MODULE

Personal Data Handling

Overview

Focus Areas

1. Applicability of the GDPR
2. Meeting GDPR obligations
3. Technical approaches to GDPR compliance
4. Principles for building GDPR-compliant solutions

Tools and Resources

5. Rapid initial analysis for GDPR compliance

Overview

When it went into effect in 2018, the European Union's General Data Protection Regulation (GDPR) offered ground-breaking protections for personal data. But it also raised questions for blockchain networks as to their new compliance obligations both within and beyond the EU.

The GDPR places obligations on what/whom it terms data controllers and processors. However, when there is no centralised service provider as on a blockchain network, who is responsible for overseeing the treatment of personal data, or paying penalties when obligations are breached? And if a chain is recording data immutably, what does that mean for erasure obligations if that data cannot be taken down? While such considerations need not be prohibitive to beginning a new blockchain project, they should be addressed early on – even, in some circumstances, by supply-chain organisations not based in the EU.

Recommended reading - Inclusive Deployment of Blockchain for Supply Chains: Part 4 – Protecting Your Data.[96]

FOCUS AREAS

1. Applicability of the GDPR

Is there a general understanding that personal data regulation will apply to your blockchain solution? What factors do organisations need to consider when determining the applicability of data protection and privacy obligations?

Personal data protection compliance requirements can dissuade the deployment of blockchain in supply chains if not properly understood, in part because the cost of non-compliance is so high. In addition, such regulations are seldom made with blockchain in mind and thus do not consider the particular nuances associated with blockchain.

While it is not possible to present a comprehensive treatment of all personal data protection regulations that might apply to an international blockchain solution, this module focuses on the European Union's General Data Protection Regulation (GDPR) as a proxy for all data protection and privacy obligations due to its broad scope and detail. The GDPR is also at the forefront of a new wave of data protection legislation globally which places strict obligations on organisations handling personal data or personally identifiable information, such as the recent California Consumer Privacy Act (CCPA).

Many non-EU entities mistakenly believe that the GDPR is not applicable to their solutions. It is important for non-EU-based supply-chain partners to consider that they may still find themselves being subject to GDPR requirements in specific circumstances.

In light of the substantial fines levied for non-compliance and the GDPR's requirement to consider data protection implications from the ground up – a concept known as "privacy by design" – review of this module and professional legal advice is strongly recommended. In addition, professional legal advice is recommended to determine if any other country-specific data protection legislation should be taken into consideration, including legislation surrounding international data transfers.

Using the GDPR as a proxy for data protection and privacy regulations, we can assume that the GDPR may apply to a blockchain solution or transaction. There is nothing about a blockchain context that exempts such things from data protection regulation.

Whether the GDPR applies will rest on the answers to two main questions:

- **Personal Data under the GDPR: Does the data in the supply chain meet the definition of "personal data" under the GDPR?**

  Personal data may include the name, identification number, location data, online identifier, or other information relating to a person. There are therefore various data points within a supply chain that could reasonably be considered to meet the definition of personal data. The GDPR's definition of personal data can also include pseudonymised data if such data can be indirectly associated with a person whether by cross-referencing with other datasets or by other means.

  For the purposes of the GDPR, transactional data stored in the blocks and public keys may be deemed to meet the definition of personal data, although again this list is non-exhaustive. Note that special categories of personal data such as data that would reveal a data subject's racial origin, religious beliefs, or sexual orientation, are defined as special category data and thus are subject to even greater protections under the GDPR.

- **Territorial Scope of the GDPR: Does the processing of such data by the blockchain solution in question fall within the territorial scope of the GDPR?**

  The question of territorial scope requires a consideration of (1) whether the controllers or processors are established within the EU (the "establishment test"); or (2) if the controller or processor is established outside the EU, whether it: (a) offers goods and services to data subjects (people) within the EU (the "targeting test"); or (b) monitors the behaviour of data subjects in the EU, where that behaviour occurs in the EU (the "monitoring test").

  The GDPR applies only when both the personal data definition and territorial scope conditions are met. If the GDPR applies, then all collection, storage and processing of personal data must be done in accordance with the GDPR's requirements, and this includes that data on the blockchain.

The GDPR applies only when both the personal data definition and territorial scope conditions are met. If the GDPR applies, then all collection, storage and processing of personal data must be done in accordance with the GDPR's requirements, and this includes that data on the blockchain.

See the World Economic Forum's white paper Inclusive Deployment of Blockchain for Supply Chains: Part 4 – Protecting Your Data[97] for a detailed description of these two issues. The correct application of these tests depends on the specific facts of the case, and professional legal advice is recommended beyond the initial assessment exercise. In addition, professional legal advice is recommended to determine if any other country-specific data protection and privacy legislation, as well as the rules governing international data transfers, should be taken into consideration.

In this toolkit, the module Digital Identity offers insights on various considerations for Personally Identifiable Information (PII) and a deeper understanding of digital identity.

2. Meeting GDPR obligations

Can a blockchain solution be GDPR-compliant given its characteristics of immutability and distributed nature?

The European Union Blockchain Observatory and Forum stated in its thematic report on Blockchain and the GDPR: "There is no such thing as a GDPR-compliant blockchain technology. There are only GDPR-compliant use cases and applications."

When the GDPR applies, obligations regarding the handling of personal data will apply to processing operations.

Achieving GDPR compliance may require a detailed legal and technical analysis. The following are important steps to take in order to approach compliance under the GDPR for a blockchain solution:

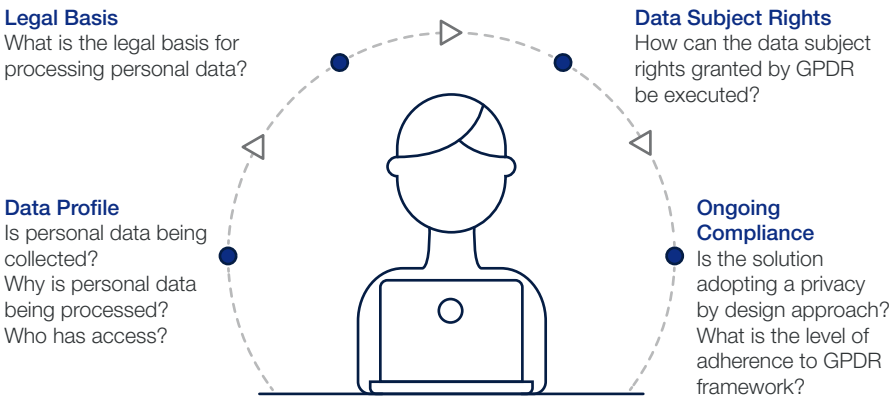

Figure 10.1 – Important steps to take in order to build a GDPR-compliant solution

- **Understand the organisation's data profile.** Engage in a detailed fact-finding exercise in relation to the organisation's data profile. This includes understanding the roles and responsibilities regarding data processing activities in order determine whether you are a data controller, joint controller, or processor. It also includes understanding what personal data is collected, who it relates to, how it is processed and for what reasons, where it is processed, to whom it is provided, who has access to it, and how long it is retained. Further guidance on these and other considerations is available from EU Member States' national data protection authorities' websites.

- **Consider the legal basis for processing.** Determining the legal basis for processing personal data on the blockchain. It is a core principle of the GDPR that lawful data processing can only take place on one of six lawful grounds – for instance, as part of a contract, legal obligation, or legitimate interest, amongst others. This is not a straightforward consideration for blockchain due to the nature of blockchain, such as the fact that, in many cases, there will not be a straightforward relationship between a given blockchain actor and the party whose data is being processed. The obligations around data controllers and data processors should be taken into account here also.

- **Consider how to uphold data subject rights.** Consider how data subject rights granted by the GDPR - including data access, correction, and erasure - can be satisfied when a data subject makes such a request. Depending on the blockchain solution used – for example, a public permissionless or private permissioned blockchain. or any other blockchain type. Whilst accessing personal data on a blockchain should be technically feasible, it may prove challenging to rectify or delete personal data on a blockchain as blockchains are generally designed to offer immutability. It is important to note that these rights can be exercised at any time by a data subject, and it is therefore important to put in place systems that allow the controller of the personal data to comply with the request within one month of the request. In this regard, consideration of accessibility at the point of engineering the data stack can make such requests easier to comply with later on if and when they occur. This is otherwise known as a "privacy by design" approach to engineering the data stack.

- **Consider the ongoing compliance process.** GDPR compliance is an ongoing process rather than a one-off exercise. Ongoing activities may include:

Example

The UK Information Commissioner's Office and the Irish Data Protection Commission provide some helpful English-language resources, which are designed to be accessible to non-specialists such as citizens and small business owners.

- Applying a “privacy by design” approach whereby data protection is made an essential part of the core functionality of the blockchain solution, and data protection requirements are considered throughout the design and implementation process
- Creating a GDPR framework and internal compliance program that includes essential policies covering data security, data breaches, and notification obligations, privacy by design, third-party vendor management, and data subject rights
- Implementing GDPR-compliant transparency notices including a website privacy policy, employee privacy policy, etc.
- Ongoing employee training to ensure that the GDPR framework and compliance program are appropriately operationalised and implemented
- Ongoing maintenance of the organisation’s security obligations
- Ongoing maintenance of processes for and adherence to data-subject rights obligations
- Ongoing review of guidance issued by the European Data Protection Board (EDPB), an independent European body, which works to ensure the consistent application of the GDPR among European Union member states, and EU member state regulators such as the UK Information Commissioner’s Office or the Irish Data Protection Commission. This function will realistically be outsourced to a legal advisor
- Good data hygiene and mapping, including at the data stack engineering level, to ensure data-subject rights obligations can be met with ease
- Incentivising a privacy-first company culture
- Staying abreast of changes in data protection laws that will affect legal obligations
- Familiarisation with and legal advice regarding the GDPR’s obligations around data transfers outside the EU

There is no single blockchain solution or set of solutions to solve the issues described above.

The most straightforward GDPR-compliant solution will always be to exclude the storage and processing of personal data on the blockchain at the outset, but where this is not possible or practical, careful consideration will need to be given to the requirements summarised above, the technical practicalities of the blockchain solution, and the specific performance factors that the blockchain solution aims to optimise. If off-chain data storage is chosen as an appropriate method of reducing risk on the blockchain itself so as to take the blockchain out of scope of the GDPR, then the GDPR obligations will still apply to the data stored off-chain if the two conditions of definition of personal data and territorial scope are met.

The following describes technology approaches which may be used as a starting point to achieve GDPR compliance.

3. Technical approaches to GDPR compliance

What features can be built into a blockchain solution to make GDPR compliance possible?

The data protection authorities of the EU Member States, the European Data Protection Board (EDPB), and the respective courts, including the European Court of Justice (CJEU) have not yet concluded which technology approaches ensure a GDPR-compliant blockchain solution. However, the following represent some technologies and common strategies that may be used to meet certain GDPR requirements when processing personal data while maintaining the desired functionality of the blockchain.

These possible solutions have been recognised by the European Union Blockchain Observatory and Forum,[98] the European Parliamentary Research Service's Panel for the Future of Science and Technology,[99] as well as the French Data Protection Authority (CNIL).[100] However, it is important to note that of the above three entities, only CNIL has legal authority to determine with a relative degree of certainty the legality of a solution, and its ruling could in theory be overturned by the EDPB and CJEU at a future point in time. Legal advice on a case-by-case basis is still recommended when in doubt.

On-chain/off-chain configurations and hashing	Role-based access controls (RBAC)	Zero-knowledge proof (ZKP)	Homomorphic encryption
Basic protections, such as on-chain/off- chain configurations, and only storing hashed data on the blockchain	Enable selective obfuscation of data depending upon the identity of a particular participant	Allows users to prove their knowledge of a value without revealing the value itself	An approach in which data is encrypted before being shared on-chain. It can then be analysed without decryption

Figure 10.2 – Major design options for data confidentiality in a blockchain solution

For a broader overview of some of the current technologies that establish data protection on a blockchain, see the module Data Protection.

The following is an explanation of each technology and its relevance to GDPR compliance.

On-chain/off-chain, obfuscating personal data and GDPR compliance

Under the GDPR, personal data must only be kept for as long as is necessary to achieve the aims for which it was collected. Being unable to delete or effectively delete such personal data from a blockchain, due to the principle of immutability, could constitute a breach of the GDPR because the "data controller" would be unable to protect the data subject's right to erasure.

Where personal data is being processed for a supply-chain solution, a potential approach would be to store only a hash of the relevant personal data on the blockchain, instead of storing the personal data itself on the blockchain. The personal data could be stored off-chain.

The hashing or obfuscation of personal data can increase the control and security of the original personal data maintained by the data controller and

allow the data controller to continue to protect and fulfill data subjects' rights requests. For example, if a data subject were to request erasure of their personal data, the data controller would be able to make such deletion by deleting the personal data stored off-chain, leaving what would then become a meaningless hash on-chain.

This solution also presumes that (a) the hash makes the original personal data inaccessible – e.g. the hash cannot be somehow processed to reveal the original personal data; (b) deleting the original personal data is enough to render the hash meaningless even when combined with any other information – e.g. any relevant keys, information stored on other parts of the distributed ledger; and (c) that regulators are satisfied by this strategy.

This approach could permit meaningful analysis to be conducted (on usage patterns, for example), while having the potential to achieve GDPR compliance if applied comprehensively across the entire blockchain – and the presumptions in (a), (b) and (c) are true.

Role-based access controls and GDPR compliance

Data integrity problems, especially non-adversarial ones, are not new to the supply-chain world. Thus, the solutions relevant to preventing benign data-integrity faults in a blockchain context don't differ much from the solutions applied to data-integrity concerns in a more traditional supply-chain context.

It is relatively straightforward to prevent benign faults, since doing so doesn't require anticipating the potential actions of intelligent and resourceful attackers. The same traditional principles and techniques apply – employing From a GDPR perspective, role-based access controls (RBAC) may be an effective tool to address compliance challenges. Ultimately, GDPR obligations largely apply to the treatment of the personal data regardless of whether the data handler is defined as a controller (who determines the purposes and means of processing personal data) or processor (who processes personal data on behalf of and at the instruction of a controller).

In the GDPR framework, the controller in a blockchain is anyone who determines the purposes and means of processing personal data by writing or adding personal data to the blockchain in a professional or commercial capacity, and a processor is anyone who processes that personal data on the controller's behalf or at the controller's instruction.

Anyone who accesses the blockchain to read the personal data needs to ensure that they have a lawful basis on which to process the data they access. In a blockchain, this lawful basis is unlikely to exist without some sort of contractual relationship between the blockchain actors.

An open-access blockchain solution without role-based access controls or other access restrictions (such as a public blockchain) does not fit neatly into this framework, primarily because it is difficult to identify any one entity who can be held liable, or be compelled to uphold data protection regulations that would apply to the personal data on that public blockchain solution. While it may be clear who the controller is before the personal data is uploaded to the blockchain and in a traditional client-provider model, it is less clear what happens after upload. On upload, the personal data will be collectively processed via a shared protocol leaving no way for the controller to ensure that any blockchain recipient maintained GDPR requirements.

Because of this shared protocol if the blockchain participant is a processor, the problem in such cases is to determine how the original controller can compel the processor to meet GDPR requirements (such as enforcing data-subject rights) and abide by the original terms of disclosure from the data subject to the controller.

If the blockchain participant is a new controller, they have no defined relationship with the data subject, so it is unclear on what legal basis the new controller would process that personal data. It is clear, then, that the disclosure of personal data to the blockchain would present significant GDPR compliance problems for the original controller and create an enforcement nightmare for the data subject.

Note that legal uncertainty remains around the status of actors who act as validating nodes in public permissionless blockchains as the GDPR was not designed with this particular scenario in mind.

Potential approaches. It may be possible to address this problem architecturally by designing a private permissioned blockchain solution, whereby all participants must agree to abide by certain GDPR-compliant terms as a condition to being granted permission: e.g. permitted uses, rules on retention periods, deletion, security and data export to foreign jurisdictions.

Moreover, as CNIL has recommended, the private consortia of the permissioned blockchain networks should also identify the controller, or joint controllers as soon as possible.[101] No public or unauthorised access to the blockchain data would be permitted, or such access should be considered carefully. However, it is unlikely that this would address concerns regarding how liability for errors would be apportioned, if at all.

Data subjects maintain a right to have inaccurate personal data rectified, and so while blockchain maintains good security over data tampering, once that data has been added, any solution will still have to deal with the problem of faulty or fraudulent data being added in the first place.

A possible countermeasure may be to cut off a "bad" participant who consistently shares faulty or fraudulent data. However, where such errors have financial consequences, the matter of enforcement among the blockchain participants – for example, who can bring a claim, what would be the quantum of such a claim, and how will liability be apportioned - becomes important and should be considered at the blockchain-solution design stage.

Zero-knowledge proof

Zero-knowledge proof allows one party to assert the validity of a statement without revealing the underlying facts that make the statement true or false. The algorithm that accomplishes this runs a statement through a true/false test repeatedly until the probability that that statement is false becomes incredibly low. At this point, one is able to confidently assert that the statement is true. One of the key advancements in zero-knowledge proof is s zero-knowledge succinct non-interactive argument of knowledge (zk-SNARK). This technology significantly reduces the time it takes a zero-knowledge proof algorithm to return a result and is one of the most powerful and promising features of blockchain. The obfuscation of the underlying data, including any personal information, may render the obfuscated personal data fully anonymised.

Fully homomorphic encryption

Fully homomorphic encryption (FHE) is a way by which mathematical calculations can be performed off-chain on encrypted data and return an encrypted result. However, the latency of a system that uses this method of computation is even slower than one that uses zero-knowledge proof. While supply-chain partners would be able to run data analytics on artificial-intelligence algorithms on fully encrypted data, and only those who should have access to the result would be provided with a key to decrypt it –and therefore the data may be fully anonymised. The slowness of FHE, however, means that it may not generally beworth the undertaking unless a supply chain

> *With an emerging technology like blockchain, the readiness or maturity of the technology is important to note when designing a solution. Regarding data protection in particular, fully homomorphic encryption (FHE) might seem ideal from a technical perspective for securing information. But from a practical standpoint, it may be slow and thus should be applied only to a limited type of data processing. Given the current state of FHE, it is more realistic to use ordinary encryption or an off-chain database as the use cases likely to benefit from FHE would be limited.*
>
> **Takayuki Suzuki,** Financial Information Systems Sales Management Division, Hitachi, Japan

has a computation for which it does not need real-time, or close to real-time, transmission.

Other methods to approaching GDPR compliance

Storing personal data off-chain with an on-chain hash and adopting role-based access controls are two of the most commonly used approaches to strive for in order to achieve GDPR compliance in a blockchain deployment. However, it is important to note that other approaches also exist. Although uncommon, editable blockchains permit data-subject rights to be respected by allowing a private permissioned blockchain administrator to delete and edit incorrect or outdated information; the trade-off is that it also sacrifices the immutable nature of blockchain. Other solutions allow for deletion by encryption, whereby a blockchain administrator makes certain data inaccessible by increasing the permission needed to access a pre-existing block on the blockchain. It is currently unclear whether this solution would be considered GDPR-compliant by data protection regulators.

4. Principles for building GDPR-compliant solutions

What are four key principles for GDPR-compliant blockchain solutions to follow?

In 2019, a report by several supply-chain industry groups and law firms identified four guiding principles for GDPR-compliant blockchain solutions.[102]

The principles as outlined in the report are:

1. **Use a private, permissioned blockchain when you collect and process personal data.** While the most common vision of blockchain is of a fully public, permissionless network, there are many private networks that are in fact private and require permission to join. Because anyone can join a public permissionless blockchain, it is impossible to ensure participants agree to necessary rules around the protection of personal data. As a result, the private, permissioned blockchains can be employed to work towards a GDPR-compliant blockchain solution. Additionally, proper mechanisms should be put in place when connecting private and public blockchains.
2. **Avoid, if possible, the storing of personal data on the blockchain.** The most obvious way to avoid GDPR compliance issues is to use a blockchain solution that does not process any personal data and minimises free-form data storage. While keeping a blockchain completely free of personal data will be very difficult to achieve, this may be done by deploying advanced cryptographic techniques such as data obfuscations, hashing and aggregation. For example, a blockchain solution can store a hashed representation of the personal data on the blockchain, with the underlying and identifiable personal data kept off-chain. Middleware can then be used to combine off-chain and on-chain data to provide a complete view that includes the off-chain personal data for authorised users only.
3. **Establish a detailed governance framework.** Given: (a) the need to adequately protect personal data; (b) the requirement to establish

contractual relationships governing the processing of personal data between parties; and (c) the legal obligations on data controllers to provide individuals with the means to uphold their personal data rights, a GDPR-compliant commercial blockchain solution will require a governance framework and data lifecycle management that is contractually binding on all participants and clearly sets out each party's rights and responsibilities.

4. **Employ innovative solutions to data protection problems.** The immutable nature of blockchain data is the one element of the technology which clashes most obviously with data subjects' rights under the GDPR. However, through use of innovative solutions such as advanced irreversible encryption as a means of deletion, it may be possible to comply with the spirit and the policy of the legislation, if not yet fully the word. While there are good arguments for irreversible encryption being adequate for GDPR compliance, definitive guidance from regulatory authorities is necessary in this area. One of the key challenges faced by regulators in this space is balancing legislation and technological advancements as, without doubt, technology is moving at a pace which lawmakers struggle to keep up with.

TOOLS AND RESOURCES

5. Rapid initial analysis for GDPR compliance

Because of the great variety of blockchain solutions and configurations, each needs to be analysed on its own distinct merits for GDPR compliance. The following decision tree (see Figure 10.3) provides a simplified summary of common approaches to approaching GDPR compliance in a blockchain context.

This tree is not intended to provide a final authoritative answer, but to assist with a simplified overview of the choices needed to be made during the design and development of a blockchain solution. A solution which is less likely to be GDPR-compliant requires further evaluation and a data protection impact assessment.

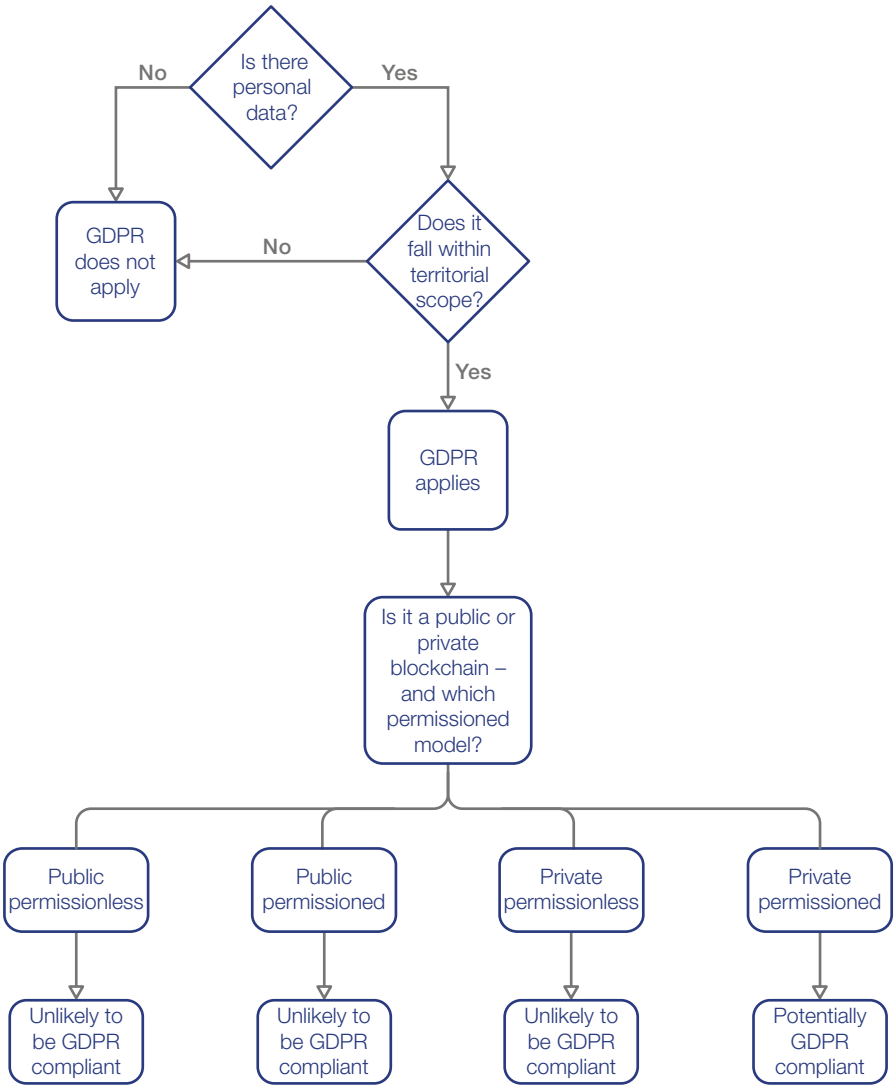

Figure 10.3 – Decision tree with a simplified summary of common approaches for GDPR compliance in a blockchain context. A solution unlikely to be GDPR compliant requires further evaluation and a data protection impact assessment

In conclusion, whilst there is inherent tension between a technology-neutral data protection law such as the GDPR and a specific technology such as blockchain, it is not impossible to become compliant. Legal advice should be sought to ensure that proposed solutions are compliant on a case-by-case basis. Demonstrating compliance where a prescription is not obvious requires a willingness to adhere to both the law (where clear) and the spirit of the law (where the letter of it is unclear) by conducting a data protection impact assessment to satisfy regulatory authorities.