the purpose of compiling viewing statistics. Even though a fan page administrator is not, of course, the designer of the 'Facebook Insights' tool, he will, *by having recourse to that tool, be participating in the determination of the purposes and means of the processing of the personal data of visitors to his page*'.[276] The Advocate General emphasised that the data processing would not occur without the decision of the fan page administrator to use this service. As such, the operator of the fan page may the 'processing of the personal data of users of the fan page possible'.[277]

Seen from this perspective, the **mere agreement** of a natural person or legal entity to the processing of personal data (which itself has been predefined by another party) is sufficient to influence the means and purposes of personal data processing, and conversely to be qualified as a data controller.[278] By implication, the position of Advocate General Bot would imply that **anyone that choses a particular technical infrastructure, such as DLT, to process data, can be a joint-controller of that system** even though they may only have limited control over the purposes and no meaningful control about the means of processing.

---

[271] Article 26 GDPR (my own emphasis).

[272] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 18.

[273] Case C-210/16 *Wirtschafsakademie Schleswig-Holstein* [2018] EU:C:2017:796, para 16.

[274] Ibid, para 23.

[275] Ibid, para 30-32.

[276] Opinion AG Bot in Case C-210/16 *Wirtschafsakademie Schleswig-Holstein* [2018] EU:C:2017:796,, paras 55-56 (my own emphasis).

[277] Ibid.

[278] Ibid.

The Grand Chamber itself recalled the need to ensure the effective and complete protection of data subjects and conversely to adopt a broad definition of controllership.[279] It observed that where Facebook was 'primarily determining the purposes and means of processing personal data' users of a fan page also qualified as joint controllers as they '**subscribe to the conditions of use of the page**' including its cookie-policy.[280] Yet, the Court noted that the educational institution had also been using the related analytics tools for its own purposes, namely 'the promotion of its activity'.[281] Contrary to its Advocate General, the Court cautioned that the 'mere fact of making use of a social network such as Facebook does not make a Facebook user a controller jointly responsible for the processing of personal data by that network'.[282]

Notwithstanding, **the threshold to be crossed to become a controller is very thin**. In this specific case, the fact that the fan page administrator was providing Facebook an opportunity to place cookies on the devices of visitors of the fan page was sufficient to qualify the institution a joint controller.[283] Wherever a fan page is created, this involves the definition of relevant parameters by the user.[284] Indeed, the fan page administrator could request the processing of specific data (such as demographic and geographic data, or information about interests, purchases and lifestyle).[285] Whereas the fan page user only receives such information in an anonymised format, the processing of personal data is nonetheless required to get to that stage.[286] These circumstances led the ECJ conclude that the fan page administrator was a joint controller.[287]

Unlike its Advocate General the ECJ thus found that the mere use of a service provided by others is insufficient to become a joint-controller, yet the fact that such a use enables the service provider to collect personal data is sufficient. One may wonder whether practically speaking, there is really a difference between both options as oftentimes, mere use will imply that this happens. The judgment has indeed been criticised by commentators that have argued that in adopting this approach, the Court weighted more heavily 'the need to ensure effective and complete protection, than a literal interpretation of the law's words'.[288]

Whereas *Wirtschaftsakademie Schleswig-Holstein* has left little doubt that a **broad interpretation of the notion of joint-control ought to be adopted**, much uncertainty remains as to the concrete practical application of this test. It is **far from clear what degree of involvement is necessary to be qualified as a joint controller**, specifically whether any involvement leads to joint-controllership or whether a *de minimis* test applies. This had been advocated by Advocate General Jääskinen in *Google Spain*, who considered that the broad definitions of personal data, personal data processing and data controller 'are likely to cover an unprecedently wide range of new factual situations due to technological development'.[289] This, he warned, should oblige the Court to 'apply

[279] Case C-210/16 *Wirtschafsakademie Schleswig-Holstein* [2018] EU:C:2017:796, para 28.

[280] Ibid, para 30-32.

[281] Ibid, para 34.

[282] Ibid, para 35.

[283] Ibid.

[284] Ibid, para 36.

[285] Ibid, para 37.

[286] Ibid, para 38.

[287] Ibid, para 39.

[288] Mahieu R et al (2018) *Responsibility for Data Protection in a Networked World. On the question of the controller, "effective and complete protection" and its application of data access rights in Europe* 17 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3256743.

[289] Case C-101/01 *Bodil Lindqvist* [2003] EU:C:2003:596, para 30.

a **rule of reason**, in other words, the principle of proportionality' in interpreting EU data protection law 'to avoid unreasonable and excessive legal consequences'.[290]

In the *Jehovan Witnesses* case, decided just a month after *Wirtschaftsakademie Schleswig-Holstein*, the Court confirmed that joint controllership is a concept that ought to be interpreted broadly. The case *inter alia* concerned the question of whether the religious community could be considered a controller in relation to the preaching activity undertaken by its members in the course of which personal data was collected.[291] The Grand Chamber again emphasised the need to adopt a broad definition of controllership to ensure the effective and complete protection of data subjects.[292] Despite the precise modalities of preaching being determined by the members of the community (and the related collection of personal data of the persons they visited) the Court considered that the Jehovah's Witness Community contributed to such activity 'by organising, coordinating and encouraging the preaching activities of its members intended to spread its faith, participates, jointly with its members who engage in preaching, in determining the purposes and means of processing of personal data of the persons contacted'.[293] They were found to be joint controllers as 'a natural or legal person who **exerts influence over the processing of personal data, for his own purposes**, and who participates, as a result, in the determination of the purposes and means of that processing, may be regarded as a controller'.[294]

According to *Jehovan Witnesses* the relevant criterion is thus that of '**exerting influence**' over personal data processing for one's own purposes. Pursuant to this expansive approach to joint controllership, many parties might qualify as joint controllers in contexts where blockchains is the chosen technology of personal data processing as will be illustrated in further detail below. It is worth noting that some of these issues might be clarified in the Court's upcoming decision in *Fashion ID*. In his opinion, Advocate General Bobek warned of an overly expansive interpretation of joint-control.[295] Here, an online retailer had embedded a plug-in in its website (the Facebook 'Like' button) and the Advocate General found that in light of existing case law, the conclusion that the website operator was a joint controller could not be avoided. Indeed, it was the owner of the website that enabled Facebook to 'obtain the personal data of the users (…) by using the plug-in at issue' in the first place.[296] At the same time, he criticised current case law as 'when pushed to the extreme, if the only relevant criterion for joint control is to have made the data processing possible'.[297] After having provided an overview of the law on controllership, it is now time to apply these findings to various forms of blockchains.

4.3. Data controllers for blockchain-enabled personal data processing

Considering the need for a broad interpretation of (joint-)controllership, many actors may qualify as (joint) controllers. It must be recalled that controllership ought to be determined on a case-by-case basis. To identify the actors determining the purposes and means of data processing in a specific use case, it is not only necessary to consider the specificities of that use case and the manner in

[290] Ibid.

[291] Case C-25/17 *Jehovan todistajat* [2018] EU:C:2018:551.

[292] Ibid, para 66.

[293] Ibid, para 73.

[294] Ibid, para 68.

[295] Opinion of AG Bobek in Case C-40/17 *Fashion ID* [2018] EU:C:2018:1039.

[296] Ibid, para 67.

[297] Ibid, para 74.

which personal data is handled, but moreover to carefully examine the governance design of a given blockchain.[298]

The below observations can thus only be of a general nature and serve to underline the difficulty of determining the controller's identity in a generalised fashion. Indeed, the relevant literature on distributed ledgers and the GDPR echoes a lack of consensus as to who should be considered as the controller of a given blockchain-enabled data processing operation. This is in part due to different understandings of what a blockchain is and how it is used, but also the different roles of various actors depending on the relevant technical and governance designs (such as what consensus protocols are used) and the uncertain legal test that ought to be applied.

First, it is worth reflecting briefly on how the notions of the 'means' and 'purposes' ought to be interpreted in contexts where DLT is used to process personal data. Regarding the **means of personal data processing,** the first difficultly is to identify the perspective that should be adopted. In cloud computing, cloud providers can be considered to determine the means of processing because they chose the software, hardware and data centers that are used.[299] By analogy, the parties that exercise influence over the software, hardware and data centres that are used to operate a specific blockchain can be considered to influence the means of processing. Further, in line with recent case law such as *Wirtschaftsakademie Schleswig-Holstein*, the mere use of a blockchain infrastructure made available by others (such as blockchain-as-a-service solutions) may be considered as an implicit determination of the means of processing.

Because blockchains are distributed databases designed to be operated by many different parties, **many actors influence the determination of the means of processing**. Regarding private and permissionless blockchains, the means are usually determined by the entity (such as a specific company) or association of entities (such as a consortium). Regarding public and permissionless blockchains the given governance arrangements influence the modalities of the means of processing. As a general rule, there is not a single legal entity that decides which software, hardware and data centers to use. Rather, these decisions are made by a range of different actors. To illustrate, in proof-of-work systems, miners make the decision of what hardware (for mining) and data centers (for mining) to use whereas core developers suggest whether and if so how software should be updated.[300] Depending on the chosen governance set-up, miners, nodes and/or coin holders then make a decision as to what software to actually implement.

The Article 29 Working Party has moreover indicated that emphasis should be placed on the '**effective means**' understood as 'substantial questions' that are 'essential to the core of processing' – the how of processing.[301] This includes what data to process and for how long, which third parties have access to the data, when and how data can be manipulated.[302] This guidance appears to indicate, that these are the core competencies of the controller in relation to the 'means' of processing whereas the proper technical and organisational measures can be delegated to processors.

Regarding the **purposes of data processing**, which ought to be treated as the main criterion to assess controllership, an equally fragmented picture emerges. This can be illustrated by the example of a simple token transaction. Clearly, the objective of transferring ownership of the given token

[298] On blockchain governance, see further Finck M (2019) Blockchain Regulation and Governance in Europe Cambridge: Cambridge University Press.

[299] Kuan Hon et al W. (2011) 'Who is Responsible for 'Personal Data' in Cloud Computing? The Cloud of Unknowing, Part 2' Queen Mary School of Law Legal Studies Research Paper No. 77, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1794130 10.

[300] Ibid.

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

[302] Ibid.

through this chosen technical infrastructure is the main objective. But could the continuous existence of the blockchain also be considered a (secondary) purpose?[303] Indeed, whoever chooses DLT as an infrastructure for such objectives, may also do so in light of the specific nature of this database, that is to say as an ever-growing ledger of transactions that may exist forever due to its resilience-by-replication characteristics. Furthermore, the relevant stakeholders do not necessarily have a guarantee that all nodes will erase their local version of the database. This is a pivotal element that necessitates further clarification as it is also crucial in relation to the purpose limitation principle that is examined below. Notwithstanding, the Court's recent case law (examined above) would indicate that an influence over any purpose of processing will be sufficient for an entity to qualify as a data controller. As a consequence, **many different entities potentially qualify as (joint-) controllers** in contexts where DLT is used. Below, an overview of the main relevant actors is provided.

4.3.1. Blockchain-based applications

The first blockchain, Bitcoin, was designed for users to directly interact with. Where this happens, controllership ought to be determined at the infrastructure level – a scenario examined just below. However, as the technology and related business models mature, the emergence of a multi-layered ecosystem can be observed, which includes an **application layer**. In these scenarios DLT merely serves as an infrastructure that anchors these applications which may themselves determine the means and purposes of personal data processing. Where an application layer exists, there is an argument to be made that **the legal entity determining the purposes of personal data processing at the application layer qualifies as the data controller.**

The Article 29 Working Party has argued in relation to social networking that the social network itself is a controller as it often determines the purposes and means of data processing.[304] Beyond, application providers are data controllers where 'they develop applications which run in addition to the ones from the SNS and users decide to use such an application'.[305] This is very similar to blockchain-based applications and indicates that in a multi-layered environment there are likely numerous (joint-)controllers that each have responsibility for various elements of the overall data processing. Where a data subject relies on an intermediary, such as a cryptoasset wallet provider, then that provider is likely also a/the controller. These intermediaries indeed generate and store the public and private keys and also transmit signed transactions to the network.[306]

It is worth noting that the CNIL has stressed in relation to relation to smart contracts that the developer of the software can be a simple external provider but, if they actively participate in the data processing they can also be found to be a processor or joint controller, depending on their role in the determination of the purposes of processing (note that also here, the CNIL mainly looks towards the purposes, not the means, of processing to determine controllership).[307]

4.3.2. Private and/or Permissionless Blockchains

In private and/or permissionless DLT, there is generally a **determined legal entity** (such as a company or a consortium) that determines the means and in many cases also the purposes of personal data processing. Where this is the case, that entity qualifies as the data controller. However,

[303] This argument is examined in further detail below in relation to the purpose limitation principle.

[304] Article 29 Working Party, Opinion 5/2009 on online social networking (WP 163) 01189/09/EN, 5.

[305] Ibid.

[306] Erbguth J and Fasching J (2017), 'Wer ist Verantwortlicher einer Bitcoin-Transaktion?' 12 *Zeitschrift für Datenschutz* 560, 563.

[307] Commission Nationale Informatique et Libertés, 'Premiers Éléments d'analyse de la CNIL : Blockchain' (September 2018), 3.

there may also be joint controllers in such circumstances. In line with *Wirtschaftsakademie Schleswig Holstein*, it can be argued that those using such infrastructure for their own purposes are joint controllers.

An example would be a consortium blockchain established between many actors that are part of the same supply chain. Clearly, the legal entity created by the consortium would be a controller considering that it exercises significant control over the purposes and also the means of personal data processing. Yet, the individual companies that have joined the consortium and are subsequently using the infrastructure for their own purposes thus enabling the DLT to process new personal data could also qualify as joint-controllers.

In its 2018 guidance on blockchains and the GDPR, the French CNIL considered that where a group collectively decides to use DLT for their own purposes, the data controller should be defined *ab initio*.[308] It moreover explicitly mentioned the options of the creation of a new legal person or the designation of an existing legal person as the data controller.[309] Indeed, particularly where these different entities qualify as joint controllers under Article 26 GDPR; they must conclude an agreement setting out their respective responsibilities.[310] This then also allows data subject to identify the entity they need to contact to enforce their rights and provides a single point of contact for data protection authorities.[311] In line with the functional approach, such an initiative would not, however, prevent the identification of other joint-controllers by subsequent court decisions, and the imposition of GDPR-related duties on these actors.

4.3.3. Public and permissionless blockchains

Where a data subject engages directly with the blockchain infrastructure level, it becomes necessary to determine controllership at the infrastructure level. This, however, is far from straightforward. Bearing in mind the need for a contextual case-by-case analysis general reflections on this topic are provided below.

It is important to stress that the identity of the data controller depends on the perspective that is adopted. Seen from a macro-level, the purpose of processing is to 'provide the associated service' (such as a Bitcoin transaction) whereas the 'means' related to the software used by nodes and miners.[312] From a micro-perspective (that is to say the individual transaction) the purpose of processing is 'to record a specific transaction onto a blockchain' whereas the means refer 'to the choice of the blockchain platform'.[313] Arguably, the micro-level is the more appropriate approach as data protection law deals with specific items of personal data.[314] With this in mind, the below analysis discusses which of the many participants in public and permissionless blockchain ecosystems are likely to qualify as data controllers.

Software developers

Of all the parties that use or contribute to the establishment and maintenance of DLT, software developers are the least likely to qualify as controllers. Developers indeed have some role in the design of the relevant software as they suggest software updates to others. However, they do not usually decide on whether such updates are adopted or not – highlighting that their influence over

[308] Ibid.

[309] Ibid.

[310] Ibid.

[311] Ibid.

[312] Bacon J et al (2018) 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' 25 *Richmond Journal of Law and Technology* 1, 64.

[313] Ibid.

[314] Ibid.

the means of processing is limited. Software updates indeed are, depending on the relevant governance structure of a given blockchain, decided by miners, nodes or other actors such as coin holders. Developers accordingly have a limited role in determining the means of processing, and generally exercise no influence over the purposes of a specific personal data processing operation as they merely make available an infrastructure for others to use to realize their own purposes. Unless the specific factual circumstances of a given use case change these assumptions, software developers are unlikely to qualify as (joint) controllers under the GDPR.

Miners

Where proof-of-work serves as the consensus protocol that enables the addition of new data to a blockchain, miners are responsible for the addition of such information. Miners are nodes that group transactions into new blocks and suggest them to the network in accordance with the consensus algorithm.[315] In return for their processing, they are rewarded with newly minted coins in the form of a block reward and they potentially also receive transaction fees paid by users to secure the fast processing of their transactions.[316]

Miners run the protocol, can add data to the shared ledger and store a (usually full) copy of the ledger on their machines.[317] Yet, there is a debate to be had as to whether their influence goes as far as to determine the 'purposes and means' of processing. Miners exercise significant control over the means in choosing which version of the protocol to run. Yet, considering that the criterion of the means has become subsidiary to the 'purposes' criterion, and miners do not determine the purposes of a specific transaction, they unlikely qualify as controllers. This led the CNIL to argue in its 2018 guidance that miners are not controllers.[318] Miners are indeed better seen as 'servants' of the overall system (that benefit financially from its maintenance, at least in a system that uses proof-of-work).[319] As such their role has been compared to that of telecommunications providers that are not legally liable for the content of the data they transmit.[320]

Nodes

The 'nodes' are the computers that store a full or partial copy of a blockchain and participate in the validation of new blocks. Once a miner finds a valid hash for a block, it broadcasts its hash to other nodes, which subsequently run a computation to verify whether the hash is valid (i.e. whether it meets the specifications of the protocol) and where this is the case, they add the new block to their own local copy of the ledger. In doing so, nodes verify whether transactions have the correct digital signatures and data format.[321] Nodes also check whether cryptoassets from the input address have been previously spent in order to prevent the 'double-spending' problem.[322]

---

[315] Blockchains' rely on ansynchorous protocols in accordance with which nodes do not wait to synchronize with their peers to validate specific blocks, rather they validate blocks on the basis of the next block available to them. Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' 25 *Richmond Journal of Law and Technology* 1, 13.

[316] Narayanan A et al (2016), Bitcoin and Cryptocurrency Technologies Princeton University Press.

[317] Note the distinction between full and lightweight nodes in some networks. The former store and entire copy of the database whereas the latter may only store those elements of the blockchain that is relevant to them.

[318] Commission Nationale Informatique et Libertés, 'Premiers Éléments d'analyse de la CNIL : Blockchain' (September 2018), 2.

[319] Martini M and Weinzierl Q (2017), 'Die Blockchain-Technologie und das Recht auf Vergessenwerden' 17 *Neue Zeitschrift für Verwaltungsrecht* 1251, 1253.

[320] Ibid.

[321] Report of the European Blockchain Observatory and Forum (16 October 2018), 'Blockchain and the GDPR' 14, https://www.eublockchainforum.eu/reports.

[322] Buocz T et al (2019), 'Bitcoin and the GDPR: Allocating Responsibility in Distributed Networks' *Computer Law & Security Review* 1, 24.

Martini and Weinzierl have suggested that each node that initiates a transaction (and thus distributes information to all other nodes) or that saves a transaction in its own copy of the database is a controller, considering that in doing so, the node pursues its own purpose: participation in the network.[323] In doing so, the node registers, orders and stores data and can freely use the data that is registered on its own node.[324]

Bacon et al have considered that nodes and miners could be compared to SWIFT, a financial messaging service that facilitates international money transfers for financial institutions and processes the personal data of the payers and payees.[325] It has already been seen above that even though SWIFT deemed itself to be a processor, the Article 29 Working Party argued that it was a controller as it exercised significant autonomy in data processing and had decided to established a US-based data center to disclose data to US authorities.[326] It has moreover been argued that nodes can be understood as joint controllers considering that they 'have equal influence and freedom to choose (or start) a certain blockchain-network – and can, for example with the necessary majority by a Fork, change the rules' is a sign of joint control.[327]

Users

Users, which can be natural or legal persons, sign and submit transactions to the given blockchain. It has been suggested that users should be considered to be the data controllers where a transaction is made directly by the user, the 'technical construct of the blockchain leads to the fact that only the user undertaking the transaction can determine the purposes and means of data processing'.[328] This is said to be the case as the user directly installs the client that connects to the network and sends transactions to other nodes. The client software can moreover be used to keep the private key (which, alternatively, can be stored on specific hardware or offline on paper).[329]

Bacon et al concur that users can be controllers where they determine the purposes of processing (namely to record a specific transaction onto the blockchain) while also determining the means in using a specific blockchain to execute their transactions. A **recent European Parliament report** embraces the same view in suggesting that users 'may be both data controllers, for the personal data that they upload to the ledger, and data processors, by virtue of storing a full copy of the ledger on their own computer'.[330]

The **French Data Protection Authority CNIL** has examined users' potential role as data controllers under the GDPR in further detail. It has suggested that where a user is a natural person, the GDPR will in some circumstances fall short of applying in light of the application of the household

[323] Martini M and Weinzierl Q (2017), 'Die Blockchain-Technologie und das Recht auf Vergessenwerden' 17 *Neue Zeitschrift für Verwaltungsrecht* 1251, 1253.

[324] Ibid, 1254.

[325] Bacon J et al (2018) 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' 25 *Richmond Journal of Law and Technology* 1, 71-72.

[326] Article 29 Working Party, Opinion 10/2006 on the processing of personal data by the Society for Worldwide Interbank Financial Telecommunication (SWIFT) (WP 128) 01935/06/EN, 11.

[327] Wirth C and Kolain M (2018), 'Privacy by BlockChain Design: A Blockchain-enabled GDPR-compliant Approach for Handling Personal Data' in 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* , 5 https://dl.eusset.eu/bitstream/20.500.12015/3159/1/blockchain2018_03.pdf

[328] Erbguth J and Fasching J (2017), 'Wer ist Verantwortlicher einer Bitcoin-Transaktion?' 12 *Zeitschrift für Datenschutz* 560.

[329] Ibid, 563.

[330] European Parliament (27 November 2018), Report on Blockchain: a Forward-Looking Trade Policy (AB-0407/2018) para 22.

exemption.[331] As suggested above, this is however unlikely to be the case where a public and permissionless blockchain is used, as in that scenario personal data would be shared with an indefinite number of people. The CNIL has also recognised that where the household exemption does not apply because the purpose of the transaction is professional or commercial, users of a given blockchain can be considered to be controllers.[332] The French DPA considers that in such scenarios, users determine the purposes of processing (their motivation for using the technology) and also influence the means – such as the format of the data and the choice to use a blockchain compared to other technology.[333]

There is accordingly broad consensus that DLT users will in at least some circumstances be considered as data controllers under the GDPR. The implications of such a finding must, however, be carefully considered. Two scenarios should be distinguished in this respect, namely whether a user processes others' or their own data.

The user as the controller regarding personal data relating to *others*

The above reasoning has revealed that the user qualifies as a data controller where they determine the purposes and means of personal data processing. Depending on the specific factual circumstances, the personal data that is processed may relate to either users themselves or to other natural persons. The latter scenario is examined first. For example, an individual initiating a Bitcoin transaction is the controller of the personal data of the party they are buying Bitcoin from or selling it to. That individual indeed determines the purposes of processing (buying or selling Bitcoin) as well as the means (choosing to rely on the Bitcoin blockchain).

It is difficult to ignore the analogies between the facts in *Wirtschaftsakademie Schleswig Holstein* (where an economic actor chose to rely on Facebook fan pages for its own purposes and was found to be a joint controller) and some DLT use cases. Where a bank relies on DLT to manage client data it would be a controller.[334] By analogy, even where the user is a natural person, they can be the controller where they **process personal data for their own purposes**. It is true that the emphasis on the choice of the given architecture can be criticised considering that there is no real choice between various providers for someone wishing to buy or sell Bitcoin, just as there are few genuine alternatives to Facebook for economic operators wishing to advertise their products in a specific manner. It may further be criticised that this rationale shifts responsibility for technology design away from the actual designers and towards users who may not only lack economic alternatives but also the required expertise to make informed decisions regarding the design of the respective technology. Nonetheless, in line with *Wirtschaftsakademie Schleswig Holstein*, the qualification of the user processing personal data relating to a natural person would call for the qualification of the former as a data controller.

This conclusion also appears to be in line with the recommendation of the Article 29 Working Party that a user of a **social media network** can be a controller.[335] This conclusion would furthermore echo the Working Party's recommendations that a **cloud computing** user is the controller of personal data processed in the cloud. Here, the cloud client is considered to be the controller as it 'determines the ultimate purpose of the processing and decides on the outsourcing of this

[331] Commission Nationale Informatique et Libertés (September 2018), 'Premiers Éléments d'analyse de la CNIL : Blockchain' 3.

[332] Ibid.

[333] Ibid, 2.

[334] Commission Nationale Informatique et Libertés, 'Premiers Éléments d'analyse de la CNIL : Blockchain' (September 2018), 2.

[335] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 21.

processing to the delegation of all or part of the processing activity to an external organisation'.[336] Similarly, where such an organisation chooses to rely on a given DLT infrastructure, whether public or private, permissioned or permissionless, it will have determined the means of personal data processing in addition to its own specific purpose for processing said data and accordingly be subject to controllership duties.

It is however worth noting that some have questioned whether users of technical infrastructure really have control over the purposes and means. Indeed (and unlike Facebook fan page administrators that may define criteria of data processing) the user of a blockchain (Bitcoin in the example of these commentators) only determines 'if a transfer is created and to whom and how much BTC are being transferred'.[337] The purpose here is always to transfer Bitcoin and this purpose cannot be altered by the user. The user moreover has no influence over how long data is stored for, which third parties have access to and when data is deleted.[338] On the other hand, the user does, however, have influence over the purposes and means such as whether to include a message in the transaction or not – showing that they have some degree of control over the means, in addition to the determination of the purposes which, in line with what was observed above, should in any event be considered to be the most important criterion.

There is accordingly consensus in the literature that a blockchain user ought to be considered as a (joint-) controller given that their choice of the relevant infrastructure qualifies as a determination or the means of processing, and their reason for using such technology qualifies as a determination of the purposes of processing. The conclusion that a user qualifies as a data controller may, however, be less straightforward where the personal data that is processed directly or indirectly relates to the user qua natural person, that is to say the data subject.

The user as the controller regarding personal data relating to *themselves*

Whenever **a user qua natural person** signs and submits a transaction, they do not just process others' personal data (such as someone else's public key) but also their own (such as their own public key). It has been outlined above that in such circumstances, the household exemption under Article 2 GDPR is unlikely to apply, considering that where private and permissioned blockchains are used, the purpose of processing will ordinarily be of a commercial or professional nature. Conversely, where public and permissionless blockchains are used, on-chain data is made available to an indefinite number of people so that, in line with the Court's settled case law on this matter the household exemption cannot apply.

To some, the possibility of data subject/data controller overlap is uncontroversial and considered to be a settled issue in EU law.[339] A closer look at existing guidance and the general scheme of the GDPR however underlines that this conclusion might not, in fact, be as straightforward. Indeed, a detailed examination of the European data protection framework and its interpretation reveals that it remains an open question whether the data subject can be considered as the data controller in relation to personal data that directly or indirectly relates to themselves. Maybe surprisingly, there seems to have been little explicit discussion of this question to date. Ongoing technical developments, such as those relating to DLT, may now compel us to answer this question explicitly, in addition to broader discussions regarding the importance and options of giving data subjects more control over personal data.

---

[336] Article 29 Working Party, Opinion 05/2012 on Cloud Computing (WP 196) 01037/12/EN, 7.

[337] Buocz T et al (2019), 'Bitcoin and the GDPR: Allocating Responsibility in Distributed Networks' *Computer Law & Security Review* 1, 24.

[338] Ibid, 1.

[339] Moerel L (2019) 'Blockchain & Data Protection…and Why They are not on a Collision Course' 6 *European Review of Private Law* 825, 843.

The Court's recent case-law on (joint-) controllership has firmly underlined that the purposes of processing ought to be taken as the main criterion to establish controllership. By analogy, any party that determines the purposes of using a specific service risks being qualified as the data controller. This conclusion appears unavoidable in ever more contexts where data is processed as nowadays – and in contrast to when the 1995 Data Protection Directive was first designed – the generation and sharing of personal data oftentimes occurs at the request of users.

To shed further light on this topic, it is useful to go back to the guidance that has already been issued. Regarding **online social networking**, the Article 29 Working Party in 2010 indicated that social media network users 'would qualify as controllers provided that their activities are not subject to the so-called 'household exemption'' in publishing and exchanging information with other users.[340] This is probably why some have considered it established that a data subject/data controller overlap is possible. However, the passages of the Working Party's guidance following this statement appear to indicate that what the Working Party here had in mind was not the personal data relating directly to that person but rather that of others (such as a picture of someone shared on the social network). Indeed, it considered that the user then needs the consent of the concerned data subject if not other lawful grounds of processing are available.[341]

It is, indeed an open question whether a data subject/data controller overlap would be compatible with the broader underlying objective of the GDPR, which was designed precisely to give data subjects rights vis-à-vis controllers in a context of unbalanced power-relations. Indeed, at first sight, a finding that a data subject may be the data controller in relation to her own data maybe understood as a finding of empowerment – the idea that the natural person would be 'in control of' her data in line with the GDPR's overarching rationale of **data sovereignty**.

A closer look reveals, however, that the opposite may be the case as considering a data subject/data controller overlap may also result in **less responsible and accountable forms of personal data processing**. Indeed, in practice the data subject is unlikely to understand the complexity of personal data processing implications and ecosystems. The data subject may be overburdened with responsibility and decisions. Social science research has furthermore revealed that it is questionable whether data subjects can really make the best decisions even if they are given sufficient information.[342]

These uncertainties echo a broader difficulty of determining the identity of the controller in polycentric networks. In the words of the Article 29 Working Party the concrete application of the concepts of controller and processor is 'becoming increasingly complex' due to the growing complexity of contemporary data environments.[343] To account for such complexity, a functional case-by-case analysis that determines why processing takes place and who initiated it has been recommended.[344]

It can accordingly be questioned whether the determination of a data subject as the data controller in relation to personal data that directly or indirectly refers to herself is compatible with the overarching spirit of the EU data protection framework. It would accordingly be important that this question is addressed explicitly to provide further clarity if the European Data Protection Board were to issue guidance on blockchain technology. It is now time to determine the consequence that flow from a finding of (joint-) controllership.

---

[340] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 21.

[341] Article 29 Working Party, Opinion 5/2009 on Online Social Networking (WP 163) 01189/09/EN, p. 6.

[342] Cranor L (2012), 'Necessary But Not Sufficient: Standardised Mechanisms for Privacy Notice and Choice' 10 *Journal on Telecommunications and High Technology Law* 273.

[343] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 2.

[344] Ibid, 8.

4.4. The importance of the effective identification of the controller

It of pivotal importance that the data subject can be easily identified to ensure that data subjects and supervisory authorities hold data controllers accountable for GDPR compliance. At the same time, potential controllers also need certainty regarding when they may qualify as such, and this even before they are addressed by a data subject in relation to their rights. Indeed, many provisions of the GDPR create obligations that data controllers must comply with at the time of, or even before, the initial personal data processing starts. For example, **Article 13 GDPR** obliges the controller to provide the data subject with information 'at the time when personal data are obtained'.[345] This information includes, inter alia, information pertaining to the identity and contact details of the controller, information regarding the purposes of processing and information regarding the recipients and categories of personal data as well as whether the controller intends to transfer the data to a third country.[346] These informational duties seek to achieve the GDPR's objectives of transparency, lawfulness and fairness.[347]

The provision of such information is also important where consent is the chosen lawful basis of processing as **Article 4(11) GDPR** requires that the data subject's consent be 'informed'.[348] More broadly, the data subject can only enforce any of her rights where they are provided with the information required under Articles 13 and 14 GDPR by the data controller. Indeed, the Court held in *Bara* that 'the requirement to inform the data subjects about the processing of their personal data is all the more important since it affects the exercise by the data subjects of their right of access to, and right to rectify, the data being processed ... and their right to object to the processing of those data'.[349] This duty applies both where the data subject knowingly provides personal data to the controller (such as where they fill in an online form that is required before they can use a certain blockchain or blockchain-based application) but also where the data controller collects personal data about the data subject through observation (such as where the controller uses tracking, RFID or sensors).[350]

The centrality of these obligations to the overall GDPR scheme and the requirement that such information is provided at the time of collection rather than after the fact underlines the importance of actors being able to determine in advance whether they qualify as a controller and must comply with related obligations.[351] The data controller is not just obliged to 'implement appropriate technical and organisational measures' to ensure and to be able to demonstrate that processing is performed in compliance with the GDPR'.[352] This may include the implementation of appropriate data protection policies where proportionate in relation to processing activities.[353]

It is worth noting that these uncertainties also make it more difficult to distinguish between the controller and processor in relation to a specific operation. To illustrate, in the cloud computing scenario the user will typically be the controller whereas the cloud provider could be a joint

[345] Article 13(1) GDPR.

[346] Article 13(1) GDPR. Note also the additional obligations in Article 13(2) GDPR as well as Recitals 60-62 GDPR. Article 14 GDPR applies where personal data is obtained from a third party.

[347] Article 5(1)(a) GDPR.

[348] Article 4(11) GDPR.

[349] Case C-201/14, *Bara*, para 33.

[350] Article 29 Working Party, 'Guidelines on Transparency under Regulation 2016/269' WP29 2018B 14.

[351] Note also that in accordance with Article 24(1) GDPR; the data subject must be in a position to prove compliance with this requirement pursuant to the principle of accountability. See also Article 5(2) GDPR.

[352] Article 24(1) GDPR.

[353] Article 24(2) GDPR.

controller, such as where the provider processes data for its own purposes.[354] It could however also be a processor, or a third party. These two categories of actors are examined next.

4.5. The consequences of controllership

This section briefly reflects on two implications of a finding of controllership. First, the nexus between responsibility and control, and second the implications of a finding of joint controllership.

4.5.1. The nexus between responsibility and control

The preceding analysis has revealed that **controllership regarding DLT-enabled personal data processing cannot be determined in a generalised manner**. Rather, a case-by-case analysis accounting for technical and contextual factors ought to be carried out. It is of course easier to determine controllership regarding private and/or permissioned blockchains as there is often a specific legal person that determines the means and purposes of processing. Determining controllership is, however, less straightforward in public and permissionless blockchains, or indeed any polycentric network the participants in which are only loosely associated.

It has been observed that depending on the respective determination of the means and purposes of data processing in each specific use case, many actors may qualify as data controllers. This could, for instance, include nodes and users. Many of these actors, however, have very limited influence over the respective means of data processing. This may then result in a situation where **a data controller may be unable to comply with GDPR obligations due to their insufficient control over the data**. This point will be made below, where it will be seen that these actors would, for example, be able to comply with a data subject's request for access under Article 15 GDPR as they may only see encrypted data, or with a request for erasure or data portability due to lacking influence over data processing. The question of whether responsibility should be possible without control is one that ought to be reflected on more broadly in European data protection law, and this could also be included in possible regulatory guidance to be issued on this matter.

For the time being, there are however further steps that actors using blockchain technology can take to bring further clarity on responsibility for GDPR compliance (a duty falling on the controller). The Article 29 Working Party has acknowledged that innovations in technology make it more burdensome to determine controllership. Indeed, 'many scenarios can be foreseen involving controllers and processors, alone or jointly, with different degrees of autonomy and responsibility'.[355] What is paramount in such contexts is that responsibility is allocated in such a way that compliance with data protection rules will be sufficiently ensured in practice'.[356]

Indeed, the most important task that falls to the controller is to 'allocate responsibility'.[357] This indicates that where a controller is not able to control all processing operations and thus give effect to data subject rights, it should make sure that its relations with other network participants that may qualify as joint-controller is such that these can be safeguarded by others. Indeed, parties that act jointly 'have a certain degree of flexibility in distributing and allocating obligations and responsibilities among them, as long as they ensure full compliance'.[358] After having introduced the key actors responsible for compliance with data protection requirements, it is now turn to turn to the Regulation's substantive requirements.

---

[354] Article 29 Working Party, Opinion 05/2012 on Cloud Computing (WP 196) 01037/12/EN, 8.

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

[356] Ibid.

[357] Ibid, 4.

[358] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 24.

4.6. The implications of joint-controllership

The GDPR's broad definition of controllership has far-reaching implications for personal data processing based on DLT. The obligations of the controller have already been examined above.

Article 26 GDPR however also explicitly addressed the **consequences of a finding of joint-controllership**. It reads as follows:

> 1. Where two or more controllers jointly determine the purposes and means of processing, they shall be joint controllers. They shall in a transparent manner determine their respective responsibilities for compliance with the obligations under this Regulation, in particular as regards the exercising of the rights of the data subject and their respective duties to provide the information referred to in Articles 13 and 14, by means of an arrangement between them unless, and in so far as, the respective responsibilities of the controllers are determined by Union or Member State law to which the controllers are subject. The arrangement may designate a contact point for data subjects.
>
> 2. The arrangement referred to in paragraph 1 shall duly reflect the respective roles and relationships of the joint controllers vis-à-vis the data subjects. The essence of the arrangement shall be made available to the data subject.
>
> 3. Irrespective of the terms of the arrangement referred to in paragraph 1, the data subject may exercise his or her rights under this Regulation in respect of and against each of the controllers.

Where there is more than one controller in relation to a specific data processing operation, Article 26(2) GDPR requires that there be an **agreement between all joint-controllers** that reflects 'the respective roles and relationships of the joint controllers vis-à-vis the data subjects'.[359] On its face, this may indicate that joint-controllers are free to share the compliance burden between them as they see fit.

It is evident from the case law on joint-controllership examined above that **a finding of joint-controllership does not presuppose that the controller is able to influence all elements of the personal data processing**. Thus, the possibility of an agreement between all controllers would appear to be a welcome tool whereby joint-controllers can make sure that they are only liable to comply with GDPR requirements that they are factually able to comply with. According to this reasoning, a joint agreement under Article 26(1) GDPR could for example provide that requests for erasure cannot be directed at users where they qualify as controllers as they do not have the required influence over the network to ensure compliance with this requirement. Importantly, **the absence of such an agreement does not, however, preclude a finding that someone is a joint-controller**.[360] Indeed, Article 26(3) GDPR indicates that a data subject may exercise her GDPR rights against any controller, and this irrespective of the terms of the arrangement concluded under Article 26(1) GDPR.

The tension between the first and final paragraph of Article 26 GDPR may thus result in an allocation of responsibility to parties unable to ensure compliance in line with what has been observed above. This tension is likely to generate considerable difficulty in relation to DLT and other polycentric data processing frameworks. It has been seen above that depending on the perspective that is adopted, a large variety of actors may be considered to influence the purposes (and means) of personal data processing. Yet, this does not necessary imply that they have actual influence of related modalities

[359] Article 26(2) GDPR.

[360] Article 26(3) GDPR.

that could ensure GDPR compliance. Even though a user may be a controller, they are most likely unable to ensure the rectification or erasure of data. Similarly, nodes only see data in its encrypted form, meaning that even where they qualify as controllers they would be unable to effectively respond to a data subject's request for access under Article 15 GDPR.

Nonetheless, the Article 29 Working Party has affirmed that 'not being able to directly fulfil all controller's obligations (…) does not exclude being a controller'.[361] This results in a **situation where many entities have legal responsibilities in relation to processing operations that they cannot control**. Article 26(3) GDPR provides that data subjects may exercise their GDPR rights 'in respect of and against each of the controllers' and this even irrespective of the possible agreement between the various controllers.[362] Thus, while such an arrangement may seek to determine the practical responsibilities of each actor, each actor maintains full legal responsibility. Moreover, as per Article 82(4) GDPR each joint controller 'shall be held liable for the entire damage in order to ensure effective compensation of the data subject' although that controller may subsequently claim back some of the funds from the other controllers under Article 82(5) GDPR.

The guidance of the Working Party on the division of control has been criticised for a lack of clarity and as it 'introduces the principle that parties can have partial responsibility, but it fails to develop a consistent framework to determine the exact scope and limit of this partial responsibility' even though the GDPR allows for the full responsibility of the data controller.[363] As a consequence there is no explicit legislative basis for partial responsibility and there is no regulating framework to distribute such partial responsibility in the law.[364] In recent case law, the Court however appears to have adopted an approach that somewhat diverges from the Working Party's stance. Indeed, the Grand Chamber held in *Wirtschaftsakademie Schleswig Holstein* that

the existence of joint responsibility does not necessarily imply equal responsibility of the various operators involved in the processing of personal data. On the contrary, 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[365]

How precisely such responsibility ought to be distributed remains, however, largely uncertain. In *Google Spain*, the Grand Chamber stated that the data controller 'must ensure, within the framework of its responsibilities, powers and capabilities' that the data subject's rights are complied with in an effective and complete manner.[366] This would imply that controllers are only liable to comply with the GDPR where they are capable of doing so. Although the Court in that case didn't make explicit whether Google was the sole controller or a joint-controller, this reasoning can make sense in situations where there are many controllers and there will always be one of them that is factually able to comply with data protection law. Where there is only one controller this reasoning amounts to an acceptance that data protection requirements will in some circumstances simply not be complied with.

[361] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 22.

[362] Article 26(3) GDPR (my own emphasis).

[363] Mahieu R et al (2018), 'Responsibility for Data Protection in a Networked World. On the question of the controller, "effective and complete protection" and its application of data access rights in Europe' 14 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3256743.

[364] Ibid.

[365] Case C-210/16 *Wirtschafsakademie Schleswig-Holstein GmbH* [2018] EU:C:2017:796, para 43.

[366] Case 131/12 *Google Spain*, para 38.

Advocate General Bobek highlighted in *FashionID* that legal protection is not necessarily more effective because more people are responsible for ensuring it.[367] There is indeed a risk that the current case law will lead to a polycentric mesh of actors and responsibilities that data subjects may find hard to navigate and which may ultimately discourage them from bringing claims. Further, in order establishing joint responsibility does not require each controller 'to have access to the personal data concerned'.[368]

This gives rise to problematic situations – in the context of blockchain technology and beyond, where data controllers are, as a matter of fact, unable to effectively ensure GDPR compliance. The current state of the law makes it moreover difficult to distinguish between data controllers and data processors. The combination of DLT's polycentric design and the current state of the law burdens the determination of data controllers in such networks.

[367] Case C-40/17 *Fashion ID GmbH* [2018] EU:C:2018:1039, Opinion of AG Bobek, para 71.

[368] Case C-25/17 *Jehovan Todistajat* [2018] EU:C:2018:551, para 69.

5. Data processors and third parties

This section briefly reflects on two other categories of actors under the GDPR, namely data processors and third parties.

Article 4(8) GDPR defines the data processor as 'a natural or legal person, public authority, agency or other body which **processes personal data on behalf of the controller**'.[369] The data processor is accordingly an entity that carries out the actual personal data processing under the instruction of the data controller, meaning that the latter and not the processor exercise determinative control over the means and purposes of processing. It is important to stress that not every personal data processing operation involves a data processor as the controller can itself carry out the processing. As such, the existence of a processor 'depends on a decision taken by the controller'.[370]

Pursuant to the Article 29 Working Party, numerous elements ought to be taken into account to determine whether someone is a data controller or processor. These include (i) the level of prior instructions received from the data controller (which determines the margin of manoeuvre left to the data processor), and (ii) the data controller's monitoring of the execution of the service. Indeed, a constant and careful supervision by the controller 'provides an indication that the controller is still in full and sole control of the processing operations'; and (iii) the 'visibility and image' given by the controller to the data subject as well as the expectations the data subject has on the basis of such visibility'.[371] In some cases, it may also be appropriate to take into account the traditional role and professional expertise of the service provider, which may entail its qualification as a data controller.[372]

The processor has a **limited number of obligations** under the GDPR. Pursuant to Article 30(2) GDPR, the processor (and, where applicable, its representative) shall maintain a record of 'all categories of processing activities carried out on behalf of the controller.[373] This should contain (i) the name and contact details of the processor or processors as well as of each controller on behalf of which they are acting (and, where applicable, the controller or processor's representative and data protection officer).[374] Under certain circumstances, the processor must also designate a data protection officer.[375] The established records should reflect the categories of processing that are carried out on behalf of the controller, and where applicable, transfers of personal data to third countries or international organisations.[376] Where possible, there should also be a general description of the 'technical and organisational security measures' that are referred to in Article 32(1) GDPR.[377] These records shall be 'in writing, including in electronic form'.[378]

It is moreover the duty of the controller or processor (and, where applicable, their representative) to make these records available to the supervisory authority on request.[379] Where a data breach has occurred, the processor must moreover notify the controller 'without undue delay' after becoming

[369] Article 4 (8) GDPR.
[370] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 1.
[371] Ibid, 28.
[372] Ibid, 24.
[373] Article 30(2) GDPR.
[374] Article 30(2)(a) GDPR.
[375] Article 37 GDPR.
[376] Article 30(2)(b) and (c) GDPR.
[377] Article 30(2)(d) GDPR.
[378] Article 30(3) GDPR.
[379] Article 30(4) GDPR.

aware of the breach.[380] The above requirements do not, however, apply to entities that employ fewer than 250 people 'unless the processing it carries out is likely to result in a risk to the rights and freedoms of data subjects, the processing is not occasional, or the processing includes special categories of data'.[381] Beyond these specific obligations, Recital 30 GDPR also requires that the processor 'should assist the controller, where necessary and upon request, in ensuring compliance with the obligations deriving from the carrying out of data protection impact assessments and from prior consultation of the supervisory authority'.[382]

To determine the whether there is a processor in relation to a specific personal data processing operation that relies on DLT, a detailed case-by-case assessment must be carried out. In some scenarios, the existence of a data processor is likely such as where a company or public authority make use of an external service provider's blockchain infrastructure. If the infrastructure is used in accordance with the procurer's wishes, the latter would be seen to determine the means and purposes and processing, meaning that the external provider is merely a data processor.[383] Moreover, users 'may be both data controllers, for the personal data that they upload to the ledger, and data processors, by virtue of storing a full copy of the ledger on their own computer'.[384]

Examples of data processors include data warehouses of out-sourcing agencies, cloud providers or those providing software, platform or infrastructure as a service ('SaaS', 'PaaS' or 'IaaS').[385] Internet Service Providers providing hosting services are also processors.[386] Should the ISP however decide to further process such personal data for its own purposes, it would become a controller.[387] By implication, it seems **likely that companies offering blockchain as a service ('BaaS') also likely qualify as data processors**.

To determine what other actors using blockchain may qualify as data processors, it must first be determined who qualifies as a controller, a determination which, as observed above, is far from straightforward. Depending on the circumstances specific to each case, the operators of blockchain infrastructure could qualify as controllers where external applications make use of this infrastructure for their own operations and it is the applications that exercise decisive influence over the means and purposes of processing.

To illustrate, the French Data Protection Authority has opined that **software developers** may qualify as data processors or data controller depending on the specific role they assume when determining the purposes of processing.[388] The CNIL considers that where a smart contract developer processes personal data on behalf of a controller, such as where it offers a technical solution to a given company.[389] Further, where multiple companies decide to together run a

[380] Article 33(2) GDPR.

[381] Article 30(5) GDPR.

[382] Recital 95 GDPR.

[383] This would also imply the need for a contract to be concluded between both parties to govern their respective responsibilities.

[384] European Parliament, Report on Blockchain: a Forward-Looking Trade Policy (AB-0407/2018) (27 November 2018), para 22.

[385] Edwards L (2018), 'Data Protection I: Enter the GDPR', in Lilian Edwards (ed) Law, Policy and the Internet Oxford: Oxford University Press 81.

[386] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 25.

[387] Ibid.

[388] Commission Nationale Informatique et Libertés (September 2018), *Premiers Éléments d'analyse de la CNIL :* Blockchain, 2 https://www.cnil.fr/sites/default/files/atoms/files/la_blockchain.pdf.

[389] Ibid, 3.

blockchain for their processing operations, they may decide that only one of them is a data controller, meaning that all others become data processors.[390]

Due to the functional criteria relied on to determine who qualifies as a processor, there may be processors that are presently unaware of qualifying as such. It is true that **the GDPR requires that there be a contract or other legal act between the controller and the processor(s)**.[391] Whereas such an agreement is needed, the controller-processor relation can exist even in its absence, in line with the GDPR's functional approach to responsibility. The existence of a contract is indeed 'neither constitutive nor decisive' for the existence of a controller-processor relationship.[392] The latter is rather established on the basis of 'the factual elements of the relations between the different subjects and the way purposes and means of the processing are determined'.[393] Where a controller-processor relation is found to exist on the basis of these criteria, the parties must conclude a contract *a posteriori*.[394]

It has been pointed out that the requirement to establish **contractual relations** between controllers and processors can be tricky if one considers the large number of participants (users, nodes and miners) in **public and permissionless blockchains**, particularly since these actors would generally not know another or have established channels of communication. In such circumstances, standard-form terms and conditions that set out the parties' respective legal obligations would need to be agreed to whenever someone first uses the platform.[395] The difficulty here resides in the fact that in public and permissionless networks, core developers (and arguably also miners) are usually the only loosely associated group that could do this, yet they are likely not to be controllers in line with the analysis above. Nonetheless, they may have incentives to promote the use of their platform, and 'find that designing it to enable compliance attracts more miners, nodes and users'.[396] The core developers could then require nodes and miners to agree to these contractual terms when they download or update the software.[397]

There is nonetheless a limitation inherent to this suggestion considering that even in such circumstances, it remains possible for users to use the infrastructure without agreeing to such contractual terms, such as where users do not directly interact with the software. Here, require user-facing intermediaries (such as wallet providers and crypto-asset exchanges) could get users to agree to the platform's terms and conditions during sign-up'.[398] There can also be circumstances where a party intervenes in data processing without being a controller. This is the case of the so-called 'third parties'.

The GDPR recognizes that there may be parties that intervene in data processing but to a degree not significant enough to be a controller or a processor. These are the **third parties** referred to under **Article 4(10) GDPR**, namely 'a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorised to process personal data'.[399]

---

[390] Ibid.

[391] Article 28(3) GDPR.

[392] Article 29 Working Party, Opinion 1/2010 on the concepts of "controller" and "processor" (WP 169) 00264/10/EN, 27.

[393] Ibid.

[394] Ibid.

[395] Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' Richmond Journal of Law and Technology 1, 74.

[396] Ibid, 74-75.

[397] Ibid, 74-75.

[398] Ibid, 75.

[399] Article 4(10) GDPR.

It has been suggested in the context of cloud computing that infrastructure cloud providers (that is to say providers of pure computer processing power) that do not share any data, and utility storage providers (which provide no substantive user applications) should not be considered as data processors. They lack knowledge of the nature of data stored and have no practical ability of accessing such data and should thus be exempted from GDPR obligations.[400] Depending on the specific circumstances of each use of blockchain technology, third parties could also form part of the actors contributing to personal data processing.

[400] Kuan Hon W et al, 'Who is Responsible for 'Personal Data' in Cloud Computing? The Cloud of Unknowing, Part 2' (2011) *Queen Mary School of Law Legal Studies Research Paper* No. 77, 22.

6. Key principles of personal data processing

Article 5 GDPR announces a number of central and overarching principles that must be respected whenever personal data is processed. First, the principles of **lawfulness, fairness and transparency** require that 'personal data shall be 'processed lawfully, fairly and in a transparent manner in relation to the data subject'.[401] Second, the principle of **purpose limitation** requires that personal data be 'collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes'.[402] Third, the principle of **data minimisation** mandates that personal data be 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed'.[403]

Fourth, the principle of **accuracy** establishes that personal data ought to be 'accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay'.[404] The principle of **storage limitation** provides that personal data must be 'kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed'.[405] Pursuant to the **integrity and confidentiality** requirement, data ought to be 'processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures'.[406]

It is the **responsibility of the data controller** to comply with, but also to be able to demonstrate compliance with these various requirements.[407] This section examines how these key principles of personal data processing can be met where DLT is the chosen processing technology by examining the various components of Article 5 GDPR.

6.1. Legal grounds for processing personal data

Personal data processing can only be lawful where there is a legal ground that permits such processing.[408] In accordance with **Article 6 GDPR**, there are various different grounds of lawful personal data processing that may be more or less suitable for a specific processing operation depending on the given circumstances.[409] Data controllers must thus make sure that one of these grounds applies before they can proceed with any specific processing operation.[410] The grounds of lawful processing provided in this list are exhaustive, meaning that Member States cannot add additional grounds or otherwise amend the scope of the six principles explicitly recognised by the GDPR. Below, the various grounds of lawful personal data processing are introduced in turn.

[401] Article 5(1)(a) GDPR.
[402] Article 5(1)(b) GDPR.
[403] Article 5(1)(c) GDPR.
[404] Article 5(1)(d) GDPR.
[405] Article 5(1)(e) GDPR.
[406] Article 5(1)(f) GDPR.
[407] Article 5(2) GDPR.
[408] It is worth noting that different principles apply to instances where special categories of data are processed. These are not examined here.
[409] To illustrate, Article 6(1)(b) GDPR can only be relied on where there is a contractual relationship between the data controller and the data subject.
[410] Article 28(3) GDPR dispenses processors from independently verifying whether controllers have such a lawful ground.

6.1.1. Consent

Personal data can be processed where the data subject has provided **consent** to such processing.[411] **Article 4(11) GDPR** defines consent as 'any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her'.[412]

Whereas there are **no specific requirements of form** regarding the provision of consent (it can be provided electronically, orally or in written form), silence or pre-ticked boxes are not acceptable forms of consent.[413] Consent should moreover cover all processing activities carried out for the same purpose(s), meaning that where there are multiple purposes 'consent should be given for all of them'.[414] Moreover, where consent is provided in the context of a written declaration that also concerns other matters, the request for consent 'shall be presented in a manner which is clearly distinguishable from the other matters'.[415] Where consent is provided by electronic means, 'the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided'.[416] Furthermore, consent can only be considered as freely given if the data subject has a genuine and free choice and is able to refuse or withdraw consent without detriment.[417] This led the ECJ to find that consent cannot be used as a legal ground enabling fingerprinting for passports as holding a passport is essential for citizens wanting to travel internationally.[418]

The GDPR also provides that consent can only be informed where the purpose of processing and the controller's identity are known to the data subject.[419] It falls on the data controller to prove that consent was lawfully given.[420] This underlines the importance of clearly being able to determine controllership in line with what was observed above. It is worth stressing that the GDPR requires that consent be '**explicit**' where special categories of data are processed, where personal data is transferred to a third country in the absence of an adequacy decision or appropriate safeguards or where the solely automated processing of personal data is based on Article 22(2)(c) GDPR.[421]

Some have suggested that consent be used to enable personal data processing through DLT, and even that a user signing up for a Bitcoin address may have 'implicitly consented to the processing of that address for transaction purposes'.[422] There are, however, two problems with such statements. First, the GDPR requires that consent be provided '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' – raising the question of the compatibility of any 'implicit' form of consent.[423] Second, once personal data is included in one of the blockchain's blocks, it will continue to be indirectly processed for as long as the ledger exists. The Regulation, however, foresees that the data subject has the right to '**withdraw his or her consent** at any

[411] Article 6(1)(a) GDPR.

[412] Article 4(11) GDPR. See also Recital 32 GDPR.

[413] Recital 32 GDPR.

[414] Recital 32 GDPR.

[415] Article 7(2) GDPR.

[416] Recital 32 GDPR.

[417] Recital 42 GDPR.

[418] Case C-291/12 *Michael Schwarz* [2013] EU:C:2013:670, para 32.

[419] Recital 42 GDPR.

[420] Article 7(1) GDPR and Recital 41GDPR.

[421] See respectively Article 9(2)(a) GDPR, Article 49((1)(a) GDPR and Article 22(2)(c) GDPR.

[422] Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' *Richmond Journal of Law and Technology* 1, 73.

[423] Recital 32 GDPR.

time'.[424] Whereas this action does not affect the lawfulness of prior processing, there is a need for a new ground of processing should the data controller wish to continue processing this data.[425] If not, the processing has to be stopped. As a consequence, unless mechanisms are implemented that can halt the processing operation in the event the data subject withdraws consent, Article 6(1)(a) GDPR is thus not a suitable ground for personal data processing on blockchains. Importantly, consent is however in no way the only or main ground of lawful personal data processing under the Regulation.

6.1.2. Contract

Personal data processing is also lawful where it is necessary 'for the **performance of a contract** to which the data subject is party or in order to take steps at the request of the data subject prior to **entering into a contract**'.[426] Where a service provider such as a bank uses blockchain technology to execute their contractual obligations towards a client they accordingly have a lawful basis for processing. It follows that where a distributed ledger is used in the context of existing formalised commercial or professional relationships (such as in a supply chain setting or where a blockchain is used for accounting purposes between many actors), the existing contractual agreements between parties can also govern the use of DLT for related personal data processing.

6.1.3. Compliance with a legal obligation

Processing can also occur where it is 'necessary for **compliance with a legal obligation to which the controller is subject**'.[427] For instance, personal data is regularly processed to comply with Know Your Customer and Anti-Money Laundering requirements, which are indeed imposed by law.[428] In the blockchain context, this may for instance be relevant for cryptoasset transactions that require compliance with AML and KYC requirements, or alternatively, where the processing of certain forms of personal data is required for compliance with tax law.

6.1.4. The protection of the vital interests of the data subject or another natural person

Personal data can also be processed where it is 'necessary in order to protect the vital interests of the data subject or of another natural person'.[429] This criterion is unlikely to be of particular relevance for most contemporary DLT uses, or to cause any particular complications in such contexts compared to other tools used to process personal data.

6.1.5. Carrying out a task in the public interest or the exercise of official authority

Under EU data protection law, personal data can be processed where this is 'necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller'.[430] This is again unlikely to be of particular relevance in the context of DLT to merit more detailed examination here.

---

[424] Article 7(3) GDPR.

[425] Article 7(3) GDPR.

[426] Article 6(1)(b) GDPR.

[427] Article 6(1)(c) GDPR.

[428] Directive (EU) 2018/843 of the European Parliament and of the Council of 30 May 2018 amending Directive (EU) 2015/849 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing, and amending Directives 2009/138/EC and 2013/36/EU.

[429] Article 6(1)(d) GDPR.

[430] Article 6(1)(e) GDPR.

6.1.6. Legitimate interests

Finally, personal data can be lawfully processed where this is necessary 'for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child'.[431]

Personal data processing can thus be carried out where this is 'necessary' from the perspective of the controller or a third party, except where these actors' interests are overridden by the interests of fundamental rights and freedoms of the data subject. As a consequence, **a balancing between the interests of the data controller and of the data subject** becomes necessary.[432]

In *Bavarian Lager*, the ECJ suggested that where the privacy of the data subject is materially affected, the interests of the company must give way.[433] In *Google Spain*, the Grand Chamber spoke of the need for a fair balance which requires that it wasn't enough that the operator had an economic interest in the processing but that moreover, there was a 'legitimate interest of internet users potentially interested in having access to that information'.[434] In addition, the Grand Chamber specified that the data subject's rights 'override, as a rule, (…) the economic interests' of the data controller.[435] This highlights that the balancing that ought to take place in this respect is in fact a **weighted balancing** based on an assumption that the data subject's interest in having their fundamental rights protected primes over the purely economic interests of the data controller.

Pursuant to **Recital 47 GDPR**, legitimate interests may exist where there is a 'relevant and appropriate relationship between the data subject and the controller' such as where the data subject is a client or in the service of the controller.[436] Much emphasis is placed on a 'reasonableness' criterion in this respect. The existence of a legitimate interest needs to be carefully assessed 'including whether a data subject can reasonably expect at the time and in the context of the collection of the personal data that processing for that purpose may take place'.[437] Moreover, the data subject's interests are considered to override those of the controller 'where personal data are processed in circumstances where data subjects do not reasonably expect further processing'.[438]

Legitimate interests is both an **attractive and difficult ground** to render personal data processing lawful. Its attractiveness relates to its flexible and general nature, and the fact that it can be used in all circumstances, whether there is an existing contractual relationship or not. It is however also a ground that can be difficult to use in practice as it is not always clear what the data controller's legitimate interests are. The Regulation's preamble considers that personal data processing for direct marketing purposes 'may be regarded as carried out for a legitimate interest' and that processing 'strictly necessary for the purposes of preventing fraud also constitutes a legitimate interest of the data controller concerned'.[439]

Notwithstanding, the application of these guidelines can be difficult in practice – also in the context of blockchain technology. For example, the case may be made that an individual that purchases a

[431] Article 6(1)(f) GDPR.

[432] See further Case C-13/16 *Valsts policijas* [2017] EU:C:2017:336. Here, the ECJ refused a public body to rely on this ground of lawful processing – this has now been codified under Article 6(1)(f) GDPR.

[433] Case T-194/04 *Bavarian Lager* [2007] EU:T:2007:334.

[434] Case C-131/12 *Google Spain* [2014] EU:C:2014:317, para 81.

[435] Ibid, para 97.

[436] Recital 47 GDPR.

[437] Recital 47 GDPR.

[438] Recital 47 GDPR.

[439] Recital 47 GDPR.

cryptoasset may be considered to 'reasonably expect' that this involves the processing of personal data (such as the public key) beyond the cryptoasset transaction itself. In reality, it may however be unlikely that most users in fact realize that public keys are personal data and that transaction may reveal information about the data subject. To what degree that criterion ought to be accounted for is not, however, entirely clear.

6.2. Fairness

Under EU data protection law, any processing of personal data 'should be lawful and fair'.[440] Fairness is a somewhat open-ended principle under the GDPR the significance of which depends on context. It requires that personal data is processed in a manner that would be expected by data subjects, forming a sort of 'reasonableness' requirement. For instance, if a data subject would be tricked into providing personal data this would unlikely to be considered fair. Whereas fairness forms an important pillar of European data protection law, and may become even more central in the context of computational intelligence, it unlikely generates any problems specific to blockchain technologies and is thus not examined in further detail here.

6.3. Transparency

The GDPR requires that it 'should be transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed'.[441] Importantly, this also requires that data subjects know 'to what extent the personal data are or will be processed'.[442] Transparency is an important building block of European data protection law. It grounds data controllers' informational duties under **Articles 13 and 14 GDPR**. This again raises the question of the nexus between responsibility and control as not all actors that may qualify as controllers may indeed be able to provide the required information. It furthermore has to be understood in relation to data subject rights as it 'affects the exercise by the data subjects of their right of access to, and right to rectify, the data being processed…and their right to object to the processing of those data'.[443]

The principle of transparency requires that 'any information and communication relating to the processing of those personal data be **easily accessible and easy to understand**, and that clear and plain language be used'.[444] This implies that the information addressed to the data subject (i) be concise, (ii) easily accessible and easy to understand, and (iii) that clear and plain language and, additionally, where appropriate, visualisation be used.[445] The necessary information can be provided in electronic form (such as through a website where it is addressed at the public).[446]

Data subjects ought moreover to be made aware specifically of the **'risks'** involved in processing.[447] Whereas the principle of transparency requires that the data subject be provided with specific information, there does not appear to be any need to inform a data subject of the specific technical infrastructure that is used to process their personal data. As such, there is **no requirement to inform data subjects that DLT would be used**, only what personal data is processed and what risks arise.

[440] Recital 39 GDPR.
[441] Recital 39 GDPR.
[442] Recital 39 GDPR.
[443] Case C-201/14 *Bara* [2015] EU:C:2015:638, para 33.
[444] Recital 39 GDPR.
[445] Recital 58 GDPR.
[446] Recital 58 GDPR.
[447] Recital 39 GDPR.

It is furthermore important to note that there is a link between transparency and the principle of **purpose limitation** discussed just below as '[w]hen the specified purpose is visible and shared with stakeholders such as data protection authorities and data subjects, safeguards can be fully effective'.[448]

There appear to be no specific technical limitations to comply with transparency requirements in relation to blockchains. Indeed, it will be seen below that, provided they are properly designed, blockchains may even aid to achieve these requirements. However, a case-by-case analysis of a specific project may reveal that the use of this technology may pose **specific data protection risks** that the data subject ought to be made aware of. Furthermore, where the **specific governance arrangements** prevent the transparent designation of the controller, or where there are no channels of communication between the controller or data subjects, such as where a simple node with no relationship to other parties in the blockchain network that only has access to encrypted data is qualified as a data controller, transparency requirements may be hard to comply with in DLT contexts.

6.4. Purpose limitation

Pursuant to the principle of purpose limitation enshrined in **Article 5(1)(b) GDPR** data shall be:

> collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes[449]

The role of the purpose limitation principle is to 'prevent the use of individuals' personal data in a way (or for further purposes) that they might find unexpected, inappropriate, or otherwise objectionable'.[450] The principle has two components, first purpose specification, and second, compatible use.[451] Pursuant to the **purpose specification principle**, personal data must only be collected for 'specified, explicit and legitimate purposes' whereas the **compatible use requirement** mandates that personal data shall not be 'further processed in a manner that is incompatible with those purposes'.[452]

Regarding blockchain technology, a particular question emerges regarding the GDPR's purpose limitation requirement, namely **whether the further processing of data added to blocks after the execution of the transaction for which it was originally added to the ledger can be considered to be compatible with the purpose limitation principle**. In light of the append-only nature of these databases, data will always continue to be processed once it is on the ledger. For example, where personal data is used on a blockchain (whether in the form of the public key or transactional data) to execute a cryptoasset transaction, that data will continue to be processed even after that transaction has been successfully completed in the sense that it remains stored on the ledger, and continues to be processed pursuant to the modalities of the used consensus algorithm. From the perspective of the purpose limitation requirement, the question emerges whether the storage and subsequent involvement of such data in other transactions can be considered as part of the original purpose of processing, or whether it is necessarily incompatible

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

[449] Article 5(1)(b) GDPR.

[450] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 11.

[451] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 3.

[452] Ibid, 3-4.

with the purpose limitation principle. This issue will be addressed having regard to both the purpose specification principle and the compatible use requirement.

6.4.1. Blockchains and the purpose specification principle

Pursuant to **Article 5(1)(b) GDPR**, the purpose limitation principle requires that the controller communicate the purposes for which data is processed, and that the purpose be made explicit and legitimate. Purpose specification can be broken down into three distinct requirements.

First, the purpose of personal data processing ought to be **specified** – that is to say that it must be 'sufficiently defined to enable the implementation of any necessary data protection safeguards, and to delimit the scope of the processing operation'. The purpose must be specified in line with the above requirements at the latest at the time of personal data collection and it must further be of sufficient detail to be useful to the data subject.[453] Second, the purpose must also be **explicit**, meaning that the purpose 'must be sufficiently unambiguous and clearly expressed'. The explicitness criterion requires that the purposes 'must be clearly revealed, explained or expressed in some intelligible form'.[454]Thirdly, the purpose of personal data processing ought to be **legitimate**. Here, the notion of legitimacy is not limited to the need for a legal ground for processing but rather requires that processing occurs in line with 'broader legal principles of applicable law' such as non-discrimination.[455]

Data controllers relying on blockchain technology should thus clearly communicate to the data subject that they are using this technology and explain related implications such as that the processing is not limited to the original transaction but that their personal data will continue to be processed thereafter. It is important to stress that while a disclosure along these lines could comply with the specificity and explicitness requirements, it wouldn't necessarily render the processing legitimate. Rather, a case-by-case analysis is needed to evaluate whether the fact that data continues to be processed past the initial transaction does not stand in the way of complying with other GDPR requirements (such as the right to erasure and the data minimisation requirement) and whether compliance with other applicable legal principles such as the non-discrimination requirement can be guaranteed. Where this is not the case, even a specific and explicit disclosure of the implications of using DLT for personal data processing would fall short of GDPR compliance.

6.4.2. Blockchains and the compatible use requirement

The compatible use requirement mandates that **personal data shall not be further processed in a manner that is incompatible with the legitimate purposes that have been communicated to the data subject in an explicit manner**. It is accordingly opportune to enquire whether even where this is not explicitly communicated to the data subject, personal data may be processed on a blockchain past the initial transaction.

The fact that personal data is processed for a different purpose from that originally communicated to the data subject 'does not necessarily mean that it is automatically incompatible: this needs to be assessed on a case-by-case basis'.[456] To determine the **compatibility of the initial purpose and further processing**, a substantive assessment should prevail over a merely formal analysis.[457] The relevant circumstances taken into account to make this determination include (i) the relationship

[453] Recital 39 GDPR. See also Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 15-16.

[454] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 17.

[455] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 12.

[456] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 21.

[457] Ibid, 21-22 and 23 ('The focus should rather be on the substance of the relationship between the purposes of collection and the purposes of further processing').

between the purposes for which personal data was collected and further processed; (ii) the context in which personal data has been collected and the reasonable expectations data subjects may have concerning further use; (iii) the nature of personal data and the impact further processing has on the data subject, and (iv) the safeguards adopted by the data controller to ensure fair processing and to prevent undue impacts on data subjects.[458]

It is again imperative to recall that each use of blockchain technology for personal data processing must be assessed on a case-by-case basis to determine its compatibility with the GDPR in general and purpose limitation more specifically. In particular criteria (ii)-(iv) are highly context specific. Regarding the first criterion, that of the relationship between the initial collection of personal data and further processing, there is however the general question, introduced above, of whether there is a clear linkage of purpose of a single blockchain-based transaction and the continued storage and in the ledger.

Regarding the relationship between the various processing operations, regard should be had to whether further processing was in one form or another implied in the original purpose of processing. The Article 29 Working Party considered that a relevant factor is what 'a reasonable person in the data subject's position would expect his or her data to be used for based on the context of the collection'.[459] In any event the nature of the contract and the relation between the data subject and the data controller ought to be considered.[460] This might indicate that **there may be circumstances where processing following a blockchain-based transaction can be considered to be covered by the purpose limitation principle**.

It is worth highlighting, however, that when considering the impact of further processing on data subject, the **public disclosure of personal data** – as would be the case on a public and permissionless blockchains is a relevant factor.[461] In case of doubt, controllers might decide to rely on consent as a basis for processing as where the data subject has provided consent, 'the controller should be allowed to further process the personal data irrespective of the compatibility of the purposes'. [462] Note, however, the difficulties regarding the revocation of consent on distributed ledgers that were discussed above.

It is important to note by way of conclusion that personal data processing that is incompatible with the requirements of Article 5(1)(b) GDPR cannot simply be legitimised through reliance on an alternative legal ground under Article 6 GDPR.[463] The close link between the purpose limitation principle, and the uncertainties regarding **anonymisation** (highlighted above) should also be stressed as 'the data controller wishes to retain personal data once the purposes of the original processing have been achieved, anonymisation techniques should be used so as to irreversibly prevent identification'.[464] This again emphasizes the need to bring further clarity to the borders between personal data and anonymous data that was introduced above. The same observation can be made in relation to the data minimisation principle.

[458] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 23-26.

[459] Ibid, 24.

[460] Ibid.

[461] Ibid, 26.

[462] Recital 50 GDPR.

[463] Article 29 Working Party, Opinion 03/2013 on purpose limitation (WP 203) 00569/13/EN, 3.

[464] Ibid, 7.

6.5. Data minimisation

Pursuant to the principle of data minimisation, **data ought to be 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed'**.[465] This means that only those data necessary for the controller's purpose are obtained and processed. Two characteristics of blockchains may be a cause of particular concern from a data minimisation perspective. First, the **ever-growing** nature of such databases causes concern. Indeed, in distributed networks, data can only be removed or altered in extraordinary circumstances – meaning that obsolete data cannot be removed. Second, the **replicated nature** of data as in such distributed networks, each node stores (in principle) a full copy of the database, leading to the manifold replication of the relevant personal data.[466]

In principle, this technology may thus appear to fall foul of the data minimisation imperative. However, whether this is really the case turns on the interpretation of a number of elements. First, **Article 5(1)(c) GDPR** requires that data ought to be adequate, relevant and limited to what is necessary 'in relation to the purposes for which they are processed'. **Recital 39 GDPR** moreover specifies that personal data 'should be processed only if the purpose of the processing could not reasonably be fulfilled by other means'.[467]

This takes us back to the discussion regarding the **appropriate interpretation of the purpose limitation principle** and its application to distributed ledgers. Indeed, if the purpose includes not only the initial transaction but also subsequent processing then arguably the replicated nature of these distributed databases and the continuous storage of data could be considered to be in line with purpose limitation. Second, there is also a debate to be had regarding the correct interpretation of the data minimisation principle in general. This principle appears to be conventionally understood to relate to the **quantity of data**. From this perspective, blockchains do not appear as a technology that can easily be squared with data minimisation. An alternative interpretation would be that data minimisation is not so much about the quantity but rather the **quality of data** meaning that what would be required is that there is no processing of special categories of data unless absolutely necessary and that data is pseudonymised or even anonymised whenever possible.[468] This, however, appears unlikely in light of the formulation of **Article 25(2) GDPR** which requires that the controller 'shall implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility'.

It is recommended that further regulatory guidance be issued on the application of the data minimisation principle's application to blockchain technology which also includes guidance on these specific points. One topic that should be addressed in this context is to what extent the **off-chain storage** of personal data may be a means of achieving the data minimisation requirement.

6.6. Accuracy

The accuracy requirement enshrined in **Article 5(1)(d) GDPR** mandates that personal data be 'accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that

[465] Article 5(1)(c) GDPR.

[466] In some blockchains, there are also nodes that only store those parts of the data that are relevant to them.

[467] Recital 39 GDPR.

[468] Of course, one can interpret the data minimisation principle as relating to both quality and quantity of personal data.

personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay'.[469]

This requirement can be problematic considering blockchains' tamper-evident nature as a consequence of which data can only be removed or changed in the most extraordinary circumstances unless the specific governance arrangements provide otherwise. As this relates closely to the right to modification under **Article 26 GDPR**, this issue will be examined in further detail below. Similarly, the reference to the 'purposes' for which data are processed under Article 5(1)(d) GDPR also raises important questions in relation to the purpose limitation principle, which was already examined above.

6.7. Storage limitation

In accordance with the principle of storage limitation, personal data should be:

> kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed; personal data may be stored for longer periods insofar as the personal data will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) subject to implementation of the appropriate technical and organisational measures required by this Regulation in order to safeguard the rights and freedoms of the data subject[470]

Article 5(1)(e) GDPR mandates that no obsolete data be retained. To ensure that personal data is not kept longer than necessary, 'time limits should be established by the controller for erasure or for a periodic review'.[471] Moreover, '[e]very reasonable step should be taken to ensure that personal data which are inaccurate are rectified or deleted'.[472] The storage limitation imperative raises the **question of when data stored on DLT becomes obsolete**, in line with the observations made above in relation to purpose limitation. This could be interpreted to relate to the completion of the relevant transaction, or it could be argued that even after this event, data is still 'necessary' for subsequent processing, in this case the continued storage of personal data on the ledger as well as its processing in the context of the given consensus protocol. As amply discussed above, data can generally only be removed from blockchains in the most extraordinary circumstances, raising the question of whether the storage limitation can be respected in such environments.

It is also worth recalling that the GDPR only applies to personal data. This implies that where data continues to be processed past the initial purpose but only in an anonymised state (noting the high threshold for **anonymisation** discussed above) then this processing no longer falls within the Regulation's scope. **Sharding** could be a useful mechanism in this respect. It has been suggested in relation to cloud computing that where the user could be the only person to access reunified shards of their stored data (where they can log into their account with the provider but no one else can or has otherwise access to the data) then, at least pursuant to a relative approach to the concept of personal data, 'the data may be 'personal data' to the user, but not to anyone else'.[473] This, however, hinges on the question from whose perspective the quality of data ought to be assessed, which was discussed above.

---

[469] Article 5(1)(d) GDPR.

[470] Article 5(1)(e) GDPR.

[471] Recital 39 GDPR.

[472] Recital 39 GDPR.

[473] Millard C (2013) Cloud Computing Law Oxford: Oxford University Press 182.

Beyond anonymisation, other forms of data handling may be considered to be equivalent to erasure for the purpose limitation principle. Indeed, the British ICO has recognised that erasing data from a digital system is not always a straightforward matter. It has suggested that putting personal data 'beyond use' may be an alternative to achieve data minimisation.[474] If there is no intention on behalf of the controller to access or use the data so that it 'is no longer live' then this may be held to be compliant.[475] This would be the case where (i) the data controller is unable or will not attempt 'to use the personal data to inform any decision in respect of any individual or in a manner that affects the individual in any way'; (ii) does not provide third parties with access to the data; (iii) 'surrounds the personal data with appropriate technical and organisational security; and (iv) commits to permanent deletion of the information if, or when, this becomes possible'.[476] It is, however, unclear whether this approach will satisfy data protection authorities in other Member States. As a consequence, regulatory guidance on this matter would provide more clarity to actors in this domain.

6.8. The accuracy principle

The principle of accuracy enshrined in **Article 5(1)(d) GDPR** requires that data be accurate, and where necessary kept up to date.[477] The principle that all inaccurate personal data be rectified or erased relates closely to the data subject rights to rectification and erasure enshrined respectively in Articles 16 and 17 GDPR, which are examined in detail just further below.

6.9. The integrity and confidentiality principle

Article 5(1)(f) GDPR requires that personal data be processed in a manner that ensures its **security** 'including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures'. This is a general obligation that does not appear to generate any particular problems where blockchains are used.

6.10. Accountability

Article 5(2) GDPR mandates that the data controller is responsible for, and should be able to demonstrate compliance with, the requirements under Article 5 GDPR (examined just above). This relates to the duties of the controller, which have already been examined above.[478]

[474] Information Commissioner's Office (26 February 2014), 'Deleting Personal Data' https://ico.org.uk/media/for-organisations/documents/1475/deleting_personal_data.pdf 4.

[475] Ibid.

[476] Ibid, 5.

[477] Article 5(1)(d) GDPR.

[478] See also Article 24 GDPR.

7. Data subject rights

Articles 15 to 22 GDPR allocate numerous specific rights to data subjects. Data controllers are obliged to facilitate the exercise of these rights and cannot delegate this task to processors.[479] The GDPR's various data subject rights are examined in turn below. It will be seen that some do not raise any specific problems in the context of blockchain technology whereas others trigger both technical and legal challenges, the possible solutions to which depend in part on the identity of the data controller and its influence over blockchain data. Of course, as always, the application of these various data subject rights to distributed ledgers can only be comprehensively assessed on the basis of a case-by-case analysis that accounts for the specific technical and contextual circumstances of each personal data processing operation.

7.1. The right to access

According to **Article 15 GDPR**

> 1. The data subject shall have 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 to the personal data and the following information:
>
> (a) the purposes of the processing;
>
> (b) the categories of personal data concerned;
>
> (c) the recipients or categories of recipient to whom the personal data have been or will be disclosed, in particular recipients in third countries or international organisations;
>
> (d) where possible, the envisaged period for which the personal data will be stored, or, if not possible, the criteria used to determine that period;
>
> (e) the existence of the right to request from the controller rectification or erasure of personal data or restriction of processing of personal data concerning the data subject or to object to such processing;
>
> (f) the right to lodge a complaint with a supervisory authority;
>
> (g) where the personal data are not collected from the data subject, any available information as to their source;
>
> (h) the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject.
>
> 2. Where personal data are transferred to a third country or to an international organisation, the data subject shall have the right to be informed of the appropriate safeguards pursuant to Article 46 relating to the transfer.
>
> 3. The controller shall provide a copy of the personal data undergoing processing. For any further copies requested by the data subject, the controller may charge a reasonable fee based on administrative costs. Where the data subject makes the request

[479] Article 12 (2) GDPR.

> by electronic means, and unless otherwise requested by the data subject, the information shall be provided in a commonly used electronic form.
>
> 4. The right to obtain a copy referred to in paragraph 3 shall not adversely affect the rights and freedoms of others.[480]

It is no coincidence that the list of the various data subject rights starts with the right to access. Indeed, the right to access ought to be considered as a **foundational right** in European data protection law considering that it enables, and is often a prerequisite for, the exercise of any of the other rights of the data subject. Indeed, accessing personal data enables the data subject to understand what data is being processed by the data controller, which may be a necessary first step before any other right can be exercised. For instance, the right to access enables the data subject to establish that personal data may be inaccurate, which may in turn motivate them to exercise their right to rectification under Article 16 GDPR. Article 15 GDPR is thus an enabling right that is of much significance for the overall structure of European data protection law.

Where a request for access is made by a data subject, the controller ought to search all of its (electronic and paper-based) records to provide related information to the data subject. Thus, where a data controller relies on DLT to process personal data in isolation or together with other means, they must enquire whether this database contains information regarding the data subject. As a general matter, there are **no principled hurdles why Article 15 GDPR could not be implemented regarding blockchains**. This, however, presupposes the existence of adequate governance mechanisms that enable effective communication and data management.

Requests for access can be addressed by the data subject to the data controller or, pursuant to **Article 26(3) GDPR** to any of the joint-controllers. It has however been pinpointed above that at least some of the entities that may qualify as (joint-) controllers under the GDPR may be factually unable to access data on the blockchain. For instance, nodes in principle only see encrypted and hashed data where data has been thus modified when put on the blockchain. As a consequence, such actors may be unable to determine whether the distributed ledger indeed contains personal data relating to the data subject that initiates the right to access. Similar problems arise in respect of the requirement that the controller shall provide a copy of the personal data undergoing processing to the data subject under Article 15(3) GDPR.[481] As a consequence, actors deciding to use blockchain technology to process personal data must make sure that there are **appropriate governance arrangements** in place that enable the effective exercise of this right.

7.2. The right to rectification

Pursuant to **Article 16 GDPR**

> The data subject shall have the right to obtain from the controller without undue delay the rectification of inaccurate personal data concerning him or her. Taking into account the purposes of the processing, the data subject shall have the right to have incomplete personal data completed, including by means of providing a supplementary statement[482]

Blockchains are an append-only ledger often purposefully designed to render the deletion and modification of data extraordinarily burdensome in order to secure data integrity and trust in the network. This fundamentally stands in tension with the GDPR's requirement that data be mutable

[480] Article 15 GDPR (my own emphasis).

[481] Article 15(3) GDPR.

[482] Article 16 GDPR.

in order to allow for its erasure, or, as required by Article 16 GDPR, its rectification. Blockchains indeed can often not support reversibility, for instance where a customer asks a service provider using blockchain to rectify information in their record.[483]

Private and/or permissionless blockchains can support such requests through an alteration of the relevant transaction record by re-hashing subsequent blocks where this is facilitated by the respective technical and governance set-up.[484] Rectifying data on **public and/or permissionless blockchains** is, however, much more difficult and individual actors are not in a position to comply with such requests. This is not because it is strictly impossible from a technical perspective to do so, much to the contrary as every single node can alter its own local copy of the ledger (provided that they can identify the relevant data to be rectified as this is far from evident where the relevant data is encrypted).[485] However, even if all nodes, miners and users were considered to in fact qualify as the data controllers liable to implement data subject rights, 'this would not necessarily provide effective protection for data subjects'. [486] This is so as even though all nodes could agree (through a contract or another form of agreement) to 'fork' to a new version of the blockchain in periodic intervals to reflect requests for erasure, this level of coordination has been said to be 'difficult to achieve among potentially thousands of nodes'.[487]

Article 16 GDPR however explicitly envisages the option to have incomplete data completed 'by means of **providing a supplementary statement**'. This is much easier to implement regarding distributed ledgers as any party that has write rights can add new data to the ledger that rectifies previous information. For example, where a user's records reflect that he is single, additional data could be added to a new block to indicate that this is no longer the case after a new marriage.

It is, however, worth questioning whether the addition of new information on-chain will in all circumstances be a satisfactory means of achieving the rationale inherent in Article 16 GDPR. It is worth noting that in the *Nowak* case, Advocate General Kokott argued that the right to rectification ought to be 'judged by reference to the purpose for which the data was collected and processed'.[488] In this case, the argument was made that Article 16 GDPR could not be invoked to obtain the rectification of answers in an examination. Adopting this **purposive approach**, it appears evident that the provision of a supplementary statement might not always be a satisfactory means of achieving compliance with the right to rectification, such as where there is a strong case that data should not just be supplemented but removed and replaced (such as in scenarios where a data subject cannot rely on the right to erasure as none of the grounds in Article 17(1) GDPR apply.[489] Conversely, it could be argued that where Article 17(1) GDPR does not apply, the data subject cannot be considered to have an interest in the erasure of data and that the mere provision of additional information ought to be considered sufficient.

This is another uncertainty of general relevance to European data protection law that could be elucidated by regulatory guidance. In addition, research could explore additional means of securing the effective implementation of the right to rectification on DLT. This could focus on the one hand on technical solutions that provide an alternative to the addition of additional statements without the deletion of the original data, and, on the other, on effective governance solutions to enable

[483] Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' *Richmond Journal of Law and Technology* 1, 76.

[484] Ibid.

[485] Ibid, 77.

[486] Ibid.

[487] Ibid.

[488] Opinion of AG Kokott in Case C-434/16 *Peter Nowak* [2017] EU:C:2017:582, para 35.

[489] On the right to erasure, see further below.

coordination among the many participants in these polycentric networks in order to secure compliance.

7.3. The right to erasure (the 'right to be forgotten')

Pursuant to Article 17 GDPR,

> 1. The data subject shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay where one of the following grounds applies:
>
> (a) the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed;
>
> (b) the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1), or point (a) of Article 9(2), and where there is no other legal ground for the processing;
>
> (c)the data subject objects to the processing pursuant to Article 21(1) and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to Article 21(2);
>
> (d) the personal data have been unlawfully processed;
>
> (e) the personal data have to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject;
>
> (f) the personal data have been collected in relation to the offer of information society services referred to in Article 8(1).
>
> 2. personal data, the 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.
>
> 3. Paragraphs 1 and 2 shall not apply to the extent that processing is necessary:
>
> (a) for exercising the right of freedom of expression and information;
>
> (b) for compliance with a legal obligation which requires processing by Union or Member State law to which the controller is subject or for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;
>
> (c)for reasons of public interest in the area of public health in accordance with points (h) and (i) of Article 9(2) as well as Article 9(3);
>
> (d)for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1) in so far as the right referred to in paragraph 1 is likely to render impossible or seriously impair the achievement of the objectives of that processing; or

(e)for the establishment, exercise or defence of legal claims.[490]

The Regulation's right to erasure is an important tool towards more **informational self-determination** as it provides the data subject with control over personal data that directly or indirectly relates to them. Article 17 GDPR enables data subjects to obtain the 'erasure' of personal data from the data controller if one of the grounds listed applies. Indeed, the right to erasure is both **a qualified and a limited right**.[491] It can only be invoked subject to the conditions in Article 17(1) GDPR and must moreover be balanced against the considerations in Article 17(2) GDPR. The ECJ has moreover stressed that the right to erasure cannot be invoked in a manner that would go counter the spirit of this provision.[492]

Many have stressed the **difficulty of applying the right to erasure to blockchains**. Deleting data from DLT is burdensome as these networks are often purposefully designed to make the unilateral modification of data hard, which in turn is supposed to generate trust in the network by guaranteeing data integrity. For example, where the relevant consensus-mechanism that is used is proof-of-work, 'the majority of all P2P connected nodes would have to verify again the legitimacy of every effected transaction backwards, unbuild the entire BC block by block and then rebuild it afterwards, with every such transaction step to be distributed block-wise to all existing nodes'.[493]

The difficulty of complying with Article 17 GDPR is thus burdened by **technical factors**, but also by **governance design**. Indeed, even if there would be a means of ensuring compliance from a technical perspective, it may be organisationally difficult to get all nodes to implement related changes on their own copy of the database (particularly in public and permissionless blockchains). In order to provide further insights on the relationship between distributed ledgers and the GDPR's right to erasure this section evaluates these elements. First, attention must be drawn to the uncertain definition of the terminology of 'erasure' in Article 17 GDPR. Indeed, it is difficult to assess whether the erasure of personal data from blockchains is possible as long as there is no precise guidance as to how this concept ought to be interpreted.

7.3.1. The meaning of erasure

Before any examination of whether blockchain technology is capable of complying with Article 17 GDPR; it must be underscored that the **precise meaning of the term 'erasure' remains unclear**.

Article 17 GDPR does not define erasure, and the Regulation's recitals are equally mum on how this term should be understood. It might be assumed that a **common-sense understanding** of this terminology ought to be embraced. According to the Oxford English Dictionary, erasure means 'the removal or writing, recorded material, or data' or 'the removal of all traces of something: obliteration'.[494] From this perspective, erasure could be taken to equal destruction. It has, however, already been stressed that the destruction of data on blockchains, particularly these of a public and permissionless nature, is far from straightforward.

There are, however, indications that the obligation inherent to **Article 17 GDPR** does not have to be interpreted as requiring the outright destruction of data. In *Google Spain*, the delisting of information from research results was considered to amount to erasure. It is important to note, however, that in this case, this is all that was requested of Google by the claimant, who did not have

[490] Article 17 GDPR (my own emphasis).

[491] See further Case C-398/15 *Salvatore Manni* [2017] EU:C:2017:197.

[492] Case C-434/16 *Peter Nowak* [2017] EU:C:2017:994, para 52 (stating that Article 17 GDPR cannot be invoked to obtain the correction of incorrect exam answers).

[493] Berberich M and Steiner M (2016), 'Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?' 2 *European Data Protection Law Review* 422, 426.

[494] https://en.oxforddictionaries.com/definition/erasure

control over the original data source (an online newspaper publication). Had the claimant wished to obtain the outright destruction of the relevant data it would have had to address the newspaper, not Google. This may be taken as an indication that what the GDPR requires is that the obligation resting on data controllers is to do all they can to secure a result as close as possible to the destruction of their data within the limits of thei own factual possibilities.

National and supranational regulators have moreover indicated that there may be **alternatives to the outright destruction of data** that could secure compliance with the GDPR's erasure obligation. In its opinion on cloud computing, the Article 29 Working Party considered that the destruction of hardware could arguably qualify as erasure for the purposes of Article 17 GDPR.[495] Furthermore, national data protection authorities have considered that erasure does not necessarily equal destruction. For example, the Austrian Data Protection Authority recently recognised that the data controller enjoys flexibility regarding the technical means of realising erasure, and that **anonymisation** can be seen as a means to realise erasure.[496] Furthermore, the UK Information Commissioner's Office has long argued that where data is '**put beyond use**' this may also be satisfactory.[497] There does not, however, appear to be consensus in all Member States on this matter.

Whether these measures will be deemed satisfactory by the Court remains to be seen. It is worth highlighting that in *Nowak*, the CJEU appeared to indicate that **erasure equals the destruction of personal data**.[498] It stated that in accordance with the right to erasure, a candidate in a written examination has 'the right to ask the data controller to ensure that his examination answers and the examiner's comments with respect to them are, after a certain period of time, erased, that is to say, destroyed'.[499] Whether this can be seen as a blanket statement that erasure always amounts to destruction in unclear, especially since the case at issue did not directly deal with the right to erasure. The statement could thus also be explained by the specific context at hand and the fact that outright destruction of the examination copy may be the most straightforward means of destruction (although the blackening out of the relevant information is another obvious option).

It is hoped that future case law on this matter will shed further light on the correct interpretation to be given to the concept of erasure. In the meanwhile, regulatory guidance could add much-needed clarity to this domain. Such guidance could consider the following technical means that have been suggested as a means of giving effect to Article 17 GDPR in relation to blockchain technology.

7.3.2. Possible alternative technical means of achieving erasure on blockchains

As awareness regarding the tricky reconciliation between Article 17 GDPR and distributed ledgers grows, a number of technical alternatives to the outright destruction of data have been considered by various actors. An often-mentioned solution is that of the **destruction of the private key**, which would have the effect of making data encrypted with a public key inaccessible. This is indeed the solution that has been put forward by the French data protection authority CNIL in its guidance on

[495] Article 29 Working Party, Opinion 05/2012 on Cloud Computing (WP 196) 01037/12/EN, 12.

[496] Austrian Data Protection Authority, DSB-D123.270/0009-DSB/2018 (05 December 2018) https://www.ris.bka.gv.at/Dokumente/Dsk/DSBT_20181205_DSB_D123_270_0009_DSB_2018_00/DSBT_20181205_DSB_D123_270_0009_DSB_2018_00.html.

[497] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-erasure/.

[498] Case C-434/16 *Peter Nowak* [2017] EU:C:2017:994, para 55. The Court considered that the candidate might indeed have an interest in the erasure of her answers in a written examination where the examination period had official closed and the result could no longer be challenged so that the document has lost any probative value.

[499] Ibid, para 55.

blockchains and the GDPR. The CNIL has suggested that erasure could be obtained where the keyed hash function's secret key is deleted together with information from other systems where it was stored for processing.[500]

Beyond, the various technical solutions introduced above in the section on anonymisation should also be evaluated for their potential to achieve compliance with Article 17 GDPR. This includes redactable blockchains, which would be 'forgetful' by design but also pruning and chameleon hashes and zero knowledge proofs.[501] It is recommended below that regulatory guidance should clarify whether any of these processes may be used to achieve 'erasure' under Article 17 GDPR. Furthermore, this is also an area where further interdisciplinary research would be of much value. Some have indeed predicted that in the future there may be new avenues for 'automating aspects of reversibility, such as corrective operation that can occur automatically through the use of smart contracts'.[502]

Regulatory guidance should provide further information on whether any of these techniques may be considered to fulfil the standard of 'erasure' under Article 17 GDPR. The challenges of compliance are not limited to technical questions as also governance design influence the ability of a given use of DLT to be fashioned in a manner that's respectful of data protection law.

7.3.3. Governance challenges

Even where technical solutions to implement the right to be forgotten on DLT can be identified, successful compliance with this data subject right (and others) might prove impossible due to a lack of mechanisms of communication and coordination between the relevant actors.

Effective compliance with Article 17 GDPR can only be given where the personal data in question is **erased from all of the nodes that participate in the network**. As a matter of fact, the Article 29 Working Party considered in the cloud computing context that where personal data is 'kept redundantly on different servers at different locations, it must be ensured that each instance of them is erased irretrievably'.[503] By analogy, personal data ought to be removed from all nodes that store this data where a request for erasure is justified.

This implies that where a data subject addresses a request for erasure to a (joint-) controller, that controller must not only remove that personal data from its own servers, but also initiate erasure from other controllers and processors that are processing that personal data. Whether a given use-case of DLT is fashioned in a manner that facilitates compliance with this obligation is a matter of fact that can only be determined on the basis of a detailed case-by-case analysis. The issue nonetheless underlines the pivotal need for **adequate governance designs** of distributed ledger technology, which will also be important to ensure compliance with legal obligations in other areas.

The controller's obligation to incentivise other controllers to undertake erasure is grounded in in **Article 17(2) GDPR**, which requires that 'the 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'.[504] It is, however,

---

[500] Commission Nationale Informatique et Libertés (September 2018), *Premiers Éléments d'analyse de la CNIL :* Blockchain, 8-9 https://www.cnil.fr/sites/default/files/atoms/files/la_blockchain.pdf.

[501] Ateniese G, Magri B, Venturi D and Andrade E (2017), 'Redactable Blockchain – or – Rewriting History in Bitcoin and Friends' https://eprint.iacr.org/2016/757.pdf.

[502] Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' 25 *Richmond Journal of Law and Technology* 1, 24.

[503] Article 29 Working Party, Opinion 05/2012 on Cloud Computing (WP 196) 01037/12/EN, 12.

[504] Article 17(2) GDPR.

worth highlighting that the obligation imposed here is an obligation of means and not an obligation of ends (when it comes to the erasure – the controller's informational duty is indeed an obligation of ends).[505] Indeed all the data controller ought to do is to take 'reasonable steps' (which are assessed in light of the available technology and the cost of implementation) to inform other controllers processing the personal data that the data subject has requested erasure.

Due to the multi-layered nature of blockchains there are likely a number of joint-controllers in respect to each transaction. In such constellations, a data subject may approach any actor of the ecosystem that qualifies as a joint controller to enforce her rights. Indeed, in *Google Spain*, the data subject's action against Google was not affected by the fact that that data could have been removed by the newspaper's website.[506] By analogy, it would not be surprising if data subjects turned to **intermediaries** such as blockexplorers to seek the removal of personal data from their own index. As blockchain ecosystems develop further this may indeed be a much more efficient solution than targeting the infrastructure level, in line with why the claimant in *Google Spain* chose to address Google rather than the newspaper that had initially published the information at issue. Future interdisciplinary research could shed further light on coordination mechanisms between various data controllers in complex polycentric networks to achieve GDPR compliance.

7.3.4. Further considerations and limitations

It is worth noting that the question of the **territorial scope of the right to erasure** is of central importance to blockchains as these often have a cross-jurisdictional nature. Whereas the precise jurisdictional scope of Article 17 GDPR is presently unclear, the upcoming Grand Chamber judgment in *Google v. CNIL* should add much-needed clarity to this area of the law.[507]

It has already been stressed above that the right to erasure is both a limited and a qualified right. Article 17(1)(e) GDPR and Recital 65 GPDR furthermore clarify that data does not have to be erased where the further retention of data is necessary for **compliance with a legal obligation**. This is a relevant consideration regarding many use cases of blockchain technologies in the financial realm such as the data retention obligations under MiFID II.[508]

7.4. Right to restriction of processing

In accordance with **Article 18 GDPR**

> 1.The data subject shall have the right to obtain from the controller restriction of processing where one of the following applies:
>
> (a) the accuracy of the personal data is contested by the data subject, for a period enabling the controller to verify the accuracy of the personal data;
>
> (b) the processing is unlawful and the data subject opposes the erasure of the personal data and requests the restriction of their use instead;

[505] See also Recital 66 GDPR.

[506] Case C-131/12 *Google Spain* [2014] EU:C:2014:317, para 80. The Court emphasised that search engines made it easier for internet users to find the relevant data and played an important role in its dissemination which was 'liable to constitute a more significant interference with the data subject's fundamental right to privacy than the publication on the web page'.

[507] Case C-507/17 *Google v CNIL* (pending).

[508] Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU

> (c) the controller no longer needs the personal data for the purposes of the processing, but they are required by the data subject for the establishment, exercise or defence of legal claims;
>
> (d) the data subject has objected to processing pursuant to Article 21(1) pending the verification whether the legitimate grounds of the controller override those of the data subject.
>
> 2. Where processing has been restricted under paragraph 1, such personal data shall, with the exception of storage, only be processed with the data subject's consent or for the establishment, exercise or defence of legal claims or for the protection of the rights of another natural or legal person or for reasons of important public interest of the Union or of a Member State.
>
> 3. A data subject who has obtained restriction of processing pursuant to paragraph 1 shall be informed by the controller before the restriction of processing is lifted.

Article 18 GDPR provides a right to the data subject to obtain a restriction of the processing of personal data relating to them in a number of circumstances and notwithstanding what specific technology is used to carry out the processing.[509] As a consequence, where distributed ledgers are used, EU data protection law requires that the data subject has the possibility of obtaining a restriction of processing, such as where the data subject contests the accuracy of personal data.[510] In order to determine whether one of the many possible joint-controllers in a given blockchain network is able to comply with the requirements of Article 18 GDPR, a case-by-case analysis of the given technical and governance arrangements must be carried out. In general, two potential overarching obstacles to compliance with this obligation can be identified.

First, there are likely **technical obstacles** to the restriction of processing in contexts of automated processing, such as blockchains. Indeed, such systems are often designed to make (unilateral) intervention in the data processing burdensome in order to increase data integrity and trust in the network. Particularly in respect of public and permissionless ledgers, there are no straightforward means of halting the processing of data contained in one of the blocks. It is worth noting that this is true both in relation to the application layer (the distributed ledger itself) but that this can also be true in relation to blockchain-based applications.

Second, there are also **governance challenges** in relation to various of the potentially many joint-controllers' ability to undertake such intervention in the network. It has been seen above that pursuant to the recent case law on joint-control any party that exercises some degree of control over the means and especially the purposes of personal data processing qualifies as a joint controller. However, some of the possible data controllers such as nodes or users lack the ability of intervening in the network in a manner that would in fact be conductive to trigger a restriction of processing – recurring theme that has been highlighted in relation to all data subject rights that are examined here. This again underlines the pivotal importance of both technical and governance arrangements that would enable data controllers to effectively comply with Article 18 GDPR.

7.5. Data controllers' communication duties

Article 19 GDPR provides that

[509] Article 18(1) GDPR.

[510] Article 18(1)(a) GDPR.

> The controller shall communicate any rectification or erasure of personal data or restriction of processing carried out in accordance with Article 16, Article 17(1) and Article 18 to each recipient to whom the personal data have been disclosed, unless this proves impossible or involves disproportionate effort. The controller shall inform the data subject about those recipients if the data subject requests it.

Article 19 GDPR requires that the data controller communicate the rectification or erasure of personal data, as well as any restriction of processing to '**recipients**' to which personal data has been disclosed. This raises the question of what parties would actually qualify as 'recipients' of personal data in contexts where DLT is used. Blockchains are oftentimes presented as being at their most useful where they are used to coordinate records between many different parties, meaning that there is at least potentially a large number of such 'recipients' for each personal data processing operation on a distributed ledger.

In **private and/or permissioned systems**, a track record of parties with permission to access and read the data usually exists, meaning that data controllers keep track of the parties personal data is disclosed to. As a consequence, contacting these parties to inform them of any actions under Articles 16-18 GDPR should be relatively straightforward.

Where a blockchain is **public and/or permissionless**, no permission is needed to obtain access to the personal data stored on such ledgers. Conversely, the parties in charge of such networks have no way of knowing what parties have gained access through related personal data, either because they have directly engaged with the network or because they have relied on tools such as blockexplorers. In these circumstances, the communication duties inherent to Article 19 GDPR can be said to 'prove impossible' or at the least 'involve disproportionate' effort. As a consequence, this may be one of the scenarios envisaged by Article 19 GDPR where the data controller gains dispensation from having to comply with their notification duties.

7.6. The right to data portability

Pursuant to **Article 20 GDPR**,

> 1. The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided, where:
>
> (a) the processing is based on consent pursuant to point (a) of Article 6(1) or point (a) of Article 9(2) or on a contract pursuant to point (b) of Article 6(1); and
>
> (b) the processing is carried out by automated means.
>
> 2. In exercising his or her right to data portability pursuant to paragraph 1, the data subject shall have the right to have the personal data transmitted directly from one controller to another, where technically feasible.
>
> 3. The exercise of the right referred to in paragraph 1 of this Article shall be without prejudice to Article 17. That right shall not apply to processing necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.
>
> 4. The right referred to in paragraph 1 shall not adversely affect the rights and freedoms of others.

The right to the portability is one of the **main innovations** of the GDPR compared to the 1995 Data Protection Directive. It is essentially a tool that allows data subjects to – in some circumstances – port data from one data controller to another. The principle of personal data portability 'aims to empower data subjects regarding their own personal data, as it facilitates their ability to move, copy or transmit personal data easily from one IT environment to another'.[511] Where a data subject's request for portability complies with the requirements of Article 20 GDPR, controllers are obliged to make the data available in a 'structured, commonly used and machine-readable format'.[512] This should moreover be in an 'interoperable format'.[513]

In order for a data subject to be able to rely on Article 20 GDPR a number of **conditions** need to be met. First, this right evidently only applies to personal data. Second, the personal data in question has to be have been provided by the data subject to the data controller. Third, personal data processing is based on consent or contract. Fourth, processing is undertaken through automated means.[514] Finally **Article 11(2) GDPR** underlines that the right to data portability does not apply if the controller can demonstrate that it is not in a position to identify the data subject unless the data subject provides more information to enable identification.

The French Data Protection Authority CNIL considers that blockchain technologies raise little problems when it comes to compliance with the portability requirement.[515] Article 20 GDPR nonetheless stresses the interest in securing **interoperability** among various DLT solutions. It has been stressed in relation to social media networks that there is little point in porting data from one social media provider to another if there are no 'friends' on the second.[516] The same concerns induced by network effects also apply to the blockchain context – both at the infrastructure and application layers. The efficient enforcement of data portability is thus one of the many reasons why the interoperability of various solutions should be encouraged.[517]

Again, it is important to recall the **necessary nexus between accountability and control**. Pursuant to the current stance of the European Court of Justice on controllership, there is a risk that entities are qualified as controllers even though they are effectively unable to comply with any of the portability requirements falling on data controllers under the GDPR. Indeed, a node may qualify as a data controller although they can only access data that may be hashed or encrypted, which in turn will defeat the use of such data for data subjects in many circumstances.

7.7. The right to object

According to **Article 21 GDPR**

> 1. The data subject shall have the right to object, on grounds relating to his or her particular situation, at any time to processing of personal data concerning him or her which is based on point (e) or (f) of Article 6(1), including profiling based on those provisions. The controller shall no longer process the personal data unless the controller demonstrates compelling legitimate grounds for the processing

[511] Article 29 Working Party (5 April 2017), *Guidelines on the Right to Portability* WP 16/EN, WP 242 rev.01, 4.

[512] Article 20(1) GDPR.

[513] Recital 68 GDPR.

[514] Article 22(1) GDPR.

[515] Commission Nationale de l'Informatique et des Libertés (06 November 2018) *Blockchain and the GDPR: Solutions for a responsible use of the blockchain in the context of personal data* https://www.cnil.fr/en/blockchain-and-gdpr-solutions-responsible-use-blockchain-context-personal-data

[516] Edwards L (2018), Law, Policy and the Internet, Oxford: Hart Publishing 109.

[517] On interoperability in the blockchain context see further the related report of the Blockchain Observatory and Forum: https://www.eublockchainforum.eu/sites/default/files/reports/report_scalaibility_06_03_2019.pdf.

> which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims.
>
> 2. Where personal data are processed for direct marketing purposes, the data subject shall have the right to object at any time to processing of personal data concerning him or her for such marketing, which includes profiling to the extent that it is related to such direct marketing.
>
> 3. Where the data subject objects to processing for direct marketing purposes, the personal data shall no longer be processed for such purposes.
>
> 4. At the latest at the time of the first communication with the data subject, the right referred to in paragraphs 1 and 2 shall be explicitly brought to the attention of the data subject and shall be presented clearly and separately from any other information.
>
> 5. In the context of the use of information society services, and notwithstanding Directive 2002/58/EC, the data subject may exercise his or her right to object by automated means using technical specifications.
>
> 6. Where personal data are processed for scientific or historical research purposes or statistical purposes pursuant to Article 89(1), the data subject, on grounds relating to his or her particular situation, shall have the right to object to processing of personal data concerning him or her, unless the processing is necessary for the performance of a task carried out for reasons of public interest.

Article 21 GDPR provides a right to the data subject to object to any processing of personal data that directly or indirectly relates to them where such data is processed by the data controller on the basis of Article 6(1)(e) GDPR (public interest) or on the basis of Article 6(1)(f) GDPR (legitimate interests).

Where the data subject exercises that right, the data controller has to **stop processing** this personal data unless it is in a position to demonstrate 'compelling legitimate grounds for the processing which override the interests, rights and freedoms of the data subject or for the establishment, exercise or defence of legal claims'.[518]

Many of the points of general significance highlighted above also apply to compliance with Article 21 GDPR, such as the controllers' factual ability to influence processing due to their limited means of interfering with the data, but also the points made in relation to the ability to halt data where it is processed automatically. One point emphasising in particular in relation to Article 21 GDPR is the interpretation to be given to the 'compelling legitimate grounds' for processing that enable the data controller to not give way to a data subject request to exercise their right to a restriction of processing. In particular, one may wonder whether the data controller's interest in the integrity of DLT records may qualify as such a legitimate interest. This is again a point that could be clarified by regulatory guidance to provide more legal certainty in this domain.

7.8. Article 22 GDPR and solely automated data processing

In accordance with Article 22(1) GDPR a data subject has the right 'not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her'. This is important in the context of blockchain technology, for instance when it comes to **smart contracts** which can be considered to make

[518] Article 21 (1) GDPR.

'decisions' in some circumstances.[519] According to the Article 29 Working Party, solely automated decision-making refers to 'the ability to **make decisions by technological means without human involvement**'.[520] A decision is hence 'based solely' on automated processing where there is 'no human involvement' in the decision-making process.[521]

Article 22 GDPR however only targets 'decisions' made through solely automated data processing, whereas Recital 71 GDPR also speaks of a 'decision, which may include a measure'. There is an argument to be made that a blockchain-based **smart contract** may qualify as a decision, at least where it leads to an outcome that would be reached through a human decision-making process in the analogue world should be considered as a 'decision', which is caught by Article 22 GDPR where it produces legal or otherwise significant effects, such as where the smart contract determines whether an insurance premium is paid, consumer rights are enforced or payment for a good or service is released.

Where a smart contract produces a 'decision' having legal or otherwise significant effects, this form of purely automated decision-making may only be used in the three distinct scenarios foreseen in Article 22(2) GDPR, which provides that the prohibition of automated processing does not apply if the decision:

> (a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;
>
> (b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or
>
> (c) is based on the data subject's explicit consent.

According to **Article 22(2)(a) GDPR**, automated execution is tolerated where it is necessary for the entering into or performance of a contract between the data subject and the controller. The requirement that the contract at issue be concluded between the data controller and the data subject again underlines the importance of being able to clearly identify the data controller in relation to blockchains. **Article 22(2)(b) GDPR** furthermore authorizes Member States or the EU to create exemptions to the prohibition of automated processing provided that data subject rights and interests are safeguarded. At this stage, no legislation has been passed at EU or Member State level to explicitly enable solely automated data processing in relation to smart contracts. **Article 22(2)(c) GDPR** allows automated data processing where it is based on the data subject's explicit consent.[522] It is worth highlighting in this respect that the observations made in relation to consent above are also of relevance in this context. Furthermore, what is required here is not just consent but 'explicit' consent, something that is not defined by the GDPR. The Article 29 Working Party has emphasised that where 'explicit' consent is needed, the data subject 'must give an express statement of consent' which could take the form of a written statement or the filling in of an electronic form or scanned documents using online signatures.[523]

Article 22(2) accordingly provides a number of options to lawfully operate smart contracts. Where this is the case, certain requirements must however be respected. Indeed, if reliance on automated

---

[519] On smart contracts and Article 22 GDPR; see further Finck M (2019) *Smart Contracts as a Form of Solely Automated Processing Under the GDPR* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3311370.

[520] A29WP, Guidelines on Automated Individual Decision-Making and Profiling (n 1) 8.

[521] Ibid 20.

[522] The GDPR does not define 'consent' – the notion however has to be construed in line with Article 29 Working Party, Guidelines on Consent under Regulation 2016/679 WP 259 (28 November 2017).

[523] Ibid.

processing occurs under Article 22(2)(a) or (c), safeguarding measures apply in the form of a **right to human intervention** (under Article 22(3) GDPR) and a **right to be informed** (under Articles 13 and 14 GDPR).[524] The Article 29 Working Party has moreover recalled that where automated processing involves a high risk, a **Data Protection Impact Assessment** ('DPIA') may be desirable.[525]

[524] A29WP, Guidelines on Automated Individual Decision-Making and Profiling (n 1) 20.

[525] Ibid.

8. Data protection by design and by default

Pursuant to **Article 25 GDPR**

> 1. Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.
>
> 2. The controller shall implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility. In particular, such measures shall ensure that by default personal data are not made accessible without the individual's intervention to an indefinite number of natural persons.
>
> 3. An approved certification mechanism pursuant to Article 42 may be used as an element to demonstrate compliance with the requirements set out in paragraphs 1 and 2 of this Article.

Article 25 GDPR imposes an obligation on data controllers to implement **technical and organisational measures** capable of ensuring respect for the principles of European data protection law. This underlines that both system design and organisational structures (which includes blockchain governance) should account for data protection principles, underlining importance of architecture and its influence on individuals.

In accordance with this obligation, the data controller ought to adopt internal policies and implement measures which meet in particular the principles of data protection by design and data protection by default which could include 'minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features'.[526] The GDPR foresees the possibility of using **certification mechanisms** pursuant to Article 42 GDPR 'as an element to demonstrate compliance' with these requirements.[527] Certification is examined separately just below.

Although the Court of Justice has not yet decided any cases on Article 25 GDPR; it held in *Digital Rights Ireland* that the **essence of Article 8 of the Charter of Fundamental Rights** requires the adoption of 'technical and organisational measures' that are able to ensure that personal data is given 'effective protection' against any risk of abuse and against unlawful access and use.[528] This indicates that it is likely that the ECJ will provide a strict interpretation of Article 25 GDPR when called

[526] Recital 78 GDPR.

[527] Article 25 (3) GDPR.

[528] Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland*, paras 40 and 66-67.

upon to adjudicate about this provision, and also underlines the central importance of data protection by design and by default to the overall structure of the Regulation.

Article 25 GDPR relates closely to the data controller accountability clause in Articles 5(2) and 24 GDPR. In the blockchain context, two overarching obligations stemming from these principles should be stressed. First, that whoever uses blockchain technology, no matter of what specific kind, ought to ensure that the **technical specificities** of this tool are such to enable compliance with the GDPR. Second, that data controllers are also obliged to make sure that the processes of **blockchain governance** to ensure that compliance with the GDPR is possible. As amply underlined above, this includes the existence of efficient challenges of communication between data subjects and data controllers but also between various joint-controllers. It will also be seen below that certification and standards may be a means to achieving the obligations enshrined in Article 25 GDPR.

9. Data protection impact assessments

Where data processing is likely to result in a **high risk to fundamental rights**, the controller ought to take preventive action and carry out a Data Protection Impact Assessment ('DPIA') to determine the impact of processing on personal data protection.[529] DPIAs are evaluations of the impact of the planned processing operations on data subjects that ought to be carried out by data controllers where the nature, scope, context and purposes of processing are of high risk to the rights and freedoms of natural parties, which can be the case in particular where new technologies are used.[530] This is required in particular where there is a 'a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person'.[531]

Under **Article 35 GDPR**, such impact assessments are recommended in particular where processing involves (i) a systemic and extensive evaluation of personal aspects of natural persons based on automated processing; (ii) sensitive data and data related to criminal convictions and offences or (iii) where the systematic monitoring of a publicly accessible area on large scale is involved.[532] Where a DPIA indicates that processing results in a high risk for data subjects and no measures to mitigate the risks can be taken, the controller is required to inform the supervisory authority.[533]

Pursuant to **Article 35(7) GDPR**, this assessment ought to provide a systematic description of the purposes and processing activities (as well as, where applicable, any assessment of the legitimate interest of the controller to process personal data), an assessment of the necessity and proportionality of the processing (in relation to the purpose), an assessment of the risks and rights and freedoms of data subjects as well as the envisaged measures to address such risks.

It is important to stress that **the need for a DPIA arises not so much because a specific technology is used but rather because the processing in question is deemed particularly risky**, such as where a large scale of special categories of data or data related to criminal convictions or offenses is processed[534] or a publicly accessible area is systemically monitored on a large scale.[535] The need for a data protection impact assessment thus arises where there is a high risk for data subjects, rather than through the use of a particular technology.

Nonetheless, the use of a new technology may in itself be considered as giving rise to a high risk. Indeed, the United Kingdom's Data Protection Authority considers that a DPIA must be carried out whenever a **new technology** is used.[536] What qualifies as a new technology is, however, notoriously difficult to define as any innovation always builds on previous innovations. Indeed, it has been noted in the introductory section that although blockchain can clearly be considered 'new' it is essentially based on a number of innovations that date back to a few decades ago. Further, one may wonder whether even though one assumes that blockchain is a new technology, for what period it can be

[529] Article 35 (1) GDPR.
[530] Article 35(1) GDPR.
[531] Article 35 (3)(a) GDPR.
[532] Article 35(3) GDPR.
[533] Article 36(1) GDPR.
[534] Article 35 (3)(b) GDPR.
[535] Article 35 (3)(c) GDPR.
[536] https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/

considered 'new'. Indeed, the first blockchain – Bitcoin – is now over ten years old. It would accordingly be helpful if regulatory guidance on blockchains and the GDPR would specify whether the mere use of blockchains creates a high risk to fundamental rights, or whether risk ought to be assessed on a case-by-case basis.

10. Personal data transfers to third countries

Chapter V of the GDPR **limits the circumstances under which personal data can be transferred from the European Union to third countries**. It clarifies that personal data can only be transferred to third countries where these (i) benefit from adequacy decisions, (ii) appropriate safeguards are offered, or (iii) on the basis of a derogation.[537] The examination of these provisions in relation to blockchain technology is important as the multiple nodes on which the ledger is kept can be located in various jurisdictions, both inside and outside the European Union. Whereas the location of the nodes can be controlled in a permissioned network, this is impossible in a permissionless system as anyone may access the network without the need for prior authorisation by a central gatekeeper.

Pursuant to **Article 45 GDPR**, transfers of personal data to third countries are possible on the basis of an **adequacy decision**. Where the European Commission has decided that a third country, territory[538] or specific sector in a third country (or an international organisation) ensure an adequate level of protection, such data transfers do not require any specific authorisation.[539] The European Commission has the ability to issue adequacy decisions taking into account factors such as the respect for the rule of law, human rights and fundamental freedoms as well as relevant legislation and practices[540], whether there is an independent supervisory authority that ensures and enforces compliance with data protection rights[541] and the relevant third country or international organisation's international commitments regarding data protection.[542] If the Commission reaches the conclusion that that jurisdiction provides an adequate level of protection, it can issue an implementing act that recognises this (the adequacy decision) which provides for periodic review (at least every four years).[543]

Adequacy is defined as a **level of protection that is 'essentially equivalent to that ensured within the Union'**.[544] This has been interpreted by the Article 29 Working Party as requiring that these foreign rules comply with a 'core' of GDPR principles, the Charter of Fundamental Rights as well as relevant international instruments (including the Council of Europe's Convention 108).[545] Where an adequacy decision with a third country exists, personal data can thus flow freely between these jurisdictions, notwithstanding whether blockchains or another personal data processing technology are used.

Where personal data is transferred to a jurisdiction that does not benefit from an adequacy decision, a controller or processor may only transfer personal data to a third country where it is able to provide **appropriate safeguards**. Under **Article 46 GDPR**, transfers to third countries are possible where the controller or processor 'has provided appropriate safeguards, and on condition that enforceable data subject rights and effective legal remedies for data subjects are available'.[546] Such safeguards do not require a specific authorisation from a supervisory authority and may include (i) legally binding and enforceable instruments between public authorities or bodies, (ii) binding corporate

[537] Note that there is a hierarchy between these different grounds. Essential equivalence can only be used where there is no adequacy and derogations can only be used where there is no adequacy decision nor essential safeguards.

[538] Territories include the Overseas Countries and Territories that have a special relationship with specific Member States but to which EU law does not apply such as Greenland or French Polynesia and the Netherlands Antilles, among others.

[539] Article 45 (1) GDPR. See also Recital 103 GDPR.

[540] Article 45(2)(a) GDPR.

[541] Article 45(2)(b) GDPR.

[542] Article 45(2)(c) GDPR.

[543] Article 45(3) GDPR.

[544] Recital 104 GDPR.

[545] WP29 2017: Article 29 Working Party, 'Adequacy Referential (updated)' (WP 254, 28 November 2017) 3.

[546] Article 46 (1) GDPR.

rules in accordance with Article 47 GDPR, (iii) standard data protection clauses adopted by a supervisory authority and approved by the Commission, (iv) binding code of conducts together with enforceable commitments of the controller or processor in the third country to apply these safeguards, and (v) approved certification mechanisms together with enforceable commitments of the controller or processor in the third country to apply these safeguards.[547]

Binding corporate rules are 'personal data protection policies which are adhered to by a controller or processor established on the territory of a Member State for transfers or a set of transfers of personal data to a controller or processor in one or more third countries within a group of undertakings, or group of enterprises engaged in a joint economic activity'.[548] These safeguards can take the **form of contractual clauses or provisions** inserted into administrative arrangements between public authorities or bodies that are subject to the prior approval from the competent supervisory authority.[549] These clauses can be included in a broader contractual framework.[550] In accordance with the *Schrems* judgment, data subjects must be able to address claims to DPAs that contest the compatibility of an out-of-EU data transfer with the EU data protection regime, which the relevant DPA must then diligently examine.[551] Where the requirements under Article 46 GDPR are met, blockchain technologies can be used to transfer personal data from an EU Member State to a third country. Indeed, some have thought about the concept of 'binding network rules' to enable such transfers.[552] This is one of the elements that could be discussed in the context of the certification mechanisms introduced further below.

Where personal data is indeed transferred to a third country through one of the various mechanisms outlined above, the data subject must be informed of this. **Article 13(1)(f) GDPR** in fact requires that the data controller also ought to provide the data subject with information about whether it intends to transfer personal data to a third country at the time of data collection.[553] **Article 15(2) GDPR** furthermore requires that where data is transferred to third countries, the data subject shall also be informed of the appropriate safeguards relating to the transfer.[554] In line with what was observed above regarding the right to access to data, blockchains may be an interesting technology to allow data subjects to obtain information about where their data has been transferred to in line with the informational duties applying to third country transfers.

[547] Article 46(2) GDPR.

[548] Article 4 (20) GDPR.

[549] Article 46(3) GDPR.

[550] Recital 109 GDPR.

[551] Case C-362/14 *Maximilian Schrems* [2015] EU:C:2015:650.

[552] See further: https://www.bundesblock.de/wp-content/uploads/.../GDPR_Position_Paper_v1.0.pdf

[553] Article 13(1)(f) GDPR.

[554] Article 15(2) GDPR. See also Article 46 GDPR.

11. Blockchains as a means to achieve GDPR objectives

Up until this stage, the debate has focused primarily on the points of tension between blockchain technologies and the Regulation. These tensions have been explained in further detail above and the subsequent section will formulate concrete policy recommendations that could be adopted in this respect.

As blockchain technologies are better understood and the subject of increased study and experimentation, some have, however, also stressed that the technology might be a suitable tool to achieve at least some of the GDPR's underlying objectives. This section will first provide an overview of blockchains as a data management tool, which may provide benefits for both personal and non-personal data, before introducing related advantages from the perspective of the EU data protection regime.

11.1. Blockchains as a tool of data governance

There is at present increased awareness that the European Union is lagging behind other jurisdictions when it comes to the development of computational intelligence. This is oftentimes traced back to a lack of fluidity in data markets.[555] Indeed, an ongoing policy debate in the EU has underlined that many actors consider that there is insufficient access to the data needed to train computational models. Blockchains have been presented as a potential solution capable of creating data marketplaces for AI development.[556]

In its April 208 data package, the European Commission has stressed that so-called **data marketplaces** – in essence digital marketplaces where personal and non-personal data can be traded as a commodity – may be used to unlock the value of data for the Digital Single Market, also in view of rendering the EU more competitive when it comes to computational intelligence. The Commission considers that data marketplaces 'will give organisations, in particular smaller ones who have datasets to sell, additional routes to market as well as easier billing and subscription mechanisms'.[557]

According to the European Commission, data marketplaces could be powered through APIs, by data marketplaces serving as intermediaries to create bilateral contracts against remuneration or data exchanges designed as closed platforms.[558] The Commission has as a matter of fact indicated that blockchains could be the technology enabling such data-sharing models.[559]

Depending on their respective design, distributed ledgers can indeed offer considerable advantages to gain more granularity over the management and distribution of data. This is due to a number of factors. For instance, blockchains can be designed to enable data-sharing without the need for a central trusted intermediary, they offer transparency as to who has accessed data, and blockchain-based smart contracts can moreover automate the sharing of data, hence also reducing transaction costs.[560] Beyond, blockchains' crypto-economic incentive structures might have the potential to influence the current economics behind data-sharing. At this stage, a number of start-

[555] It is worth noting, however, that future machine learning processes may need much less (non-synthetic) data to be trained than is currently the case.

[556] For an overview, see further Finck M (2019) Blockchain Regulation and Governance in Europe, Cambridge: Cambridge University Press, Chapter Five.

[557] European Commission (2017), 'Commission Staff Working Document on the free flow of data and emerging issues of the European Data Economy' SWD, 2 final 13.

[558] Ibid, 5.

[559] Ibid, 5.

[560] Finck M (2019) Blockchain Regulation and Governance in Europe, Cambridge: Cambridge University Press, 136.

ups are experimenting with this idea to enable new data markets in the European Union.[561] At the same time, other institutions and organisations have initiated projects that use blockchain technologies in order to stimulate data sharing.[562]

It is evident that, if successful, such projects could present broader benefits to the data economy. Blockchains' characteristic as a data management tool may, however, also provide specific benefits to realize some of the GDPR's overall objectives.

11.2. Blockchains as a tool to achieve GDPR objectives

The above-documented characteristics of blockchain technologies as an instrument of data governance can present distinct benefits to realize some of the objectives inherent to European data protection law.

A recent European Parliament report highlighted that 'blockchain technology can provide solutions for the 'data protection by design' provisions in the GDPR implementation on the basis of their common principles of ensuring secured and self-governed data'.[563] Providing data subjects with **control over the personal data** that directly or indirectly relates to them is one of the various objectives pursued by the Regulation. **Recital 7 GDPR** foresees that '[n]atural persons should have control of their own personal data'. This rationale can also be observed on the basis of data subject rights, such as the right of access (Article 15 GDPR) or the right to data portability (Article 20 GDPR) that provide data subjects with control over what others do with their personal data, and what they can do with that personal data by themselves.

Seen from this perspective, control implies on the one hand that data subjects can monitor what happens to personal data relating to them, and, on the other that they can decide who should have access to their personal data. At present, these dual objectives can be hard to pursue in practice. Commentators have observed that at this stage, personal data is only purported to be processed in accordance with law and the data subject has little means to verify whether that is actually the case.[564]

There is, however, precedent of how blockchain technologies could be used to provide data subjects with increased control over their personal data. In **Estonia**, a blockchain-like technical infrastructure has long been used to provide data subjects with more control over their health data.[565] This structure enables data subjects to 'a patient can assess any and all authorisations regarding her data access. By default medical specialists can access data, but any patient can choose to deny access to any case related data, to any, or all care providers; including one's own general practitioner/family physician'.[566]

There is currently broader experimentation being undertaken in relation to blockchains as a control-bestowing tool regarding **health data**. In this context, the role of the blockchain is to (i) secure

[561] Instead of many, see https://oceanprotocol.com/.

[562] See, by way of example, World Economic Forum (23 January 2018), *Harnessing the Fourth Industrial Revolution for Life on Land*, https://www.weforum.org/reports/harnessing-the-fourth-industrial-revolution-for-life-on-land.

[563] European Parliament (27 November 2018) Report on Blockchain: a Forward-Looking Trade Policy (AB-0407/2018) para 14.

[564] Wirth C and Kolain M (2018), 'Privacy by BlockChain Design: A Blockchain-enabled GDPR-compliant Approach for Handling Personal Data' in 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*, https://dl.eusset.eu/bitstream/20.500.12015/3159/1/blockchain2018_03.pdf.

[565] Note that there is a debate as to whether the relevant infrastructure is really a 'blockchain' or just a similar technology. This does not influence on the point made in this section.

[566] Priisalu J and Ottis R (2017) 'Personal control of privacy and data: Estonian experience' 4 *Health and Technology* 441.

uploaded data; (ii) use a decentralised permission management protocol to manage access control to the data, and (iii) record all access activity. Research projects are for instance exploring the potential of data sharing solutions based on blockchains in the health sector.[567] Patientory uses distributed ledgers to encrypt and shred medical records to prevent data breaches.[568] MedRec uses smart contracts as a record management system for EMRs in multi-institutional settings.[569] In the United States, leading healthcare organisations have collaborate don a pilot project to explores how data can be shared through DLT to improve data quality and reduce administrative costs.[570] Another project encourages breast cancer victims to use a distributed ledger to share medical data with researchers.[571] The objective is to train AI algorithms to detect cancer on mammograms while giving patients the option to revoke access to their data.[572].

Similar mechanisms could be designed to allow data-sharing solutions in **other sectors**. It has indeed been emphasised that, more broadly, blockchains could ensure that there is both high availability or, as well as full control over, personal data by offering solutions whereby users keep pointers to the origin of the data.[573] Research has pointed out that 'blockchain is an important technology enabling us to rethink obsolete design models and establish new standards for trust, transparency and privacy under which personal data could be handled in the future'.[574] It thus appears that, if properly designed, the technology may enable alternative forms of data management that present advantages compared to current models.

It is said that DLT can manage access and the further processing of personal data through third parties. The idea here is that the data subject would have a private key that can control access to their personal data to third parties on a case-by-case basis.[575] Faber et. al. have suggested a multi-layer system that could provide users with more control over their data. First, the smart contract layer would store conditions for data exchanges between the user and service providers or purchasers.[576] The access layer would serve to connect an offline storage with the blockchain. This framework would enable users 'to control and own their personal data, while service providers are guests with delegated permissions. Only the user can change this set of permissions and thereby access to the connected data'.[577] Finally, the hash storage layer would store hashes of data, which are created when 'personal data of the user is verified by certain trusted authorities like government organisations who could verify the user's personal data'.[578] In this fashion, the blockchain would store a hash of the verified data, allowing a service provider to verify the user's personal data.[579] Finally, an off-chain repository (any external online database, such as the cloud) will store the actual

[567] Liang X et al (2017), 'Integrating blockchain for data sharing and collaboration in mobile healthcare applications' https://ieeexplore.ieee.org/document/8292361/.

[568] https://patientory.com/our-solution/> accessed on 24 April 2018.

[569] See https://medrec.media.mit.edu/.

[570] United Health Group (2 April 2018) *Humana, MultiPlan, Optum, Quest Diagnostics and UnitedHealthcare Launch Blockchain-Driven Effort to Tackle Care Provider Data Issues* http://www.unitedhealthgroup.com/Newsroom/Articles/Feed/Optum/2018/0402HumanaMultiplanOptumUHCBlockchain.aspx.

[571] Maxmen A (9 March 2018), *AI researchers embrace Bitcoin technology to share medical data* https://www.nature.com/articles/d41586-018-02641-7

[572] Ibid.

[573] Ibid, 4.

[574] Ibid.

[575] Michael Isler (2018), Datenschutz of der Blockchain, JusLetter, 17-18.

[576] Benedict Faber et al, 'BPDIMS: Blockchain-Based Personal Data and Identity Management System (2019) Proceedings of the 52nd Hawaii International Conference on Systems Science, 6859-6860.

[577] Ibid.

[578] Ibid.

[579] Ibid.

user data and be connected to the data pointers of the access layer.[580] This entails that 'data can be fragmented and is less attractive for hacking, while accessing and finding the data in the database is highly efficient'.[581] Off-chain storage allows for deletion of data (though question of the hash) and 'all the user data in these off-chain repositories will be stored in an encrypted form using symmetric encryption keys that are owned by the respective user who owns the data'.[582]

This architecture would store the hashed data pointers pointing to off-chain personal data and provide 'guarantees that the user data has not been altered by the user or anyone else'.[583] The advantage of using blockchains to facilitate such a system is that the provide 'complete transparency and verifiable proofs about various transactions related to the user data and identity management, which will enhance trust and confidence in the system to all the stakeholders such as users, service providers and data purchasers'.[584] Smart contracts on the other hand facilitate fully-automated self-enacting agreements. For the user, the advantage would be to 'be able to grand and revoke access to personal data, but also to monitor who has access to it and what it is being used for'.[585] Another work has also stressed DLT's potential for provenance tracking, which could also extend to personal data.[586]

Such solutions could, for example, be helpful in ensuring compliance with the **right to access** to personal data that data subjects benefit from in accordance with Article 15 GDPR. Furthermore, Isler has argued that DLT can have the potential to support control over personal data in allowing them to monitor respect of the **purpose limitation principle**.[587] In the same spirit, the technology could be used to help with the **detection of data breaches and fraud**. Furthermore, these tools could be used to allow users to track – at least up to a certain degree – what happens with their data, including whether it is transferred to third countries.

More generally, blockchains could be experimented with to determine whether they may be suitable tools that enable data subjects to independently **monitor the data controller's compliance** with its obligations under the GDPR. Research has highlighted the potential usefulness of blockchain to provide a data subject with control over her personal data.[588] At present a data subject may consent to a specific use of personal data but thereafter has little choice than to trust the data controller that it indeed treats the personal data in the agreed manner, and more generally in accordance with the GDPR. Beyond access requests, a data subject however has little means to exercise controls over the actual handling of personal data. In this context, it has been argued that blockchain 'can bring personal data management to a level of privacy and security that prioritizes individual sovereignty and shared transparency'.[589] Indeed, blockchain has been presented as a means to enable the data subject to keep pointers to the origin of the data and smart contracts could be used to provide access to data to a third party whenever this is required.

It is worth highlighting that the European Union is already supporting a number of projects that seek to achieve these objectives through blockchain technology. The **DECODE project** is a

[580] Ibid.

[581] Ibid.

[582] Ibid.

[583] Ibid.

[584] Ibid.

[585] Ibid.

[586] Neisse R et al (2017), 'A Blockchain-based approach for Data Accountability and provenance Tracking' https://arxiv.org/abs/1706.04507.

[587] Isler M (2018), Datenschutz of der Blockchain, *JusLetter*, 1.

[588] Wirth C and Kolain M (2018) Privacy by BlockChain Design: A Blockchain-Enabled GDPR-Compliant Approach for Handling Personal Data, https://dl.eusset.eu/bitstream/20.500.12015/3159/1/blockchain2018_03.pdf

[589] Ibid.

consortium of fourteen organisations from across the European Union that is funded by the European Union's Horizon 2020 programme. It seeks to provide tools that 'put individuals in control of whether they keep their personal data private or share it for the public good'.[590] The DECODE project combines blockchain technology with attribute-based cryptography to provide stronger control over personal and non-personal data. This could lead to a situation where entitlements attached to the data would be searchable in the public domain but will nonetheless only grant access to only those parties entitled to access.[591] The idea is that parties themselves could decide through smart contracts how their data is used, by whom, and on what basis.[592]

Further, **MyHealthMyData** is a project that is also funded under the EU Horizon 2020 scheme that uses blockchain technology to create a structure where data subjects can allow, refuse and withdraw access to their data according to different cases of potential use.[593] Smart contracts are used to implement such choices in view of furthering data sovereignty.[594] In the future, further research could build on these projects to determine whether blockchains can be used to further GDPR objectives also in other regards. Indeed, one of the policy recommendations formulated below is that the EU should continue supporting interdisciplinary research on blockchains potential as a tool to further the objectives inherent to the GDPR.

The above overview has highlighted that there may be room for experimentation with blockchain technologies to function as tools capable of achieving GDPR objectives. It must, however, be stressed that blockchains by no means automatically fulfil these aims. Rather, they must be purposefully designed to do so. The capability for blockchains to both enables new forms of data management and sharing, as well as its ability to function as a tool at the service of blockchain compliance should be further examined through interdisciplinary research.

[590] https://decodeproject.eu/

[591] https://decodeproject.eu/have-more-questions

[592] Ibid.

[593] <http://www.myhealthmydata.eu/> accessed on 24 April 2018.

[594] Panetta R and Cristofaro L, 'A closer look at the EU-funded My Health My Data project' (2017) 4 Digital Health Legal 10.

12. Policy options

This study has examined the relationship between blockchain technologies and European data protection law. It has been seen, firstly, that there is a significant tension between the very nature of blockchain technologies and the overall structure of the GDPR. Whether specific blockchain use cases are compliant with the supranational legal framework can, however, not be examined in a generalised fashion but rather ought to be determined on the basis of a case-by-case analysis. Secondly, the study has also highlighted that in specific cases, this class of distributed technologies may offer distinct advantages that can be helpful to achieve some of the GDPR's objectives. It is on the basis of the preceding analysis that this section develops concrete policy options that could be adopted to ensure that these distributed technologies develop in a manner that is aligned with the legal framework's objectives.

12.1. Regulatory guidance

The key point highlighted in the first and main part of the present study is that there is currently a lack of legal certainty as to how various elements of European data protection law ought to be applied to blockchains. This uncertainty is anchoredin two overarching factors. First, it has been seen that oftentimes, the very technical structure of blockchain technology as well as its governance arrangements stand in contrast with legal requirements. Second, it has also been observed that trying to map the Regulation to blockchain technologies reveals broader uncertainties regarding the interpretation and application of this legal framework. The GDPR is indeed legislation that is based on broad general principles. This bears flexibility and adaptability advantages in an age of fast technological change, yet also has downsides, such as that it can be difficult determine with certainty how a specific provision ought to be applied in a specific context.

Indeed, one year after the GDPR became binding and although the legal regime is largely based on the previous 1995 Data Protection Directive, it is evident that many pivotal concepts remain unclear. Many instances of that phenomenon have been highlighted above. For example, it is currently unclear where the dividing line between anonymous data and personal data due to conflicting statements to this effect in the Regulation and the Article 29 Working Party's interpretation thereof. Moreover, whereas the GDPR recognises a right to 'erasure' that data subjects are free to exercise in some circumstances, there is no indication regarding what 'erasure' actually requires. As such, it is unclear whether erasure in the common-sense understanding of the word is required or whether alternative technical approaches with a similar outcome may be sufficient. These are important questions as erasure in the common-sense understanding of the word is difficult to achieve on DLT whereas alternative technical approaches have been envisaged. Oftentimes, the interpretation of core GDPR concepts is burdened by a lack of harmonious interpretations between the various supervisory authorities in the European Union.

Furthermore, there is currently – in the blockchain context and beyond – an on-going debate regarding the allocation of responsibility for GDPR compliance. The Regulation considers that the data controller is the entity determining the purposes and the means of personal data processing. Yet, in practice only the purposes are taken into account to make that determination. This has led to an expanding number of actors that may be qualified as data controllers – particularly joint-controllers, as is also obvious from recent case law of the CJEU. In addition, there is a lack of legal certainty as to what consequences flow from a finding of controllership, precisely whether the (joint-) controller ought to comply with all GDPR requirements, only those assigned to it in an agreement with other joint-controllers, or only those that are effectively within its responsibilities, powers and capabilities. It is hoped that future case law, especially the upcoming judgment in *FashionID*, will clarify at least some of these questions, which are important for blockchains but also beyond.

This study has furthermore observed that blockchain technologies challenge core assumptions of European data protection law, such as that of data minimisation and purpose limitation. At the same time, however, this is a broader phenomenon, as these principles are also hard to map to other elements of the contemporary data economy such as big data analytics facilitated by 'Artificial Intelligence' techniques such as machine learning or deep learning. Indeed, the interpretation to be given to the overarching requirements of data minimisation and purpose limitation is not obvious in such contexts.

Whereas some have called for a revision of the GDPR, it is not evident that this is necessary. The Regulation was designed as a form of principles-based regulation that is technologically neutral and stands the test of time in a fast-changing data-driven economy. Thus, it is not the structure of the GDPR as such that causes confusion, rather the lack of certainty as to how specific concepts should be interpreted. This could be addressed through regulatory guidance without the need for legislative reform, which would itself come with significant limitations and disadvantages.

Regulatory guidance could as a matter of fact provide add much legal certainty compared to the current status quo. This could take the form of various regulatory initiatives. On the one hand, supervisory authorities could coordinate through the European Data Protection Board to draft specific guidance on the application of the GDPR to blockchain technologies at supranational level, preventing the risk of fragmentation that would result from a number of independent initiatives in the various Member States. Whereas such specific guidance would be important to generate more legal certainty, a revision of other, more general guidance documents of the Article 29 Working Party would also be helpful. Indeed, it has been observed above that these have sometimes themselves generated uncertainty as to how specific provisions of the GDPR should be applied. It has, for instance, been seen that whereas the GDPR itself adopts a risk-based approach to anonymisation, the Article 29 Working Party has endorsed a somewhat divergent test. Updating some of these more general guidance documents, in particular those that have not been endorsed by the EDPD would be helpful to address outstanding questions in the context of blockchain technologies but also beyond.

The provision of regulatory guidance would indeed achieve a dual objective. On the one hand, it would provide further certainty to actors in the blockchain space, which have long stressed that the difficulty of designing compliant blockchain use cases relates in part to the lack of legal certainty of what exactly is required to design a compliant product. On the other hand, regulatory guidance on how the GDPR is applied to blockchains, as well as on specific elements of the GDPR that more generally have been the source of confusion could add more certainty and transparency in the data economy more broadly.

On the basis of the analysis carried out in this study, the questions to be addressed in this context should include the following:

- ◊ Can the household exemption be invoked in relation to public and permissionless blockchains where data is shared with an indefinite number of people?
- ◊ Is anonymisation an effective means of provoking the 'erasure' of data for the purposes of Article 17 GDPR?
- ◊ Should the anonymisation of data be evaluated from the controller's perspective, or also from the perspective of other parties?
- ◊ Can a peppered hash produce anonymous data?
- ◊ What is the status of the on-chain hash where transactional data is stored off-chain and subsequently erased?
- ◊ What is the status of anonymity solutions such as zero knowledge proofs under the GDPR?

- ◊ Is there a *de minimis* test regarding influence over the purposes and means of processing that must be crossed for an actor to qualify as a (joint-) controller?
- ◊ What is the scope of a data controller's responsibility under the GDPR? Is responsibility limited to the (joint-) controller's responsibilities, powers and capacities?
- ◊ How is the purpose of personal data processing to be evaluated in relation to blockchains in light of the purpose limitation principle? Does this only encompass the initial purpose (the transaction) or does it also encompass the continued storage of the data and its further processing, such as to achieve consensus?
- ◊ Can a data subject be a data controller in relation to personal data that relates to themselves?
- ◊ What is the relationship between the first and third paragraph of Article 26 GDPR? Is there a need for a nexus between responsibility and control?
- ◊ How ought the principle of data minimisation to be interpreted in relation to blockchains?
- ◊ Is the off-chain storage of transactional data a means of complying with the data minimisation principle?
- ◊ Is the provision of a supplementary statement always sufficient to comply with Article 16 GDPR?
- ◊ How ought 'erasure' to be interpreted for the purposes of Article 17 GDPR? Can the deletion of a private key satisfy lead to the erasure of on-chain data?
- ◊ How ought Article 18 GDPR regarding the restriction of processing to be interpreted in the context of blockchain technologies?
- ◊ Does the continued processing of data on blockchains satisfy the compelling legitimate grounds criterion under Article 21 GDPR?
- ◊ Does the mere use of a blockchain trigger a need to carry out a data protection impact assessment?

12.2. Support codes of conduct and certification mechanisms

As a technologically-neutral legal framework, the GDPR was designed in a manner that should enable its application to any technology. This design presents many advantages, such as that it is supposed to stand the test of time and that it does not discriminate between particular technologies or use cases thereof. Indeed, as an example of principles-based regulation, European data protection law devises a number of general overarching principles that must then be applied to the specificities of concrete personal data processing operations.

The technology-neutrality of the GDPR however also entails that it can at times be difficult to apply its obligations to specific cases of personal data processing, as evidenced by the analysis above. It is important to note that the Regulation itself provides mechanisms specifically designed to deal with this: certification mechanisms and codes of conducts. These tools were included in the Regulation specifically to enable the application of the GDPR's overarching principles to concrete contexts where personal data is processed. In contrast to the adoption of regulatory guidance as suggested above, certification mechanisms and codes of conducts exemplify a co-regulatory spirit whereby regulators and the private sector collaborate to devise principles designed to ensure that the principles of European data protection law are respected where personal data is processed. This has, for instance, been done in relation to cloud computing where many of the difficult questions examined above have also arisen when these solutions were first deployed. The EU Cloud Code of Conduct was defined between the major cloud-computing providers as a means of securing GDPR compliance in collaboration with the European Commission and the Article 29 Working Party.[595]

[595] See further: https://eucoc.cloud/en/about/about-eu-cloud-coc.html.

Like blockchain now, cloud computing has raised many difficult questions regarding GDPR compliance and the code of conduct was seen as one means to introduce more legal certainty in this area and ensure a higher adherence to the objectives of the Regulation. As such, the establishment of codes of conduct and certification mechanisms could be very useful also in the context of blockchain technologies. This could, for instance, include the design of binding network rules regarding international data transfers.

Article 40 GDPR foresees the establishment of codes of conduct by associations and other bodies that represent categories of data controllers or processors. Article 42 GDPR moreover encourages that data protection certification mechanisms be established in the form of data protection seals and marks to demonstrate compliance with the GDPR. The notion of the certification mechanism is not defined although the reference to 'data protection seals and marks' would indicate that this could take the form of a trustmark visible through the user interface or similar mechanisms. Companies that are using DLT in their operations should accordingly be encouraged to develop codes of conduct and certification mechanisms specifically tailed to DLT. Whereas these initiatives do not remove the need for a case-by-case compliance assessment, they are valuable starting points for such analysis. Moreover, codes of conduct and certification mechanisms are valuable steps towards ensuring that technical systems are designed to be compliant-by-design in line with the data protection by design and data protection by default obligations enshrined in the Regulation.

Actors relying on approved codes of conduct under Article 40 GDPR or certification mechanisms under Article 42 GDPR moreover benefit from a risk-management perspective. As a matter of fact, adherence to these standards can be used by the data controller to demonstrate compliance with its obligations under Article 24 GDPR. The European Union could accordingly encourage the initiation of related procedures, which are complementary to the provision of regulatory guidance in order to resolve some of the uncertainties in this area.

12.3. Research funding

Regulatory guidance as well as codes of conduct and certification mechanisms could add much legal certainty where the tension between the GDPR and blockchain technologies stems from a lack of legal certainty as to how specific provisions of the GDPR ought to be applied.

This, however, will not always be sufficient to enable the compliance of a specific distributed ledger use case and European data protection law. Indeed, it has been amply underlined in the above analysis that in some cases, there are technical limitations to compliance. In such instances, regulatory guidance, certification mechanisms and codes of conduct arguably will not go far enough to resolve a lack of compliance. In other cases, the current governance design of blockchain use cases stands in the way of compliance. These technical and governance limitations could be addressed by interdisciplinary research on these matters.

Such interdisciplinary research could, for example, define governance mechanisms that enable various controllers in decentralised networks to coordinate effectively in order to enforce data subject rights, something that, as has been seen above, is not straightforward in the current state of affairs – in DLT and beyond. Other interesting topics would include the design of mechanisms that enable the effective revocation of consent in contexts of automated personal data processing as well as definitions of technical solutions to comply with Article 17 GDPR. More broadly, such research could also focus on data protection by design solutions under Article 25 GDPR; for instance the development of protocols that would be compliant by design.

This would benefit the development of compliant blockchain solutions in the European Union but could also more broadly serve to design solutions of, for instance, anonymity and data-sharing that would be of much broader relevance to the Digital Single Market as they could also be deployed in other contexts. As data-ecosystems are increasingly decentralised even beyond the DLT realm, such research could benefit the digital domain more generally. This would benefit the Digital Single Market, support the EU's global leadership role in data protection and the digital economy and lay the groundwork for suitable and sustainable future regulation.

13. Conclusion

This study has discussed the application of the European Union's EU General Data Protection Regulation to blockchain technologies. It has been observed that many points of tension between blockchains and the GDPR can be identified. Broadly, it can be maintained that these are due to two overarching factors. First, the GDPR is based on the underlying assumption that in relation to each personal data point there is at least one natural or legal person – the data controller – that data subjects can address to enforce their rights under EU data protection law. Blockchains, however, often seek to achieve decentralisation in replacing a unitary actor with many different players. This makes the allocation of responsibility and accountability burdensome, particularly in light of the uncertain contours of the notion of (joint)-controllership under the Regulation. A further complicating factor in this respect is that in light of recent developments in the case law, defining which entities qualify as (joint-) controllers can be fraught with uncertainty. Second, the GDPR is based on the assumption that data can be modified or erased where necessary to comply with legal requirements such as Articles 16 and 17 GDPR. Blockchains, however, render such modifications of data purposefully onerous in order to ensure data integrity and increase trust in the network. Determining whether distributed ledger technology may nonetheless be able to comply with Article 17 GDPR is burdened by the uncertain definition of 'erasure' in Article 17 GDPR.

The study has concluded that it can be easier for private and permissioned blockchains to comply with these legal requirements as opposed to private and permissionless blockchains. It has, however, also been stressed that the compatibility of these instruments with the Regulation can only ever be assessed on a case-by-case basis. Indeed, blockchains are in reality a class of technologies with disparate technical features and governance arrangements. This implies that it is not possible to assess the compatibility between 'the blockchain' and EU data protection law. Rather, this study has attempted to map various areas of the GDPR to the features generally shared by this class of technologies, and to draw attention to how nuances in blockchains' configuration may affect their ability to comply with related legal requirements. Indeed, the key takeaway from this study should be that it is impossible to state that blockchains are, as a whole, either completely compliant or incompliant with the GDPR. Rather, while numerous important points of tension have been highlighted and ultimately each concrete use case needs to be examined on the basis of a detailed case-by-case analysis.

The second key element highlighted in this study is that whereas there certainly is a certain tension between many key features of blockchain technologies setup and some elements of European data protection law, many of the related uncertainties should not only be traced back to the specific features of DLT. Rather, examining this technology through the lens of the GDPR also highlights significant conceptual uncertainties in relation to the Regulation that are of a relevance that significantly exceeds the specific blockchain context. Indeed, the analysis has highlighted that the lack of legal certainty pertaining to numerous concepts of the GDPR makes it hard to determine how the latter should apply to this technology, but also others. This is, for instance, the case regarding the concept of anonymous data, the definition of the data controller, and the meaning of 'erasure' under Article 17 GDPR. A further clarification of these concepts would be important to create more legal certainty for those wishing to use DLT, but also beyond and thus also to strengthen the European data economy through increased legal certainty.

The study has, however, also highlighted that blockchains can offer benefits from a data protection perspective. Importantly, this is by no means automatically the case. Rather, blockchains need to be purposefully designed in order for this to realize. Where this is the case, they may offer new forms of data management that provides benefits to the data-driven economy and enable data subjects to have more control over personal data that relates to them.

It is on the basis of these observations that the study has formulated three broad policy recommendations, which have been broken down into various elements. First, it was suggested that regulatory guidance on the interpretation of certain elements of the GDPR when applied to blockchains should be provided to generate more legal certainty in this area. Second, it was recommended that codes of conduct and certification mechanisms should be encouraged and supported. Third, it was recommended that funding be made available for interdisciplinary research exploring how blockchains' technical design and governance solutions could be adapted to the GDPR's requirements, and whether protocols that are compliant by design may be possible.

REFERENCES

Acar G (9 April 2018), *Four cents to deanonymize: Companies reverse hashed email addresses* https://freedom-to-tinker.com/2018/04/09/four-cents-to-deanonymize-companies-reverse-hashed-email-addresses/.

Antonopoulos A (2017), Mastering Bitcoin, O'Reilly.

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

Article 29 Working Party, Opinion 05/2012 on Cloud Computing (WP 196) 01037/12/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 5/2009 on Online Social Networking (WP 163) 01189/09/EN.

Article 29 Working Party, 'Guidelines on Transparency under Regulation 2016/269' WP29 2018B.

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

Ateniese G, Magri B, Venturi D and Andrade E (2017), 'Redactable Blockchain – or – Rewriting History in Bitcoin and Friends' https://eprint.iacr.org/2016/757.pdf.

Austrian Data Protection Authority, DSB-D123.270/0009-DSB/2018 (05 December 2018).

Bacon J et al (2018), 'Blockchain Demystified: A Technical and Legal Introduction to Distributed and Centralised Ledgers' 25 *Richmond Journal of Law and Technology* 1

Beck R, Müller-Bloch C and King J (2018) *Governance in the Blockchain Economy: A Framework and Research Agenda.*

Berberich M and Steiner M (2016) 'Blockchain technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?' 2 *European Data Protection Law Review* 422.

Böhme R and Pesch P (2017), 'Technische Grundlagen und datenschutzrechtliche Fragen der Blockchain Technologie' 41 *Datenschutz und Datensicherheit* 473.

Blocher W, Hoppen A and Hoppen P (2017) 'Softwarelizenzen auf der Blockchain' 33 *Computer und Recht* 337.

Brakerski Z and Gentry C and Vaikuntanathan V (11 August 2011), Fully Homomorphic Encryption without Bootstrapping, https://eprint.iacr.org/2011/277.

Buocz T et al (2019), 'Bitcoin and the GDPR: Allocating Responsibility in Distributed Networks' *Computer Law & Security Review* 1.

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

Conte de Leon D et al (2017), 'Blockchain: Properties and Misconceptions' 11 *Asia Pacific Journal of Innovation and Entrepreneurship* 286.

Edwards L (2018), *Law, Policy and the Internet*, Oxford: Hart Publishing.

Erbguth J and Fasching J (2017), 'Wer ist Verantwortlicher einer Bitcoin-Transaktion?' 12 *Zeitschrift für Datenschutz* 560.

European Parliament (27 November 2018), Report on Blockchain: a Forward-Looking Trade Policy (AB-0407/2018).

Benedict Faber et al, 'BPDIMS: Blockchain-Based Personal Data and Identity Management System (2019) Proceedings of the 52nd Hawaii International Conference on Systems Science, 6859.

Felten E (22 April 2012), *Does Hashing Make Data 'Anonymous'* https://www.ftc.gov/news-events/blogs/techftc/2012/04/does-hashing-make-data-anonymous

Felten E (26 February 2018) *Blockchain: What is it good for?* <https://freedom-to-tinker.com/2018/02/26/bloc>.

Finck M (2019) Blockchain Regulation and Governance in Europe Cambridge: Cambridge University Press.

Huckle S et al (2016), 'Internet of Things, Blockchain and Shared Economy Applications' 98 *Procedia Computer Science* 461.

Michael Isler (2018), Datenschutz of der Blockchain, JusLetter.

Information Commissioner's Office (26 February 2014), 'Deleting Personal Data' https://ico.org.uk/media/for-organisations/documents/1475/deleting_personal_data.pdf

Kuan Hon W et al (2011), 'Who is Responsible for 'Personal Data' in Cloud Computing? The Cloud of Unknowing, Part 2' *Queen Mary School of Law Legal Studies Research Paper* No. 77.

Mahieu R et al (2018) *Responsibility for Data Protection in a Networked World. On the question of the controller, 'effective and complete protection' and its application of data access rights in Europe* 12 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3256743

Martini M and Weinzierl Q (2017), 'Die Blockchain-Technologie und das Recht auf Vergessenwerden' 17 *Neue Zeitschrift für Verwaltungsrecht* 1251.

Matzutt R et al (26 February 2018) *A Quantitative Analysis of the Impact of Arbitrary Blockchain Content on Bitcoin* https://fc18.ifca.ai/preproceedings/6.pdf 1.

Meyer D (27 February 2018), *Blockchain technology is on a collision course with EU privacy law* <https://iapp.org/news/a/blockchain-technology-is-on-a-collision-course-with-eu-privacy-law/>.

Millard C (2013) Cloud Computing Law Oxford: Oxford University Press.

Moerel L (2019) 'Blockchain & Data Protection…and Why They are not on a Collision Course' 6 *European Review of Private Law* 825.

Mourby M et al (2018), 'Are 'pseudonymised' data always personal data? Implications of the GDPR for administrative data research in the UK' 34 *Computer Law & Security Review* 222, 223.

Nakamoto S (2009), *Bitcoin: A Peer-to-Peer Electronic Cash System* https://bitcoin.org/bitcoin.pdf

Narayanan A et al (2016), Bitcoin and Cryptocurrency Technologies Princeton University Press.

Narayanan, A and Clark J (2017) 'Bitcoin's academic pedigree' 60 *Communications of the ACM*.

Narayanan A and Shmatikov V (2010) 'Myths and Fallacies of Personally Identifiable Information' 53 *Communications of the ACM* 24.

Neisse R et al (2017), 'A Blockchain-based approach for Data Accountability and provenance Tracking' https://arxiv.org/abs/1706.04507

Nisbet K (30 April 2018), *The False Allure of Hashing for Anonymization* https://gravitational.com/blog/hashing-for-anonymization/

Ohm P (2010) 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' 57 *UCLA Law Review* 1701.

Panetta R and Cristofaro L, 'A closer look at the EU-funded My Health My Data project' (2017) 4 Digital Health Legal 10.

Purtova N (2018) 'The law of everything. Broad concept of personal data and future of EU data protection law' 10 *Law, Innovation and Technology* 40.

Sikorski J, Haughton J and Kraft M (2017), 'Blockchain technology in the chemical industry: Machine-to-machine electricity market' 195 *Applied Energy* 234.

Sillaber C and Waltl B (2017), 'Life Cycle of Smart Contracts in Blockchain Ecosystems' 41 *Datenschutz und Datensicherheit* 497.

Singh J and Michels J (2017), Blockchain As a Service: Providers and Trust *Queen Mary School of Law Legal Studies Research Paper* No. 269/17, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3091223.

Stalla-Bourdillon, S and Knight, A (2017) 'Anonymous data v. personal data—a false debate: an EU perspective on anonymization, pseudonymization and personal data' *Wisconsin International Law Journal*.

Sweeney L (2000) 'Simple Demographics Often Identify People Uniquely' *Data Privacy Working Paper* 3 https://dataprivacylab.org/projects/identifiability/paper1.pdf.

Walch A (2017), 'The Path of the Blockchain Lexicon (and the Law)' 36 *Review of Banking and Financial Law* 713.

Werbach K and Cornell N (2017), 'Contracts Ex Machina' 67 *Duke Law Journal* 313, 335.

Wirth C and Kolain M (2018), 'Privacy by BlockChain Design: A Blockchain-enabled GDPR-compliant Approach for Handling Personal Data' in 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* , 5 https://dl.eusset.eu/bitstream/20.500.12015/3159/1/blockchain2018_03.pdf

Wu H and Feng Wang F (2014), A Survey of Noninteractive Zero Knowledge Proof System and Its Applications, *The Scientific World Journal*, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4032740/

Van der Sloot B (2015), 'Do Privacy and Data Protection Rules Apply to Legal Persons and Should They? A Proposal for a Two-Tiered System' 31 *Computer Law and Security Review*.

Veale M (2018) et al, 'When data protection by design and data subject rights clash' 8 *International Data Privacy Law* 105.

Zuiderveen Borgesius F (2016), 'Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation' 32 *Computer Law & Security Review* 256.

Blockchain is a much-discussed instrument that, according to some, promises to inaugurate a new era of data storage and code execution, in turn potentially stimulating new business models and markets. The precise impact of the technology is, of course, hard to anticipate with certainty, not least since many remain sceptical of blockchain's possible impact. In recent times, there has been much discussion in policy circles, academia and the private sector regarding the tension between blockchain and the European Union's General Data Protection Regulation (GDPR). This study examines the European data protection framework and applies it to blockchain technologies in order to document these tensions.

QA-02-19-516-EN-N

This is a publication of the Scientific Foresight Unit (STOA)
EPRS | European Parliamentary Research Service

ISBN 978-92-846-5044-6
doi: 10.2861/535
QA-02-19-516-EN-N