Abstract: The compatibility between emerging technology such as Blockchain and the new data protection obligations has been a matter of discussion since the European General Data Protection Regulation (“GDPR”) came into force back in 2018. The main reason is that GDPR is considered the Regulation that has incorporated the strongest obligations and enforcement consequences. There are some areas of tension on how to comply with these obligations, some of which are still a matter of discussion that have not been concluded by regulators or in court. In this chapter, I set out an overview of the main obligations, relevant areas of tension and set out an opinion on what I have seen that in practice is working the best, and how I recommend dealing with these issues in practice, in order to do an assessment applied to a particular context in which a Blockchain network is used.

Keywords: Blockchain; Privacy; Data Protection Laws; General Data Protection Regulation (GDPR); Data Protection

1. Introduction: Privacy Laws landscape towards GDPR

The European Union’s General Data Protection Regulation (“GDPR” or “Regulation”) [1] became binding in May 2018. Based in the previous data protection regime, it introduced additional obligations placed on organisations that process personal data either on their own discretion or under the instructions of other organisations.

The GDPR main goals are to facilitate free movement of personal data between the European Member States at the time that ensures a safe environment to protect people whom the personal data is about (“data subjects”) in line with the fundamental right to data protection (Article 8 of Charter of Fundamental Rights of the European Union) [2] and the human right to Privacy (Article 12 of the 1948 Universal Declaration of Human Rights) [3].

The application of the GDPR is considered by entities irrespective of their worldwide location due to the fact it has been provided with an extraterritorial effect. This means that European data protection regulators have the power to enforce this Regulation over entities that have no physical presence in the European Union. However, GDPR will not apply by default to every organisation processing personal data of European citizens. The Regulation will only apply to:
organisations that are offering goods or services to individuals that are located in the European Union (irrespective of whether or not there is a payment involved);
organisations that monitor individuals while they are in the European Union; and
organisations that have some sort of presence in the Union (for example by way of having an establishment which might be just an office or a branch).

Moreover, privacy is not only a matter of interest of the European countries, and there are other jurisdictions in which a similar approach has been implemented or it is currently being considered by their central governments. For example, each of the countries that the European Union has certified as offering an adequate level of protection of personal data (the "Adequacy Decision") [4], have been approved on the basis that, to some extent, there are similar principles and obligations in place. These countries are, to date: Andorra, Argentina, Canada (only covering commercial organisations), Faroe Islands, Guernsey, Israel, Isle of Man, Japan, Jersey, New Zealand, Switzerland, Uruguay and the United States of America (limited to those organisations that are Privacy Shield certified only). In addition, South Korea is currently under talks in order to be declared adequate.

In addition, there are other legislations protective of the data that organisations process, as it occurs in countries like China, Singapore, India, Brazil and Dubai.

The US is also taking steps to mirror the GDPR standards and so California has approved the California Consumer Privacy Act (the "CCPA") [5], and countries like Texas, New York and Washington are attempting to follow this path.

For the purposes of this article we will refer to the legal regime that shares the basics of GDPR as "the data protection laws".

2. The general data protection principles

To meet general standards of compliance when processing personal data, organisations need to determine the purposes for which a category of personal data is used. This is known as the purpose limitation principle and although it is not the first principle in the GDPR list (Article 5 of GDPR) it is, to my view, the most relevant one because a particular purpose needs to be defined in order to ensure compliance with the remaining principles.

For example, if a research institution plans to contact the subjects of a trial to seek for an update of how they feel after trying a product for a period of time, the institution will need the subjects' contact details. The institution may say that it needs to contact the subjects of the trial to complete the research exercise, however the purpose for which contact details are collected is "to contact the subject and gather updated information", and the objective or the reason why this is required is because otherwise the research exercise cannot be completed.

Once the purposes are determined, then the following actions should be considered to meet the remaining principles:

- To find a legal basis allowing the use of data (in GDPR this is stated in Articles 6 and 9);
- To be fair and transparent with data subjects (so they will need to be informed of the processing activities carried out, amongst other things);
- To only use the personal data that is necessary to achieve the purpose (known as the minimisation principle);
- To ensure that the data processed is accurate;
- To delete the data when this is not needed anymore;
• to put in place appropriate technical and organisational measures to ensure the confidentiality and resilience of the data; and
• to be able to demonstrate compliance with these principles (the so-called "accountability principle").

Since many global organisations need to comply with GDPR and the data protection regimes in other jurisdictions, finding the ways to achieve harmonised and pragmatic compliance is becoming a crucial task in order to ensure that the data processed meets standards that are becoming more global with the time. Therefore, a recommended starting point is setting out rules and procedures based on the above principles in order to build up compliance from there.

3. Additional obligations

The GDPR also places certain obligations on the organisations that determine the means and purposes of the processing of personal data ("data controllers"). Whereas the legislation in other countries do not refer to these categories of organisations as "data controllers", they usually incorporate similar obligations for the organisations responsible for the processing of personal data.

In addition, the development of new technology such as Blockchain facilitates the collaboration of multiple parties on one Blockchain. Some of these parties will be either subcontracted by data controller and processing data according to the specific and detailed instructions of the latter. This type of role when it comes into the processing of personal data is defined under GDPR as a "data processor" (who processes personal data in a way in which is determined by a data controller, and for the purposes for which the data controller decides).

One of the novelties introduced by the GDPR is that it places direct obligations on both controllers and processors.

Although the list of obligations placed on controllers and processors is longer than what we stress below, the most relevant obligations that impact on projects where Blockchain technology is involved are, for data controllers:

• To identify which party or participant might be a data processor and put in place the mandatory agreements with data processors (Article 28 GDPR);
• To ensure that there are suitable mechanisms in place allowing the International transfers of personal data (Articles 44 to 50 GDPR);
• To ensure that data subjects can exercise their rights (Articles 12 to 23 of the GDPR);
• To identify the scenarios in which there is more than one data controller jointly determining the purposes for which and the manner in which personal data is processed, and if this is the case, to put in place suitable arrangements in order to set out the responsibilities of each joint controller (Article 26 of the GDPR); and
• To comply with the minimisation principle.

For processors, the mandatory scheme differs and is more focussed on following the controller's instructions and ensuring the security of the data including complying with the requirements for International transfers of personal data.

Below I set out the areas of tensions identified by regulators and their most relevant views on how to deal with these principles in a Blockchain environment, as well as the pragmatic solutions I have to date recommended in similar scenarios that I found they worked the best in practice.
4. Data subjects’ rights

When personal data is processed in a Blockchain environment individuals whom personal data concerns will interact with controllers due to their entitlement to exercise data protection rights. These rights, depending on the jurisdiction may include (amongst other rights) the right to request for a copy of their personal data, the right to request a rectification of information that in their opinion is incorrect, the right to object to certain processing of personal data should in their opinion does not meet the legal requirements, the right to request for their data to be erased (known as the "right to be forgotten") and the right to issue a complaint to the competent data protection regulator.

Indeed, one of the main concerns in the Blockchain world has been how to ensure the exercise of some of these rights. Below I explain the pragmatic solutions I have been considering to deal with this issue.

5. Privacy Laws applied to Blockchain: Areas of tension and actions to consider to mitigate risks

It has been argued that Blockchain technologies may be unable to comply with European data protection law (and by extension, to similar legal regimes) due to the Blockchain very own nature.

In order to understand what is causing these areas of tension it is relevant to note that while the protection of personal data has been considered in Europe for more than twenty years, being GDPR a result of how the previous legislation has been interpreted by the Courts, we cannot ignore the fact that the GDPR is still quite a novel piece of legislation which has incorporated or revised terms and obligations (such as what is a "special category of data", or what obligations affect a "joint controller" scenario).

This means that there is an expectation on how the Courts and regulators will interpret some of these terms when they are applied to a Blockchain case scenario. And so, the sector is both keeping an eye on case law developments and also demanding regulators to issue further guidance and agree on certification schemes which brings clarity and a harmonised approach. If we usually affirm that certain technology is still in an early stage subject to testing, investigation and further developing, I would say that so the development of privacy laws are.

My view is that while there will be, with no hesitation, challenges to face, it is possible to use Blockchain and achieve a reasonable level of compliance with the data protection laws and in particular in line with the GDPR principles on which I focus in this chapter.

Having advised on complex global projects, one thing I recommend bearing in mind at each stage of a Blockchain project is something as simple as coming back to the fundamentals of data protection.

One of the tasks that I found very useful in practice is focussing on responding to questions that can be applied to any scheme in which Blockchain is used, irrespective of its nature, whether private or public, the type of persons involved, reasons why and type of personal data processed. And so I produce checklists for organisations to use, ideally when the use of Blockchain is still under consideration. These include questions such as:

- What do we want to achieve?
- What data needs to be stored or otherwise used on the chain to meet this purpose?
- Is there any personal data?
- We check it twice: Are we sure that this data is not personal data (e.g. keys are most likely to be personal data)?
We are confident that under no circumstances we can identify a person by putting together this data and any other information we may have access to from different sources.

- Who are the people to whom the data concerns (participants, miners, others?)
- What can be stored off chain?
- What technology, other than Blockchain, could I use to achieve the same purpose?
- If other technology is suitable, why do we need to use Blockchain?
- How do we obtain this data?
- How will the data be verified?
- Who else will need to use this data and in which countries are they located?
- Can we achieve the same goal if all information is:
  - Anonymised,
  - Encrypted,
  - Obfuscated?
- If this is a privacy Blockchain, do we really need to connect it with a public Blockchain?
- How can we inform individuals on how their personal data is used on this Blockchain, or for the purposes of this Blockchain?
- How can we keep individuals informed on a regular basis?
- What are the views of our legal colleagues or the person responsible for data protection in our organisation / project?

This will help to get a clear overview of the types of personal data used, reasons why, parties involved, alternatives and (human and technical) resources available, which are, in my view, crucial to take risk-based approach decisions when dealing with the areas of tensions most discussed by regulators and experts in the field: "How to decide what personal data is used on Blockchain?"; "How to allocate responsibilities between the parties?"; and "How to guarantee the exercise of data subjects' rights?"

Once the project is considered and overviewed from a data protection perspective (by discussing over the above, and related questions), then it is relevant to pay attention to the main areas of concern. Going through these areas of concerns after discussing and agreeing on these questions it is much effective in my opinion because there will be then a clear picture on things like how important it is to use personally identifiable information (commonly known as "PII", it is the other term used to refer to personal data, meaning any information that can be used to potentially identify an individual, from that information, or from putting together such information with any other available data) or whether it really needs to be stored on the chain, and so, agreeing on the appropriate approach turns to be a more straightforward route.

Below I focus on the three aforementioned areas of tension only because in practice, I saw these causing a greater level of concerns up to the present time.

6. How to decide what personal data is used on Blockchain?

The use of personal data needs, by default, to be considered in a "need to be used" basis. This means that it should be kept to the minimum amount of data that is necessary to achieve each of the purposes for which Blockchain technology is put in place.
In order to comply with the minimisation principle organisations should consider the implementation of anonymisation procedures to ensure that only when it is strictly necessary, personal data is stored on the Blockchain network. The anonymisation of personal data is one of the areas causing a greater level of discussion to agree on the steps that need to be taken because while it is certain that GDPR does not apply to data that is "fully" anonymised, for data to be qualified as "fully" anonymised there needs to be certainty that not a single person can be identified and that there are no re-identification process allowing the process to be reversible. Hence, not all data that has been subject to an anonymisation process is out of the scope of the application of GDPR. For example, the French data protection regulator ("CNIL"), who has issued guidance on the application of the GDPR on Blockchain and smart contracts technology [7] recommends registering personal data:

a. in the form of a commitment (defined by the CNIL as the "cryptographic mechanism that allows one to "freeze" data in such a way that it is both possible – with additional information – to prove that has been frozen and impossible to find or recognise such data by using this sole "commit").); or if this is not suitable

b. in the form of a hash using a hash function with a key; or

c. encrypting the data by choosing an encryption option that ensures a high level of confidentiality.

The solution I recommend here to minimise risks of breaching the law and/or facing a data breach incident, is anonymising the personal data to the maximum extent that still allows the Blockchain achieve it purpose. Then, to analyse the anonymised data from a testing point of view that the UK data protection authority (Information Commissioner’s Office or ICO) calls: "the intruder test" [7], according to which the likelihood of identifying a person from anonymised information is considered from the perspective of an intruder trying to do so by compelling the anonymised information with any other information that it is easily available and in the public domain. My view is that when in doubt, the data should be treated as personal data. Irrespective of the outcome such a test being carried out, it is also recommended to combine encryption technology with the use of anonymisation technique.

Many regulators and experts on the sector agree on that another relevant solution that in my view should be considered in addition to the anonymisation and encryption techniques is storing the personal data off-chain where possible, and only including personal data on-chain if this is necessary, for example when the existence of that personal information needs to be proved and Blockchain is used for this purpose.

Before using personal data on-chain, under GDPR it is mandatory to carry out a data protection impact assessment ("DPIA") to assess the risks associated to the use of Blockchain and consider what measures should be put in place to minimise the risks to a level to which it is acceptable in the data protection field.

Again, the questions above and the consideration of the data protection principles applied to the context, are a useful starting point should a DPIA needs to be carried out. Another relevant element here is being accountable (since it is also mandatory under GDPR). Thus, I always recommend keeping records of the relevant considerations and discussions over the questions and the risks to the individuals that may be stressed, as well as the decisions taken to mitigate or minimise those risks. For example:
• When anonymising data: Keeping records of the procedures followed, an intruder test carried out, a conclusion as to whether the data is, in the reasonable opinion of the people involved, fully anonymised, and if it is not fully anonymised, the additional safeguards in place and reasons why.

• If personal data can be stored off-chain: Records of what personal data will be kept off-chain, reasons why, measures to avoid such data being added by participants, measures to reverse personal data that should be off-chain and has been added and validated on the Blockchain by using hashing technology, and the impact of the use of data off the chain on the decentralised used of data. On this point, perhaps it might be necessary to explore whether this operation can be duplicated so there are two sets of off-chain data with two different hashes linking this data, so if one set of data is compromised, the personal data is not lost. If this is debated it will be relevant to keep records of the final decision made and the reasons why.

• If considering obfuscation techniques, keep records of the rationale behind choosing an obfuscation technique (for example between third-party indirection services or ring signatures) all of which should be considered on a case-by-case basis, although I found that the use of ring signatures might have a greater impact on minimising risks in comparison with other alternatives. An analysis of the risks associated with each solution and choosing the safest alternative is crucial since it seems that none of the current alternatives fully mitigate all risks.

In a similar way, any decision taken towards encryption must be documented. It is relevant to note that encrypted personal data is most likely to be considered pseudonymised personal data. Therefore, the data protection laws will most likely still apply to encrypted data. The report published by the European Union Blockchain Observatory and Forum “Blockchain and the GDPR” (the Blockchain Report) [8] deals with this issue stating that while “hashed personal data is a grey area”, “reversibly encrypted personal data is personal data” and it stresses that with regards of the use of hashing, whether or not the outcome will be depends on the case. Therefore, when considering hashing as part of a Blockchain project it is recommended to pay special attention to the likelihood of data being personally identifiable information anymore. The report recommends to use the most advanced hashing algorithms so risks are minimised to the maximum extent.

7. How to allocate responsibilities between the parties?

In order to assess what party is responsible for complying with depending which data protection obligation, it is necessary to determine who is a data controller and whether any of the parties involved is a data processor, which needs to be considered on a case by case basis, depending on the technology applied, the purposes for which personal data is used that the ultimate role adopted by each party for each of the purposes for which personal data is used.

The challenge of confirming data controller and processor roles in a Blockchain context comes from the fact that originally the data protection laws were designed in a context in which the processing of personal data was centralised, and so, the entities responsible for the handling of personal data were easier to identify.

As the Blockchain Report points out, the Blockchain model differs from the original scheme considered by the data protection laws due to it is a decentralised based technology in which multiple parties/individuals are directly involved in the processing of personal data.
On this point, the CNIL’s guidance on Blockchain considers those who have some sort of administration functions over the chain (which capability to give access, send data for validation and add data) as data controllers. Although they will not be data controllers if the Blockchain is used only by individuals who are not using the technology for business purposes (e.g. a group of family members using it to build a database of their family tree –for purely household purposes only).

Therefore, the first step I recommend when dealing with this area of tension is clarifying whether:

- The Blockchain will be used for private, personal, purely household purposes (because in this context privacy laws are most likely to not apply- although my view is that while specific data protection obligations do not apply, the Human Right to Privacy must always be considered, even in this context); or
- Used for business purposes, in which case identifying who has administration functions and taking decisions on who is granted with access, how data is sent for validation, who validates the data and the level of hash rates, because those taking these decisions will be most likely to be data controllers.

Hence, individuals using this technology as part of their business functions are meant to be data controllers, as well as legal entities will. The CNIL brings some illustrative examples in its guidance: "For example, a natural person who buys or sells Bitcoin, on his on her own behalf is not a data controller. However, the said person can be considered a data controller if these transactions are carried out as part of a professional or commercial activity, on behalf of other natural persons."

The CNIL also recommends to constitute legal entities if a group of natural persons aim to use Blockchain technology to achieve a common goal (so they will jointly decide on the reasons why and the manner in which personal data will be processed). In this context the responsibilities of each party must be determined and a solution for this is putting in place a contract between all parties. The contract should determine the responsibilities of each of the participants when it comes to data protection obligations; for example, who is responsible for giving access to other participants, who is responsible for adding data, who will be responsible for determine hash rates or whether any data is stored or otherwise used off-chain. Also, if an individual wishes to exercise their data protection rights, who will be responsible for being the main point of contact, and prior to this, who will be responsible for informing all individuals affected as it is mandatory to comply with the transparency principle and the data protection laws. In a data protection legal context in which participants are taking joint decisions, putting in place this contract is either mandatory (e.g. under GDPR it is always mandatory to do so in this context) or otherwise highly recommended for the responsibilities and liabilities of each participants to be agreed up front.

It is also relevant to confirm who may be a data processor. An example of this would be when an insurance company contracts a developer to put in place a smart contract according to which passengers are automatically reimbursed if their flight is delayed, provided that the developer only processes passengers' details as instructed by the insurance company [8]. In this context, my recommendation are:

Firstly, focussing on the instructions that one party (the data controller) is giving to the other party (the data processor), because these instructions are what determine the extent to which the processor is a data processor. Note that these are instructions only in relation to the use of personal data, as opposed to the Blockchain technology per se (e.g. if a party instructs another party to use Blockchain for a purpose for which there no personal details are used, so the one receiving the
instructions has not have a discretion to decide on the number of participants, verification methods, etc. then none of them is a controller or a processor since these data protection rules would not apply because no personal data is involved).

The reason why focussing on the instructions is that important is because when a processor acts beyond the instructions (and hence takes decisions over the how and the why personal data is used), it becomes a data controller for that processing going beyond a data processor scheme. Also, because under GDPR the relationship between controllers and processors needs to be set out in a written agreement in which the inclusion of certain clauses is mandatory (as set out in Article 28 of the GDPR), and because a processor should not be engaged in a project unless it demonstrates that it can effectively comply with all GDPR obligations set out in such a mandatory contract.

A simple task that in my opinion helps identifying each of the above roles (independent controllers, joint controllers, or processors) and relationships is (to the extent that it is possible to do so in a Blockchain scheme) to send out a checklist where each participant will have the opportunity to consider, from their view, whether they are actually taking decisions on the processing of personal data used within the Blockchain, by responding to simple questions like:

- “Do you have administration functions?”
- “Do you decide who to grant with access to personal data?”
- “Do you decide how data is sent for validation”
- “Do you decide who validates the data?”
- “Do you use the Blockchain for your own purposes? If so, why?”
- “Are you receiving any instructions on what to do with the personal data that is on the chain? If so, whom from? Please detail the instructions”.

In some cases the roles are obvious and there is no need to get into this level of detail. However, in more complex Blockchain schemes this exercise helps to clarify positions. What I found to be a challenge in certain schemes and something that I have discussed long is whether to send questionnaires to every participant (as they may take a different view as to whether they receive instructions or not), and if so, how, in particular in public Blockchains. The solution I have been working towards this issue so far has been to send questionnaires to selected persons to cover, at least, people by categories given their functions performed in the Blockchain.

Putting in place the named Article 28 GDPR agreements with every data processor is also a challenge in certain Blockchain schemes, in particular in public Blockchains. If the technology allows it, a reasonable option I have been considering might be for controllers to include the terms of the contract on a chain and for all processors involved to validate it, the action of validation meaning an opt-in action in agreement of the terms and conditions and the instructions to follow.

8. How to guarantee the exercise of the data subjects’ rights?

Generally speaking, facilitating the exercise of data subjects’ rights such as the right of access or the provision of information should not represent a serious issue to organisations. However, the right to be forgotten has been largely discussed by experts on privacy laws and professionals of the technology sector as it seems to be incompatible with the nature of the Blockchain technology.

It seems that finding a pragmatic approach in which the risks to individuals and organisations are valued and minimised requires working towards the minimisation principle and the idea of
storing the less possible amount of personal data on the Blockchain as well as considering techniques in which the data is not identifiable or accessible.

The CNIL has also considered this issue and provided some useful recommendations implying that when an organisation uses data recorded that is a commitment which fits a purpose in an anonymised form there should not be a need of requesting such data to be deleted or if so, this is likely to be considered as a reasonable approach that gets close to data being deleted from a system.

Regarding the right to be informed, and in line with the approach I suggested above for contracts, perhaps a solution might be to include an on-chain block with privacy information that needs to be verified by all participants as an acknowledgement of having read it. The privacy information block should include as a minimum information which provision is mandatory, including how data subjects can exercise their rights, and who is responsible for dealing with such a request should the occasion arises.

9. Conclusions

Compliance with the data protection laws is a crucial element for Blockchain technology both from a regulatory point of view and a business perspective in order to gain the trust of participants and encourage its use.

Regulators and privacy practitioners are aware of the issues concerning the addressing of all obligations and requirements when it comes to apply them on certain technology such as Blockchain, which means that we expect to see additional developments, guidance, certification schemes and case law (some of which have been expressly demanded by the European Parliament in July this year, in its report on "Blockchain and the General Data Protection Regulation: Can distributed ledgers be squared with European data protection law?"[9].

It is therefore relevant to observe the GDPR principles and considering full compliance with GDPR and the data protection laws should they apply to an organisation. To do so, as I have detailed above, I recommend starting by taking the following actions:

1- Focus on the data protection principles and the big picture:
   a. The purpose of using Blockchain;
   b. The personal data that is used and whether it is necessary to achieve the purpose;
   c. The manner in which data is used and validated; and
   d. Identify the persons who take decisions on all the above.

2- If possible, do not store personal data on the Blockchain and consider full use of data obfuscation, encryption and aggregation techniques as recommended in The Blockchain Report.

3- Seek for collecting personal data off-chain or otherwise a private permissioned Blockchain network. Otherwise carefully consider the options in point 2 above and the risks associated with the use of personal identifiable data on the Blockchain.

4- Keep records of the analysis and decisions taken.

5- Finally, explore ways to inform users and if possible provide the information on the Blockchain and seek for a high consensus mechanisms to ensure the provision of information is validated for as much users as possible.
References


