A primer to GDPR, blockchain, and the Seven Foundational Principles of Privacy by Design.

Author:
Silvan Jongerius (silvan@techgdpr.com), TechGDPR (techgdpr.com)

Reviewers:
Greg McMullen, COALA (coalaip.org)
Abigail Garner, TechGDPR

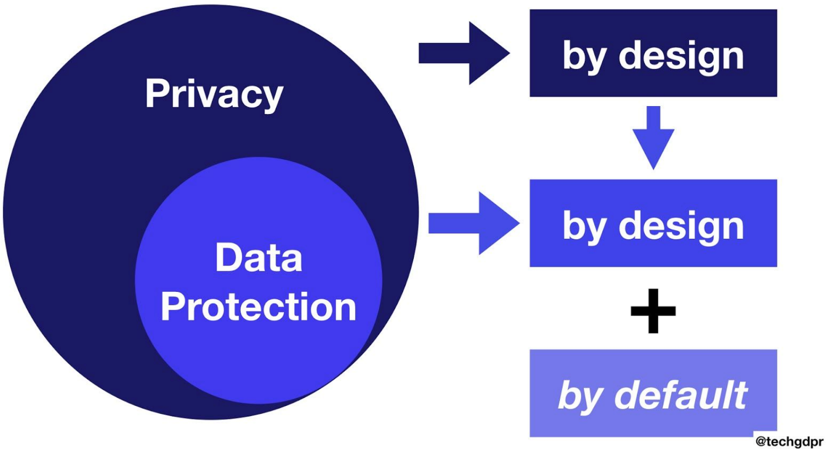
Version 1.2, December 4th, 2018
This is a working document; the author welcomes your feedback.

Throughout my ongoing meetings and conferences about the GDPR and how it relates to cutting-edge technologies, I encounter many ‘add-on’ privacy solutions such as ‘secure your Office 365’, ‘breach detection’ and ‘fix your network.’ Such options, which aim to patch privacy vulnerabilities as they arise, also have the secondary effect of demonstrating the importance of implementing privacy from the start. While some of these add-on options are essential in business environments, it is almost impossible to ensure all vulnerabilities have been repaired. But when privacy is by default and implemented from the start, it is much more likely to be effective. If we are talking about blockchain, the stakes are even higher. With blockchain, privacy by design is the only option. Since blockchains are immutable, there is no way to ‘fix’ things once data is out in the open, and the source code is public.

With the introduction of the GDPR, the concept of *data protection by design and by default* has been signed into law. It is no longer just great advice as outlined by Dr. Ann Cavoukian in the mid-1990s in Privacy by Design: The Seven Foundational Principles. Now these principles are also fully enforceable, and non-application may result in fines of the second highest level: up to €10 million or when greater, 2% of worldwide turnover. Dr. Cavoukian, the former Privacy Commissioner of Ontario, Canada, continues to be to be a champion for privacy by design.

While we speak in general terms about privacy by design, the GDPR deals with the specifics of *data protection*, which I approach as a subset of privacy. Article 25(1) details the specifics of Data Protection by Design and Article 25(2) discusses Data Protection by Default.

The European Data Protection Supervisor has published Opinion 5/2018 on Privacy by Design, one of the first opinions published under the GDPR—again emphasising the importance of implementing privacy from the very beginning.



From a high-level perspective, blockchain effectively *enforces* parts of the Privacy by Design framework, and practically leaves it as the only option. Due to the immutable nature of blockchain, there is simply no other way than to apply it from the very beginning. Protocol or dApp developers could face liability issues if this is not implemented correctly from the start.

The Seven Foundational Principles of Privacy by Design are:

1. *Proactive not Reactive*; *Preventative not Remedial*
2. *Privacy as the Default Setting*
3. *Privacy Embedded into Design*
4. *Full Functionality—Positive-Sum, not Zero-Sum*
5. *End-to-End Security—Full Lifecycle Protection*
6. *Visibility and Transparency—Keep it Open*
7. *Respect for User Privacy—Keep it User-Centric*

Analyzing each of these principles within the context of blockchain offers insight into the greatest challenges for blockchain in privacy by default and privacy by design.

1. Proactive not Reactive; Preventative not Remedial

The GDPR defines “proactive privacy” in this way: *“the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures”.*

The basics are easy – and really the only option for blockchain privacy. If you work with blockchain/DLT you can only be proactive, because if you are not, and personal data make it onto an immutable ledger, you can no longer comply with subjects’ rights.

But you will have to think this through to the end, and that can be challenging. If usage patterns or interaction can be linked back to a natural person, you should already be careful about having this information out in the open.

While encryption can bring a lot of advantages, you should still live under the assumption that it can be broken one day – and in the case of public ledgers, you will end up facing a problem.

2. Privacy as the Default Setting

The GDPR defines privacy by default as *"implement[ing] appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed".*

Users often don't understand the decisions they are being asked to make regarding their privacy. So it is important that when no action is taken by the user, the default option is the most privacy-friendly option, which collects only data needed for the legitimate purpose that has been explicitly defined. (Article 5(1)(b), purpose limitation).

For example, a private transaction within a cryptocurrency, that can be revealed if the parties so desire, is a much better option *from a privacy point of view*, than one that is public but pseudonymised through a wallet address. Perhaps usage patterns, either now or in the future, may reveal (more) personal data than originally intended or foreseen.

While encryption is helpful, it should, as the GDPR and the guidance suggests, be seen as a technical measure to protect the data, and not as a way to anonymize it. Encryption algorithms may be broken in the future, at which point all encrypted data on a public ledger is indeed, public. Major data breach!

3. Privacy Embedded into Design

The GDPR describes this in Article 25(1) in the following way: "the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures".

Privacy should be considered at the design stage, and be applied consistently throughout the product or service. Small flaws or overseen points can have oversized bad outcomes, particularly in a blockchain environment. One should also look beyond the protocol or dApp itself: how do the other system layers impact users' privacy? Is additional information collected on another level, and can this be related back to a natural person or lead to singling out? To do this properly, all system layers should be evaluated.

4. Full Functionality — Positive-Sum, not Zero-Sum

The GDPR does not have specific requirements for this.

In blockchain, privacy and functionality can co-exist. Most implementations dealing with personal data will only store the verification on the blockchain, and keep other parts of information elsewhere. This way, there is no need for a trade-off between privacy and functionality. Innovation can still happen in a privacy-preserving regulatory environment, and the GDPR is no exception. Blockchain offers great possibilities for giving the data subjects full, uncompromised control without the need for personal data to live in central containers that are not only in the control of a single party, but are also vulnerable as high-reward attack targets.

Decentralized, blockchain-based self-sovereign identity solutions are a great example of increased privacy with a user experience that meets or exceeds current identity services. By applying privacy by design, the paradigm of trade-offs such as: “if we want to live in a more secure world, we need to compromise our privacy” is simply not true, and blockchain can actually be a great help in achieving both privacy and security.

5. End-to-End Security — Full Lifecycle Protection

The GDPR defines this in Art 17 – right to erasure, and Art 5(1)(f) – integrity and confidentiality.

When using blockchain, it is easy to deliver on the integrity of the data. The concept of blockchain is immutability, and if therefore tamper-proof, ensuring that data can be deleted only by having a clearly breakable link to any personal data. This means no information on the blockchain should be able to be re-linked to any personal data once the link is purposefully broken. To achieve this, the current best practice is storing only hash-proofs on the blockchain that include a secret piece of information that can be permanently deleted, or even better, by using zero-knowledge proofs.

It should be emphasized that even a hash, if it can be re-constructed, could be considered personal data *under certain circumstances*, following the same reasoning as the judgment of the CJEU in Case 582/14 – Patrick Breyer v. Germany.[1] This is not always the case, but in most cases, this will not be known with certainty and should therefore be assumed to be personal data if any doubt exists.

6. Visibility and Transparency — Keep it Open

The GDPR phrases this principle as “the controller should adopt internal policies and implement measures which meet in particular the principles of data protection by design and data protection by default. Such measures could consist, inter alia, of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features” in Recital 78, and in Article 5(1)(a) (“lawfulness, fairness and transparency”).

Blockchain is transparent by nature, so this principle should be easy to uphold. The recommendation of Privacy by Design is to verify that the business practices or technology involved are, in fact, operating according to the promises and objectives they have defined. Within public blockchain environments, not only the code, but also the data generated, is public and can be verified.

At times, blockchain could be a little too transparent; every participant in public networks, can see everything. While this makes it a great feature for certain purposes, you also want to make sure not to store any personal data on a blockchain (a nuance here is in place, as there are ways of storing pseudonymized data, but that’s beyond the scope of this article).

[1] http://curia.europa.eu/juris/document/document.jsf?text=&docid=184668&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=1116945

7. Respect for User Privacy — Keep it User-Centric

In Recital 7, the GDPR states that, "Natural persons should have control of their own personal data."

Operators and architects of blockchains should keep the data subject and their privacy in mind in all stages of designing and operating the system. This is particularly important at the start of any blockchain protocol design as what has been stored on a blockchain is immutable, and cannot be deleted or reversed in the future. Ideally, it should be made difficult or actively discouraged to store any sort of personal data on an 'immutable ledger', not only for privacy reasons but also because software developers could risk liability as certain regulators are of the opinion that they could be seen as Data Processors under the GDPR.[2]

On the other hand, blockchain technology also has the ability to give users control over their personal data through self-sovereign identity systems. The user has control over where the data is stored, on their own device or elsewhere. This is protected against falsification by storing the third-party proofs (and no actual personal data) on the blockchain. This allows users to reveal their identity or only parts or proofs thereof to other parties of their choice.

Conclusion

Privacy by design includes a set of important principles now deeply embedded into the GDPR. It is no longer optional to consider and implement the principles and related requirements. Especially with blockchain, there is no alternative to implementing privacy by design from the start, as the usual add-on privacy enhancements simply will not satisfy the requirements of the GDPR.

This short paper is to be seen as a primer of potential further work on the subject of how to properly practice privacy by design using blockchain technology. There are a variety of examples and ideas available, most of them initiated or developed by Dr. Ann Cavoukian, to be evaluated with the constraints and opportunities of blockchain. Future work could also include evaluations of how specific blockchain projects implement or practice good privacy by design.

[2] Opinion of the French data protection authority CNIL in September 2018 (French): https://www.cnil.fr/sites/default/files/atoms/files/la_blockchain.pdf