R3 Reports

The Application & Impact of the European General Data Protection Regulation on Blockchains

Jana Moser

Contents

R3 Research aims to deliver concise reports on DLT in business language for decision-makers and DLT hobbyists alike. The reports are written by experts in the space and are rooted in practical experience with the technology.

The Application and Impact of the European General Data Protection Regulation on Blockchains

Jana Moser

February 15, 2017

Abstract

This paper assesses if and how the European Data Protection Regulation (GDPR) applies to public and private or consortium blockchains. The paper focuses on the crucial question of whether blockchains fall within the scope of GDPR, especially if personally identifiable information is processed. The paper proposes that this is most likely true and that for public blockchains the data is not simply anonymous. Finally, the paper describes the main obligations and requirements under the GDPR by which blockchain companies must abide.

1 Regulatory background

The European General Data Protection Regulation (GDPR) was adopted by the European Parliament on 4th April 2016 and will come into effect on 25th May 2018. Some companies are established outside the European Economic Area (EEA) and therefore do not process data within the EEA. Once this piece of regulation takes effect, we should get more clarity on if, and to what extent, current European data protection and privacy laws apply to these companies.

To understand the impact of the GDPR on blockchains – either processing data on permissioned private networks, such as Corda, or public blockchain(s), such as Ethereum – two main questions must be considered:[1]

1. Does the GDPR apply to a company running a public, private or consortium blockchain?
2. Is there data in blockchains that is considered Personally Identifiable Information (PII)?

Only if both questions are true will further regulatory requirements arise (see **Section C** below).

2 Regulatory scope

2.1 Material scope of the GDPR

The European data protection regulation is meant to be technically neutral. The focus lies on harmonizing privacy regulation, strengthening the protection of natural persons and enabling free flow of data within the European Economic Area (EEA). Thus, any processing of personal data (Art. 2 par. 1 GDPR) falls within the scope of the GDPR except where a statutory exemption applies.

A "household exemption" is defined in Article 2 paragraph 2 (c) of the GDPR:

[1]For the purposes of this paper, a private blockchain means that an identified, known set of individuals or entities validate and maintain the network whereas a public blockchain is one in which non-identified (pseudonymous) participants validate and maintain the network. A blockchain such as Ethereum involves stringing together a chain of containers called blocks, which bundle transactions together like batch processing, whereas a distributed ledger, like Corda, does not and instead validates each transaction (or agreement) individually. The fact that blockchains are set up as distributed ledgers does not affect the application of European data protection law. For simplicity, the term blockchain and distributed ledger technology (DLT) are considered to be interchangeable in this paper.

"This Regulation does not apply to the processing of personal data: ... (c) by a natural person in the course of a purely personal or household activity"

This exemption applies to data processing from regulatory scope if it is done for purely personal or household activities. With public blockchains, this "household exemption" is arguably not relevant as it does not apply even to private blockchains because blockchains are connected to some kind of professional or commercial activity.[2]

Even if public blockchains were only used on personal computers, the household exemption would still not apply: the European Court of Justice explained in the Lindqvist verdict in 2003 that data accessible to an indefinite number of people contradicts the household exemption.

2.2 Territorial scope of the GDPR

One main purposes of the GDPR is to establish a level playing field by clarifying the territorial scope of GDPR. From 25 May 2018, any company in the world must assess whether the GDPR applies. That includes companies that are responsible for data processing, i.e. for storing, analysing or other means of processing (called a data controller), as well as (sub)contractors who process data only on behalf of another company (data processors). Companies not established in the EU must expect to abide by the GDPR too, if they either offer goods (free or paid) to EU citizens, or monitor their behaviour in the EU.

Figure 1: Territorial scope of the GDPR

Applicable to companies	Establishment	Data processing related to the EU	
Data processing in the context of the activities	European Union	Regardless of whether the processing takes place in the Union or not	Art. 3 par. 2 (a) GDPR
of a controller			
of a processor			
Data processing in the context of the activities	Not in the European Union	To the processing of personal data of data subjects who are in the Union where the processing activities are related to - Offering goods (free or paid) - Monitoring of their behaviour in the Union	Art. 3 par. 2 (b) GDPR
of a controller			
of a processor			

Consequently, it is not of primary importance where data processing takes place, to affirm the territorial scope of the GDPR. If a company established in the EU uses a public, private or consortium blockchain or distributed ledger it must comply with the GDPR. It cannot avoid it by using US-based service providers. If, on the other hand, non-European companies run the blockchain, it depends on whether data of European citizens is processed (if EU citizens are either the target or subject of any monitoring). European domains such as German (*.de), French (*.fr), or support telephone numbers (+49*, +33*) are strong pieces of evidence to affirm that offerings are made to EU citizens or their behaviour is monitored.

2.3 Personally identifiable information

The GDPR applies only to processing of "personal data" (Art. 1 par. 1 GDPR) which is legally defined as,

"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 specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person;" (Art. 4 par. 1 GDPR)

Basically, the future regulation aligns with the currently applicable EU privacy directive which says,

[2]Any business activity is not covered by the household exemption. This also means supply chain management etc. is not excluded from GDPR.

"'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;" (Art. 2 (a) EU Directive 65/46/EC)

This means that any processing related to an identified or identifiable natural person is relevant from a privacy perspective. The Article 29 Data Protection Working Party 136[3] adopted an opinion on the concept of personal data and said that, "in order to consider that the data "relate" to an individual, a "content" element OR a "purpose" element OR a "result" element should be present".

Examples of data considered as "relating to an individual":

- Content element: a person has signed a contract.
- Purpose element: 100 ad views are attributed to a website visitor
- Result element: an analysis of browsing history shows that a visitor is most probably interested in sports.

Consequently, the term "identifiable" is subject to interpretation, and the understanding of most privacy professionals and all European data protection authorities is rather conservative: an individual is identifiable if either the data processor or any other person/company has the knowledge to identify the data subject using this information and the data processor has the technical and/or legal means to procure this additional information to identify the individual.

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 (see Recital 26 GDPR). That means it doesn't matter whether or not the company has the intention to identify the data subject. Only the technical and legal potential possibility are relevant.

2.3.1 Anonymous data

The principles of data protection should therefore not apply to company or anonymous data. The latter is 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 no longer identifiable.

a) Definition of PII

Personally identifiable information (PII) includes names, postal addresses, email addresses, telephone numbers, complete IP addresses, user IDs and any other directly or indirectly related information. This data is not anonymous, and is therefore subject to privacy regulation.

Although cookie IDs, device IDs, MAC addresses, UDIDs etc. are pseudonymous, they are still deemed to be PII. Using pseudonyms, identity is disguised, but not erased, from the data set – it is possible to collect additional data relating to the same individual without having to know his identity. A typical example of pseudonymous data is web tracking based on cookies or similar means. Furthermore, simply replacing data with other data (e.g. replacing a name with an ID) does not anonymize it.

b) Creating anonyms

There are various technologies that can anonymize data to a certain extent. Aggregation, encryption, and tokenization are the most common techniques. However, due to the increasing computing capabilities no technology or technique can be determined as the "perfect" anonymization solution. The entire data processing and data separation concept must be considered as a whole.

Even data encryption does not automatically change PII to anonymous data – encryption does not amend the value and content of data, but rather creates pseudonymized data where the relation can be re-created and the individual re-identified with proportionate means (Art. 29 Data Protection Working Party, WP 136).

[3]The Article 29 Data Protection Working Party is composed of: (1) a representative of the supervisory authority (ies) designated by each EU country, (2) a representative of the authority (ies) established for the EU institutions and bodies, and (3) a representative of the European Commission. It has no investigative powers but examines privacy questions and publishes opinions about data protection.

Although disguising identities can be done in a way that no re-identification is possible, e.g. one-way cryptography, which creates anonymized data, simple hashing - even with salted hash algorithms – is mostly considered an insufficient method of anonymization.

Full anonymization requires the removal of the relation to an individual person, with the result that no natural person can be connected to, or singled out from, the information given.

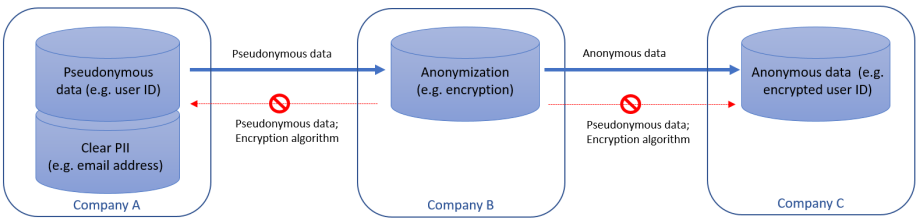
To create anonymous data, additional measures need to be considered, like generalizing of attributes, deletion of data sources or strong data aggregation. For instance, concerning IP addresses, the deletion of the last two octets is suggested.

The assessment as to whether data is anonymous should be done from an attacker's point of view who intends to misuse and/or re-combine data to identify a natural person (Art. 29 Data Protection Working Party, WP 216). Consequently, anonymization is meant to prevent potential de-anonymization by making it disproportionate or even impossible. Basically, data is only anonymous if a person is neither identifiable by other existing data, correlation (e.g. search results in a search engine (Art. 29 Data Protection Working Party, WP 148) nor by additional data. That means that data profiles can be anonymous only if any identification is irreversible from data processor's "relative" point of view. It is crucial what he knows and if he has reasonable technical and legal means to get the knowledge of others – the "absolute" perspective so to say – needed to identify a person.

c) Anonymization through knowledge distribution

Consequently, a good way to make de-anonymization disproportionate for a company is distribution of knowledge and control, as many parties must work together to re-identify a natural person. Data is anonymous if a data processor transfers pseudonymous data to a third and independent party which has no means to re-identify the natural person. The third party could, for instance, encrypt pseudonymous data. The anonymizer gets data from the data processor and both parties agree that the anonymizer is obliged to hold this data and the encryption algorithm strictly confidential. The data protection authorities seem to have accepted this set up as privacy compliant for web tracking and profiling solutions like targ.ad and nugg.ad. Both use anonymizer solutions provided by an independent third party who erases the relation to a natural person by deletion of IP addresses.

Figure 2: Anonymization of data



For Google Analytics "anonymizeIP" it seems even to be accepted to encrypt data on Google's frontend servers in Europe. That refers to anonymization in the environment of the responsible data processor and not a third party, although the requirements for a technical and organizational separation of anonymous and personal identifiable data must then be far higher.

All in all, simple cryptography alone is not enough to affirm anonymization but data and "knowledge federalism" is an additional means to increase the likelihood that re-identification is disproportionately difficult.

2.3.2 PII in blockchains

Information related to or processed in a blockchain is often considered to be anonymous. If this were true, the GDPR would generally not apply to blockchains. As described above, the assessment of whether or not the data processing is anonymous depends very much on the entire technical set up and process, including the sources of information being added to the blockchain. Moreover, anonymization should not be regarded as a one-off exercise and the attending risks

should be reassessed regularly by data controllers (Art. 29 Data Protection Working Party, WP 216).

a) Information on blockchains

Any function making up the blockchain, is simply a function and therewith non-PII if it does not contain any address (e.g. IP address, Bitcoin address) or other information (e.g. names, banking information). So, there is no difference to the privacy assessment of any other software code. But as soon as information is used to process a transaction, or even stored on a blockchain, it is related to a transaction: it contains a history, a state and a state transition rule. No matter what the subject of a transaction, it always needs an address to which the current state is related. For Bitcoin, for instance, a transaction relates to the amount of bitcoins held by a specific address (e.g. address 39BaMQCphFXyYAvcoGpeKtnptLJ9v6cdFY has 522.11790015 bitcoins, address 375zAYokrLtBVv6bY47bf2YdJH1EYsgyNR has 375 bitcoins ...).

b) Bitcoin's blockchain is pseudonymous not anonymous

The original purpose of Bitcoin-like blockchains is that all transactions can be checked all the way back to the genesis block by everyone. It is a permanent, public record of the transactions that occur. Although the bitcoin address itself is not directly related to a natural person, at some stage the person who transfers bitcoin bought bitcoin (e.g. exchanged fiat money for bitcoin). If people don't circumvent banking regulation this very first transaction is mandatorily related to a natural person and the following bitcoin transactions too. Consequently, it is generally possible for someone to relate a bitcoin address to a natural person, though only through legal means (e.g. right to information).[4]

So, the Bitcoin protocol contains not anonymous, but pseudonymous (and therewith personally identifiable) information to which the GDPR applies.[5]

2.3.3 PII in Ethereum

Whereas Bitcoin is a ledger containing a list of ownership of bitcoins, Ethereum can be viewed as a transaction-based state machine and the Ether token (ETH) was originally not intended to be used as or considered a currency, asset, share or anything else (Wood, 2014).

The state of Ethereum at any point in time can be described as the state of all the accounts on Ethereum, where each account is either an externally owned account (EOA) or a contract (Buterin, 2016). The state of each account can include such information as account balances, reputations, trust arrangements, or data pertaining to information of the physical world; in short, anything that can currently be represented by a computer. A transaction in Ethereum specifies a destination address, a quantity of ether to transact and a "data" field which theoretically can contain any information (and also a sender address, although this is implicit in the signature and therefore is not specified explicitly).

In contrast to Bitcoin, not all data in Ethereum must be stored on the public blockchain, especially the mapping table between addresses (Ethereum accounts) and account states, which is stored in a simple immutable database backend. In private or consortium blockchains that run a cloned copy of Ethereum code, the Ethereum state transition rules can be separated from the Ethereum public blockchain consensus mechanism. However, the information on the blockchain is inherently linked to a database backend and vice versa. Here, Bitcoin and Ethereum work similarly.

In the future, Ethereum could potentially store contracts encrypted in the blockchain. The obfuscated accounts could only be read by the contract if the contract decrypts the storage and therewith the account internally, i.e. on a person's machine, and the smart contract code checks if the person is entitled to decrypt the code. That means that an authorization concept is implemented in the code. Consequently, from a privacy perspective it is basically not important what the contract code or the underlying rule is. It is only relevant if it contains PII or if someone might have access to PII. And this is obviously true for the entitled person who decrypts the information. That means that PII is processed on the decrypting server/computer.

Additionally, contracts have their own addresses, and so can serve as owners of digital assets in the same way that users can (Buterin, 2016). There is a strong parallel to permanent IP addresses:

[4] There are several companies (e.g. Chainalysis, Blockseer) that provide forensic and tracking tools to compliance teams and law enforcement.

[5] It is basically possible to use tor networks and other tools to remain as anonymous as possible. However, here the basic case is assessed.

IP addresses are related to a computer to send and receive data packages over the internet. The IP address is connected to the Internet and is related to one natural person who pays the access provider. An address for a smart contract is permanently related to a specific contract and can be used by others who use the same communication protocol to transfer information to or receive information from this address. Although contracts don't have "owners" as IP addresses usually have, a contract will be used by a natural person, sooner or later, directly or indirectly. Like an IoT device with its own IP address, as long it is just the device, the IP address is anonymous. But if a log file contains that an IP address related to an individual has requested information from the IoT device, this information becomes PII because it is connected to personally identifiable information.

All in all, personal data on a public blockchain is considered at least pseudonymous data and therefore PII. Further, private or consortium blockchains, such as Corda, will face the same challenges as any other outsourcing to technical service providers.

3 Main obligations under GDPR

Current applicable data protection laws in the European Union are mainly based on the EU Directive 95/48/EC ("Data Protection Directive") and 2002/58/EC ("ePrivacy Directive"). This means that each European Member state has its own data protection laws, although harmonization and free flow and movement of data within the EU was one main goal of the EU privacy regulation (Art. 1 (2) Data Protection Directive and Art. 1 (1) ePrivacy Directive). The result has been a "privacy patchwork", though, which the European legislators have again sought to harmonize with the GDPR.

However, the GDPR also contains so called "opening clauses" regarding specific topics. They grant Member States leeway regarding various fields, so that Member States are entitled to address their local specific needs by additional national laws. These laws and a revision of the ePrivacy Directive[6] must come into effect on the same day as the GDPR, but they are still being drafted. That's why the following explanations concentrate on the obligations and requirements under the GDPR deemed most important for blockchain companies.

3.1 Right to erasure and right to be forgotten

The "right to be forgotten" was the subject of a very famous ruling of the European Court of Justice in 2014 (Google Spain) and is now codified in Art. 17 par. 2 GDPR. It is a part of a bundle of rights that an individual has under the GDPR and a subcategory of the right to erasure (Art. 17 par. 1 GDPR). Its basis is Art. 8 par. 2 of the Charter of Fundamental Rights of the European Union under which everyone has the right to have the data rectified, which has been collected concerning him or her.

3.1.1 Preconditions

Generally, under Art. 17 par. 1 GDPR an individual has the right to demand an immediate erasure of his or her PII, if

- the processing of these data is not necessary anymore for the pre-defined purposes;
- the individual has withdrawn his or her consent or objected to any data processing in the future;
- the data processing was illegal;
- the controller is obliged to delete these data under EU or member state law; or
- the individual was a child in the moment of data collection and processing.

However, there are exemptions defined in Article 17 par. 3 GDPR which give the controller means to balance the affected rights so that the controller does not need to erase data:

[6]For the provisions of this new ePrivacy Regulation particularise and complement the general rules on the protection of personal data laid down in the GDPR as regards electronic communications data that qualify as personal data, see Recital 5 Proposal for a regulation of the European Parliament and of the Council, 10 January 2017 COM (2017), 10 final 2017/0003 (COD).

- for exercising the right of freedom of expression and information;
- for compliance with a legal obligation which requires processing by Union or Member State law;
- for reasons of public interest in the area of public health;
- for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes; or
- for the establishment, exercise or defence of legal claims.

If data must be erased the controller shall communicate the erasure to each recipient to whom the personal data have been disclosed (Art. 19 GDPR).

3.1.2 Erasure of public data

The right to be forgotten contains not just an obligation of the controller to delete data in the situations mentioned above. It is only applicable to public data and includes a special obligation how to inform other controllers who are processing these data:

"Where the controller has made the personal data public and is obliged pursuant to paragraph 1 to erase the 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." (Art. 17 par. 2 GDPR)

Firstly, the right to be forgotten only applies, if the controller has made the respective data public. What "public" means is not defined in the GDPR. However, the judgements in the cases of Google Spain and Bodil Lindqvist and the defined material scope of the GDPR that excludes so called "household activities" (Art. 2 par. 2 (c) GDPR) indicate that "public" means "data are made accessible to an indefinite number of people". So, making information publicly available on a website or in any other electronic or non-electronic way means it opens up the application of the right to be forgotten.

Secondly, the controller must take reasonable steps to inform other controllers, which are processing this data. The information must contain that the data subject has requested the erasure of his or her data and any link or reference to it. The controller is free on how to implement this communication, taking technical and financial aspects into account, as long as the dissemination of the information, especially the erasure of references (e.g. links), is effective. The difference to the notification obligation under Art. 19 GDPR is that, in the case of public data, the controller does not know each recipient of this data. That's why he must take reasonable and appropriate measures to communicate the data erasure by deleting any reference to it that the controller has in his hands.

3.1.3 Erasure of blockchain data

The right to erasure under the GDPR clashes heavily with the fact that blockchains are intended to maintain data immutability.

There is no general exemption under the GDPR that allows the controller to keep data in a blockchain because it is technically unfeasible for him to delete it. Moreover, controllers are forced to design the operating systems in a privacy friendly manner (Art. 25 GDPR) so that any system is set up in accordance with the legal obligations and principles under the GDPR.

The available technologies and implementation costs are only relevant for public data and for the right to be forgotten. But that special argument applies only to the "references to PII". That means, nevertheless, that the controller must ensure the erasure of the relevant PII – like any other non-public data processing.

There is just one important exemption a controller can pull to deny an individual's request to delete his or her data: European or Member State law sets forth that data must be kept (e.g. book keeping, archiving obligations etc.) (Art. 17 par. 1 (e) GDPR). That works especially for commercial blockchains (e.g. banking, health or automotive industry). However, even in these cases, after the retention period elapses, data must be erased.

a) Data erasure and private blockchains

Generally, a way to comply with the GDPR is to exclusively use pseudonymous data in the blockchain and refrain from processing clear PII in the blockchain. The latter can be stored and processed outside the blockchain together with or separate from a reference table.

If the controller must delete any data of an individual, he simply deletes clear PII, so that he is unable to create a reference to the individual anymore. If it is fine with the individual to process PII but the processing of it in the blockchain is an issue, only the reference table is to erase. In both cases, former pseudonymous, i.e. personally identifiable information in the blockchain converts to anonymous data to which the GDPR does not apply anymore.

b) Data erasure and consortium blockchains

Data erasure effectively works with a private blockchain within one company. If it is a joint private operation or consortium that involves several companies at least the others are still able to refer to the individual. Consequently, the data in the blockchain remains PII even if one company has deleted the PII on its side.

So, if more than one company uses PII in a blockchain the controller should reply to an individual's request with the question, if all PII controlled by the addressee shall be erased or just data processed in the blockchain. In the first case, the addressee must delete all PII and the reference table, in the second case the addressee must additionally notify the other controllers to delete the reference table.

c) Data erasure and fully public blockchains

Finally, concerning fully public blockchains, PII is processed in the distributed network that "belongs" to no specific person or company.[7] So, an addressee of an individual's request to erase data is only able to delete PII he controls but not throughout the entire network. Unlike the Google Spain judgement, there is no specific controller who can be forced to ensure that data is no longer accessible.

Thus, a controller using a fully public blockchain must only delete PII on his side (see above). Bob can't be called to account for PII processing on the entire public blockchain. However, he remains responsible if he makes PII public on the blockchain and therefore transfers PII to third parties; he should not try to story any PII on the blockchain.

3.2 Data processing agreements

The GDPR distinguishes between a controller-processor and a controller-controller-relationship. To understand explanations on that topic it's important to remember the following:

If PII is subject of the agreement or data flow between parties and these data are deemed to be PII for just one party, the entire agreement and/or data flow falls within the scope of European privacy regulation. Any company that runs an IT project, especially an outsourcing project, must consider the need of special privacy clauses and/or agreements.

3.2.1 Controller-processor relationship

A data processing agreement ("DPA") is an agreement that contains that a party processes PII on behalf of another party. A DPA is mandatory under the current law, in Art. 17 (3) Data Protection Directive, if no other statutory provision legalizes the processing by a processor, and under future European law if there is a controller-processor-relationship (Art. 28 (3) GDPR).

A controller is the legal or natural person that determines the purpose and means of the processing of PII (Art. 4 (7) GDPR). The commissioned party can be a legal or natural person and is called a processor (Art. 4 (8) GDPR).

The key element of a DPA is, that the processor acts on behalf of the controller. To assess, in an individual case, if a DPA is needed (i) the technical set up and the technical architecture of the data storage and flow, (ii) the access to the processed PII, and (iii) the contractual relationship between the parties must be considered.

[7] If an agent of influence can link mining pools and farms back to real world identities, they could coerce a sufficient proportion of the hashrate to mount a 51% attack and rewrite the chain and/or censor transactions.

In simple terms, a company is a processor under Articles 4 (8), 28 GDPR if

- it *manages and/or stores and/or otherwise processes* PII for his client. (e.g. dedicated servers, cloud computing (SaaS));
- it has *authorized access* to the servers run by or on behalf of the controller. (e.g. server administration, IT security, development)

A company is not considered as processor, if

- it has no access to the databases at all (e.g. simple rack space rent);
- it receives or has access to data to use it, *but not on behalf of the controller* (e.g. data providers and sellers)

In a DPA relationship the controller is literally "in control" of the purposes and means of data processing. That's why he must comply with various obligations under the GDPR. A controller shall use processors only if they provide sufficient guarantees to implement appropriate technical and organizational measures (Art. 28 (1) GDPR). The GDPR contains further obligations of the controller concerning the circumstances and means of the data transfer to or the data access by a processor, e.g. a contract under Art. 28 (3) a) GDPR including regulations regarding sub-contractors, detailed description of technical and organizational measures, and the processor's obligation to comply with the controller's instruction.

3.2.2 Controller-controller relationship

The opposite of a DPA relationship is a controller-controller-relationship or "joint controllership" (Art. 26 (1) GDPR):

"Where two or more controllers jointly determine the purposes and means of processing, they shall be joint controllers. [...]"

A joint controllership characterizes that more than one natural or legal person is responsible and defines the purposes and means of data processing. Both parties are independent from each other and fully obliged under the GDPR. Working together they must define in a contract who is in charge of which legal requirement under the GDPR and especially towards the data subject. The essence of the agreement shall be made available to the data subject (Art. 26 (2) GDPR).

3.2.3 Fully private blockchains

From a privacy perspective, the main characteristic of a fully private blockchain run by one single organization is an elaborated control of access permissions. That means that not every involved party is able to write or amend data on the blockchain because the write permissions are kept centralized whereas read permissions may be public or restricted to an arbitrary extent.[8]

Consequently, the centralized entity can determine how PII is processed and which data is accessible and readable. So, the centralized party is the controller under Art. 4 (7), 28 GDPR as it determines the purpose and the scope of data processing.

This means, in a fully private blockchain, the relation between the central controlling entity and other parties depends on the specifics of their relation, although in most cases it will be a controller-processor-relationship such that a DPA is needed.

3.2.4 Consortium (private) blockchains

A consortium blockchain is defined by a pre-selected set of nodes which are run by several entities who jointly control the consensus process. From a privacy perspective the joint approach of a consortium that runs a private blockchain is the relevant distinguishing factor, because the participants of a consortium need each other to control the consensus process.

Thus, the blockchain consortium is a joint controllership under Art. 26 GDPR, that must comply with the legal stipulations the GDPR imposes on joint controllers.

[8] A caveat is an organization running a centralized blockchain - depending on how it is architected - could in theory change the rules of a blockchain, revert transactions, modify balances.

These processes are specifically needed to:

- sign an agreement and jointly agree on
  - the purposes and means of data processing,
  - determine who is in charge of which obligation under the GDPR, esp. who must inform the data subjects under Art. 13 and 14 GDPR,
  - who grants the data subjects' statutory rights (e.g. right to delete and correct data);
  - a description of the relation between the parties and their roles towards the data subjects;
- provide the essence of the agreement to the data subject;
- be prepared that the data subject may exercise his or her rights under the GDPR in respect of and against each of the controllers.

3.2.5 Fully public blockchains

A public blockchain is a blockchain that anyone in the world can read, anyone in the world can send transactions to and expect to see them included if they are valid, and anyone in the world can participate in the consensus process – the process for determining what blocks get added to the chain and what the current state is. That means that all participants have agreed on the same consensus process. However, an entry or exit of a participant does not affect the process. Furthermore, the processed data is publicly available on the network and the purpose of data processing is not being determined between the parties. They only agree on how data is being processed.

From a privacy perspective, this structure is comparable with linked websites on the Internet. Using specific protocols and programming formats the website providers agree that websites can be linked and interact with each other (e.g. weblinks, iframes etc.). However, the website owners have no relation to each other nor does a deletion of a website affect the other websites. Consequently, they are all independent data controllers and no joint controllers.

To conclude, parties of a fully public blockchain are neither processors nor joint controllers but independent controllers. That means that they must comply with GDPR requirements independently from each other. They must assess separately from each other if they are allowed to process PII or not, including making PII accessible via the public blockchain and any transfer of data to the other members via public blockchains.[9]

3.3 Data transfer

The goal of the GDPR is to harmonize the data protection level in the European Union. Any processing outside of the EU, including any access from a third country, is subject to special preconditions under the GDPR as most privacy and data protection regulations in third countries are deemed to be of "lower" or at least "non-equivalent" standard. That is why instruments like "Privacy Shield".[10] (Art. 45 GDPR), contractual model clauses[11] (Art. 46 par. 2 (c) and (d), 93 par. 2 GDPR) or corporate binding rules (Art. 47 GDPR) are subject of the GDPR to ensure an adequate level of privacy outside of the European Union (Art. 44-49 GDPR).

Companies using blockchains – public or private – face the same challenges regarding the transfer of data outside the European Union as any other company. In each case the transfer of data needs a thorough legal assessment. However, using a fully public blockchain it will be impossible to comply with the GDPR because the data recipients are unknown as is as the location where data is processed. Without this information no company can ensure an adequate level of data protection.

[9] In practice this would rule out using a public blockchain for any personal data. For instance, if Alice gives Bob a USB stick with data, Bob possesses this data. Bob must ensure that he is entitled to store or use it. Bob can't just refer to the person who gave him the stick. If he is not allowed to use these data Bob must delete the data or not accept any "defective data set" i.e. data that can't be used legally.

[10] For more information on the Privacy Shield Frameworks visit here.

[11] Model contracts for the transfer of personal data to third countries can be downloaded here.

3.4 Data protection impact assessment

Coming into effect the GDPR will introduce a new obligation for controllers called "data protection impact assessment" (DPIA or PIA for "privacy impact assessment") (Art. 35 GDPR). That means that controllers must carry out an assessment of the impact of envisaged processing operations on natural persons.

3.4.1 Scope

A DPIA is not comparable with an assessment of the impact on anything else, such as reputation of the institution or agency, information security, or any other general risk assessments. The supervisory authority established in the EU Member States establishes and makes public a list of the kind of processing operations which are subject to the requirement for a data protection impact assessment. The authority shall also communicate those lists to the European Data Protection and establish and make public a list of the kind of processing operations for which no data protection impact assessment is required and also communicate those lists to the European Data Protection Board (Art. 35 par. 3 and 4 GDPR).

3.4.2 Applicability

Generally, the obligation to run a DPIA applies only if the operation is likely to result in a high risk to the rights and freedoms of natural persons. A DPIA is in particular required in cases of

"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"

and also in cases of processing on a large scale of sensitive data, (e.g. biometric and health data, data about sexual interests or ethics) criminal data, or systematic monitoring of publicly accessible areas on a large scale (Art. 35 par. 2 GDPR). So, for instance, systematic and extensive solvency checks on natural persons by a bank in the course of credit approval processes may impose a DPIA as well as an extensive collection of driving behavior of car owners to use these data for car insurance.

The controller shall consult the competent supervisory authority prior to processing where a DPIA indicates that the processing would result in a high risk in the absence of measures taken by the controller to mitigate the risk. It is a question of each individual use case whether a DPIA is necessary.

3.4.3 DPIA for blockchain projects

Blockchain projects are not special: blockchains face the same challenges and requirements as any other IT project. In most cases a systematic and extensive evaluation of personal aspects relating to a natural person is the subject of blockchain based processing on which decisions are based that produce legal effect or similarly affect a natural person.[12] However, whether a DPIA is needed depends on each single project and use case.

It is important, though, that the obligations regarding a DPIA must be fulfilled by each controller. It is not possible to trust the DPIA of other controllers, like other controllers in a blockchain network or consortium.

3.5 Data protection officer

From 25th May 2018, a company[13] is obliged to designate a data protection officer (DPO), if its core activities consist of processing operations (1) which, by virtue of their nature, their scope and/or their purposes, require regular and systematic monitoring of data subjects on a large scale; or (2) on a large scale of special categories of data (Art. 9 GDPR) or personal data relating to

[12] It does not matter if Bob uses smart contracts or simply stores tracking data on the blockchain. The purpose of the data processing is relevant, i.e. if the operation is likely to result in a high risk to the rights and freedoms of natural persons.

[13] Special requirements and provisions apply to a public authority or body (Art. 37 GDPR).

criminal convictions and offences. Union or Member State law may define additional circumstances when a company must designate a DPO (Art. 37 par. 4 GDPR).

3.5.1 Designation

The DPO must be designated on the basis of professional qualities and, in particular, expert knowledge of data protection law and practices and the ability to fulfill the tasks referred to in the GDPR (Art. 37 par. 5 GDPR). It doesn't matter if the DPO is an employee or a commissioned expert (Art. 37 par. 6 GDPR). Furthermore, a DPO who is also an employee of the respective company may fulfill other tasks and duties within that company but these obligations must not result in a conflict of interest (Art. 38 par. 6 GDPR). This would be true for instance, if a head of the IT department or legal department were to be the designated DPO.

There is no designation form (e.g. written, text form) set forth in the GDPR. And a group of undertakings may appoint a single DPO provided that a DPO is easily accessible from each establishment (Art. 37 par. 2 GDPR). Anyway, the company must publish the contact details of its DPO, typically in the information to the data subject (Art. 13 par. 1 (b), Art. 14 par. 1 (b) GDPR) and must additionally communicate them to the supervisory authority (Art. 37 par. 7 GDPR).

A DPO may not be dismissed or penalized for performing his tasks under the GDPR (Art. 38 par. 3 GDPR). However, it is possible to limit the term of a DPO contract and/or regularly terminate a DPO contract. Member States may not set forth further obligations or limitations concerning the designation of a DPO under Member State law.

3.5.2 Position and tasks

A DPO is an independent body with a direct report to the highest management level. He or she is free of any instruction regarding the DPO tasks and bound by secrecy and confidentiality by law. The DPO's main tasks are exclusively described in Art. 39 par. 1 GDPR as follows:

"(a) to inform and advise the controller or the processor and the employees who carry out processing of their obligations pursuant to this Regulation and to other Union or Member State data protection provisions;

(b) to monitor compliance with this Regulation, with other Union or Member State data protection provisions and with the policies of the controller or processor in relation to the protection of personal data, including the assignment of responsibilities, awareness-raising and training of staff involved in processing operations, and the related audits;

(c) to provide advice where requested as regards the data protection impact assessment and monitor its performance pursuant to Article 35;

(d) to cooperate with the supervisory authority;

(e) to act as the contact point for the supervisory authority on issues relating to processing, including the prior consultation referred to in Article 36, and to consult, where appropriate, with regard to any other matter."

Overall, the DPO supports company's compliance with data protection laws. In the course of this, he or she shall have due regard to the risk associated with processing operations, taking into account the nature, scope, context and purposes of processing (Art. 39 par. 2 GDPR). That means that a DPO's assessment must include the risks for the data subject as well as the needs and reasons of the operating company. The same applies to any communication of the DPO with the supervisory authorities or third parties who contact him concerning data protection issues.

3.5.3 DPO in blockchain projects

Usually, the question as to whether a DPO must be designated is independent from the type of technology implemented. Only the scope and/or purposes of data processing as well as the category of processed data is relevant. So just using blockchain technology does not change the legal assessment concerning the requirement for a DPO.

If a company must designate a DPO under GDPR it is not sufficient to refer to the DPO of another company in a blockchain consortium. Only companies with a corporate law relationship are allowed to "share" a DPO. This would also be true for an acquired blockchain company, for instance, which doesn't have a DPO and provides the platform for other companies of its group.

3.6 Investigative powers and administrative fines

The supervisory authorities have various investigative powers[14] to enforce compliance with the GDPR, e.g. request information, issue reprimands, impose a limitation including a ban on processing (Art. 58 GDPR). And Member States are entitled to enact further powers for the supervisory authorities (Art. 58 par. 6 GDPR).

In the event of an infringement of the statutory requirements under the GDPR the supervisory authorities can impose administrative fines up to 20 million Euro, or in the case of an undertaking, up to 4% of the total worldwide annual turnover of the preceding financial year, whichever is higher (Art. 83 (6) GDPR).

Figure 3: Overview of fines under GDPR in the most relevant cases.

	Fines up to 10 mil. EUR or 2% annual turn over	Fines up to 20 mil. EUR or 4% annual turn over	Fines up to 20 mil. EUR or 4% annual turn over
Principles of data processing (Art. 5-7, 9 GDPR)	n/a	Art 83 par. 5 (c)	Non-compliance with an order by the supervisory authority, Art. 83 par. 6, 58 par. 2 GDPR
Right to erasure and right to be forgotten (Art. 17 GDPR)	n/a	Art 83 par. 5 (b)	
Controller-processor relationship and data processing agreement (Art. 28 GDPR)	Art 83 par. 4 (a)	n/a	
Data privacy impact assessment (Art. 35 GDPR)	Art 83 par. 4 (a)	n/a	
Data protection officer (Art. 37 GDPR)	Art 83 par. 4 (a)	n/a	
Data transfer outside of the European Union (Art. 44-49 GDPR)	Art 83 par. 4 (c)	n/a	

[14] Infringement of data protection law can lead to claims of the individuals (e.g. under consumer protection laws) as well as class action suits, and competitors could file a suit based on competition law.

References

Bodil Lindqvist v Åklagarkammaren i Jönköping (2003). European Court of Justice. Case C-101/01. http://curia.europa.eu/juris/liste.jsf?num=C-101/01.

Buterin, Vitalik (2016). Ethereum: Platform Review, Opportunities and Challenges for Private and Consortium Blockchains. R3 Research. https://www.r3.com/research/.

European Commission. Article 29 Data Protection Working Party, WP 136 (2007). http://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf.

European Commission. Article 29 Data Protection Working Party, WP 148 (2008). http://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2008/wp148_en.pdf.

European Commission. Article 29 Data Protection Working Party, WP 216 (2014). http://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

European Union Directive 2002/58/EC ("ePrivacy Directive") (2002). European Parliament. / http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:en:HTML.

European Union Directive 95/48/EC ("Data Protection Directive") (1995). European Parliament. http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:31995L0046.

General Data Protection Regulation (2016). European Parliament. Regulation (EU) 2016/679. http://eur-lex.europa.eu/eli/reg/2016/679/oj.

Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González (2014). European Court of Justice. Case C-131/12. http://curia.europa.eu/juris/liste.jsf?num=C-131/12.

Langfeldt, Owe, and Tereza Struncova (2016). Data Protection Impact Assessment. DPO-EDPS meeting. https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/DPO_Corner/Trainings/16-10-27_TS_OL_presentation_EN.pdf

Wood, Gavin (2014). Ethereum: A secure decentralized generalized transaction ledger. EIP-150 Revision.

R3 is an enterprise software firm using distributed ledger technology to build the next generation of financial services infrastructure.

R3's member base comprises over 80 global financial institutions and regulators on six continents. It is the largest collaborative consortium of its kind in financial markets.

Consortium members have access to insights from projects, research, regulatory outreach, and professional services.

Our team is made of financial industry veterans, technologists, and new tech entrepreneurs, bringing together expertise from electronic financial markets, cryptography and digital currencies.

corda is an open source, financial grade distributed ledger that records, manages and executes institutions' financial agreements in perfect synchrony with their peers.

Corda is the only distributed ledger platform designed from the ground up to address the specific needs of the financial services industry, and is the result of over a year of close collaboration between R3 and its consortium of over 80 of the world's leading banks and financial institutions.