Executive Summary: GDPR Compliance Challenges in Blockchain by Jørgen Leiros

Jørgen Maurstad Leiros, as part of his MSc in Business Administration and Data Science, conducted a **regulatory and technical analysis** on the compatibility of **blockchain technology with GDPR (General Data Protection Regulation)**. His research explores the legal ambiguities surrounding **data processing, user rights, and compliance risks**, proposing potential **blockchain design solutions** to align with European data protection laws.

Key Focus Areas:

1. **When Does GDPR Apply to Blockchain?** –
    - GDPR applies to **organizations processing personal data via blockchain**, including **entities within the EU** and **global firms handling EU user data**.
    - **Challenges arise in defining "personal data"**, as blockchain transactions use **pseudonymous identifiers**, which may still be traceable.
    - **Public blockchains (Ethereum, Bitcoin) create compliance risks** due to **immutable, decentralized data storage**, complicating **data deletion and user rights enforcement**.
2. **Legal Uncertainty: Who is the Data Controller?** –
    - GDPR mandates **identifying data controllers and processors**, but in **decentralized blockchain networks, accountability is unclear**.
    - **Possible controllers:**
        - **Blockchain nodes and miners** (since they validate and distribute transactions).
        - **Smart contract developers** (as they define how data is processed).
        - **Organizations using permissioned blockchains** (which control access and data usage).
    - **No legal consensus** exists on who holds primary responsibility in **public blockchains**, creating **compliance gaps**.
3. **GDPR User Rights vs. Blockchain Immutability** –
    - **Article 16: Right to Rectification** – Blockchain does not allow altering stored data, conflicting with GDPR's requirement for **correction of inaccurate personal data**.

    - **Article 17: Right to Erasure ("Right to be Forgotten")** – Deleting blockchain records is nearly impossible, raising concerns about **permanent data storage**.
    - **Article 18: Right to Restrict Processing** – In **public blockchains, data is permanently replicated**, making it **impossible to prevent further processing**.
4. **Proposed GDPR-Compliant Blockchain Design:**
    - **Use of Permissioned Blockchains** – Restricting node access ensures **GDPR-compliant data governance**.
    - **Off-Chain Storage & Hashing** – Storing personal data **off-chain** while keeping **hashes on-chain** allows compliance with **erasure and rectification rights**.
    - **Enhanced Smart Contract Design** – Implementing **legal safeguards within smart contracts** to allow **user data control and compliance with GDPR**.

Key Takeaways:

- **Public blockchains face major GDPR compliance issues**, particularly in data deletion, user rights enforcement, and unclear legal responsibility.
- **Permissioned blockchain models and off-chain data storage** offer potential solutions but **require further legal and technical refinement**.
- **Regulatory clarity is needed to align blockchain technology with European data protection laws**, ensuring **legal certainty for businesses and developers**.

Through this research, Jørgen has demonstrated expertise in **data privacy law, blockchain governance, and regulatory technology (RegTech)**, contributing valuable insights into **GDPR-compliant blockchain applications**.

Introduction

The General Data Protection Regulation (GDPR) is a part of the secondary legislation of European Union law, and aims to provide protection for the personal data of natural persons (Directive 95/46/EC, Article 1). This legislation comes as a consequence of an increased amount of data gathering and digitally stored personal information, made possible by technological development. The GDPR has an important function in empowering data subjects, but in doing so, it may impose unintended restrictions that hinder the

implementation and development of exciting new technologies such as artificial intelligence and blockchain technology (Li et al., 2019).

Distributed Ledger Technology (DLT) is a technology that allows for storing and exchanging data on a decentralized network of *nodes* [1] in contrast to centralized servers controlling traffic which is the case with *Web2.0*[2] (G. Srivastava et al., 2019). Information about every transaction on the network is stored on what is popularly called a blockchain as *hash values*[3]. All nodes on the network have a copy of the blockchain and when a new block of transactions is validated, it is added to the blockchain, and each copy is updated. These copies are used to ensure the integrity and transparency of the data that is stored on the blockchain. These inherent qualities make blockchain technology a great alternative in use cases where integrity and transparency is important, or in the absence of a trusted third party, but may cause issues in complying with GDPR.

Ethereum introduced "smart contracts" as a way to launch computer programs on the blockchain in 2016. This expanded the potential applications of blockchain technology significantly. As smart contracts were introduced the same year as the GDPR was drafted, it is not unlikely that the legislators have put less emphasis on this technology than what is warranted with today's increased use cases. We are going to explore this issue by discussing when a blockchain network is within the scope of GDPR, who the responsible parties of the decentralized network are, central problems regarding lawfulness of processing and data subjects' rights, as well as potential ambiguity and its effects. Based on the compatibility issues identified through the discussion, this synopsis will conclude with suggesting a potential solution to blockchain design for GDPR compliance.

Blockchain in the Scope of GDPR

Articles 2 and 3 of the GDPR describe the material and the territorial scope of the regulation. To lie within the material scope means that the organization needs to be "processing personal data wholly or partly by automated means, and to the processing other than by automated

[1] A point of connection within a data communication network (Rouse, 2023)
[2] The widespread adoption of the internet known for corporate digital platforms and advertising-driven revenue (Nabben, 2023)
[3] A string of fixed size, consisting of numbers and letters (Loo, no date)

means of personal data which form part of a filing system or are intended to form part of a filing system."(Directive 95/46/EC, Article 2). Article 4 goes on to define personal data as data relating to a natural person who can be identified directly by that information or indirectly in combination with other information.

Determining whether data on the blockchain should be regarded as personal data can be challenging. One of the most important aspects of blockchain technology is its great dedication to anonymity and pseudonymity for its users, but the issue of complete anonymity still remains (N. Andola et al., 2021). Most blockchains are public, meaning they are available to everyone on the platform, and in principle the data is encrypted so that a real-life user operates under a pseudo identity. This makes it possible for the real-life individual to be almost entirely anonymous. From the perspective of an organization however, this does not apply.

The organization storing the personal data of a user on the blockchain will, in many cases, still need to access all of the user's data for the same purpose as an organization using a standard database. The processing can, however, be done without combining the identifiable personal data of the data subjects, but when combining is required, the organization finds itself within the boundaries of the regulation. The data is encrypted, but the encryption can potentially be reverse engineered in the future, which is why it is not regarded as completely anonymous. Hashing is often used to pseudonymize information on the blockchain, but with ever evolving computing power, reverse engineering of today's hashing algorithms is becoming increasingly plausible (E. Felten, 2012).

The GDPR governs the processing of personal data within the European Union (Directive 95/46/EC, Article 3). Article 3, the territorial scope, shows that it applies to organizations, whether controllers or processors, with establishments in the Union, regardless of where the actual processing occurs. Additionally, GDPR extends its jurisdiction to entities outside the EU processing the personal data of individuals within the Union. This includes situations where goods or services are offered to EU residents, irrespective of payment, or when monitoring their behavior within the EU occurs. Furthermore, GDPR covers the processing of personal data by controllers located outside the Union but subject to Member State law through public international law.

An organization that processes personal data on the blockchain with establishments in the Union will always be subject to the GDPR requirements. The same applies to an organization that is not based in the Union, but still has operations there and relies on blockchain to process personal data (European Parliament, 2019). If neither of these scenarios are relevant, organizations may still need to adhere to the GDPR. This is the case for any organization that offers goods or services to data subjects in the Union, for instance when making their blockchain platform available for the data subjects in the EU, as this can be perceived as offering a service. Additionally, the GDPR applies when an organization outside of the Union processes data with blockchain in the context of monitoring data subjects in the EU.

Transfer of Personal Data

Chapter 5 of the GDPR presents the restrictions of personal data transfers from the EU to third countries. For an organization to be allowed to transfer personal data in these scenarios, it must follow three conditions: (1) if there is an approved adequacy decision given to the country by the Union, (2) there are appropriate safeguards in place in form of standard contractual clauses (SCC), or (3) based on specific derogations (European Parliament, 2019). In the case of an organization using blockchain to process personal data, it is important to take these conditions into account as the distributed ledger facilitates multiple nodes which can be located both within and outside of the Union. Using a public blockchain allows unrestricted access to participants outside the EU, leading to potential clashes with the regulations for the organization. However, having a permissioned blockchain allows the organization to decide the location of the nodes on the network, ensuring that processing nodes are either within the EU or comply with one of the three conditions stated above.

Controller and Processor

In the context of GDPR, we separate the entities responsible for data processing into three roles; data controller, joint controllers and data processors. Article 4(7) defines a data controller as: "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data […] " (Directive 95/46/EC, Article 4(7)).

In the context of using blockchain, the data controller is the entity that decides on the "why" (purpose) and the "how" (means) of processing of personal data, thus bearing the responsibility for compliance with GDPR. If an organization makes use of permissioned blockchains for data processing, it likely becomes the data controller and must ensure that the processing adheres to GDPR principles. This is because the organization is the sole decision-maker in regards to the means and, in most cases, the purposes (European Parliament, 2019). There are, however, conflicting opinions on who should be liable as the data controller, and this is mainly the issue when discussing public blockchains. It has been argued that nodes, miners and software developers could each be classified as data controllers. The miners could for instance qualify as data controllers since they "determine why and how their own local version of the block is processed." (R. Belen-Saglam et al., 2023). Nodes are argued to be controllers because of the fact that while completing a transaction on the network, it distributes this information across the whole platform, thus pursuing its own purpose (European Parliament, 2019). In the case of software developers, it has been argued that the first designers of the blockchain should keep the regulations in mind and create a system that is compliant by design. If this is not taken into account, the first designers should be prepared to be held liable by private parties or authorities (G. Jaccard & A. Tharin, 2018). However, a counter argument would be that the designers are only facilitating the use of the service, and therefore they have limited influence on the means of how the personal data is to be processed (European Parliament, 2019).

In Article 26, the term joint controllers is described as multiple parties jointly deciding on the purposes and means of processing (Directive 95/46/EC, Article 26). They need to agree clearly on who is responsible for what, especially concerning data subjects' rights and information duties referred to in Article 13 and 14, unless Union or Member State law specifies their roles (R. Belen-Saglam et al., 2023). The arrangement needs to be disclosed to the data subjects, and regardless of this agreement, data subjects can exercise their rights under this regulation against any of the joint controllers . In the context of blockchain, it is argued that nodes on the platform are joint controllers. In the case where nodes form a majority number of all mining power, more than 50%, they should qualify as joint controllers. However, in order to meet the requirements of Articles 13 and 14, the distribution of

responsibility between the joint controllers should be clear, and because of the vast amount of nodes in a public blockchain, this becomes an impossible task.

In Article 4(8), a data processor is defined as: "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller" (Directive 95/46/EC, Article 4(8)). The obligations of the data processor are further explained in Article 28 of the GDPR. The data controller shall only use processors who provide sufficient guarantees to implement appropriate measures, ensuring that the processing is following the GDPR requirements while protecting the data subjects' rights. The processor must follow a binding legal agreement, detailing processing specifics and controller rights. This includes assisting the controller in fulfilling GDPR obligations, and processing data only under instructions made by the controller (Directive 95/46/EC, Article 28). When determining the data processors in context of blockchain use, it is important to emphasize that such decisions should be made on a case-by-case basis (European Parliament, 2019). The identification of processors is dependent on the determined data controller, and this classification, as seen above, can be difficult. There are a lot of cases where entities can be classified as data processors, but given the scope of this synopsis, we choose to address a selection. Examples of data processors could include data warehouses of out-sourcing agencies, cloud providers or those providing software, platform or infrastructure as a service (European Parliament, 2019). Additionally, smart contract developers could be classified as data processors considering the fact that they are responsible for the processing of personal data under instructions by the data controller (R. Belen-Saglam et al., 2023).

Lawful Processing of Personal Data

Under the GDPR, all processing of personal data must comply with the six basic principles stated in Article 5 and at least one of the criteria for legitimate basis stated in Article 6 & 9 (Trzaskowski & Sørensen, 2022, p. 79). Although blockchain technology can align with most of these principles, there could be challenges related to *storage limitations*. The principle concerns the duration of the data storage, and that the data cannot be stored longer than necessary. Because the data can’t be erased from the blockchain, the controller might have a problem with ensuring compliance. Maintaining *accuracy* of the data could also be

problematic, since the principle states that inaccurate personal data must be erased or rectified without delay (Trzaskowski & Sørensen, 2022, pp. 83-84).

Following Article 5, Article 6 serves an important role in establishing the lawful basis of processing data, and outlines the conditions under which such processing is legal (Directive 95/46/EC, Article 6). The legitimacy of data processing in blockchain depends on the specific use case, requiring a careful selection of a valid basis that involves understanding the purpose, functions, and data flows to fulfill the intended purpose(s). (Trzaskowski & Sørensen, 2022, p. 87).

The characteristics of blockchain technology introduces unique challenges when applying a legal basis. If consent is applied, the EDPB's (European Data Protection Board) guidelines states that it " [...] can only be an appropriate lawful basis if a data subject is offered control and is offered a genuine choice with regard to accepting or declining them without detriment" (Trzaskowski & Sørensen, 2022, p. 89). In the context of the decentralized and immutable nature of the blockchain, challenges may arise in updating or revoking consent due to the difficulty of altering information once recorded. If this results in the data subject losing control, consent will be an invalid legitimate basis (Trzaskowski & Sørensen, 2022, p. 90). As consent may be an invalid legal basis on the blockchain, legitimate interest and performance of contract seem to be the only viable options for legal basis if processing is not carried out as part of a legal obligation, to protect the data subjects' vital interests, or public interests.

In the immutable blockchain network, personal data, transaction history, and user behavior is stored to maintain the integrity of the platform. Introducing the balancing test, the controller has a (legitimate) interest in ensuring that fraud is prevented and the data is processed in a way that minimizes the impact on individual privacy (Trzaskowski & Sørensen, 2022, p. 102). In case the legal basis is the performance of a contract, personal data is processed as part of the contractual arrangement between parties. To fulfill the obligations of the contract, the personal data (e.g., account details for billing) is necessary. Without this information the blockchain technology is unable to provide the service.

Data Subjects' Rights

If we find a blockchain network to be within the scope of GDPR, articles related to data subject's rights are particularly problematic for the use of blockchain technology. As mentioned in the introduction of this synopsis, blockchain technology has inherent qualities that can make it favorable to traditional databases. These qualities are guaranteed by the protocol of the blockchain network. The protocol determines how blocks are validated through a consensus algorithm (G. Srivastava et al., 2019). It is this algorithm that removes the need for a trusted third party, and allows for a distributed network. The most widely used consensus algorithms differ somewhat in logic, but they are all based on the ability for nodes to compare new proposed copies of the blockchain to their own copy. So a basic principle, and inherent value of blockchain technology, is the immutable nature of the blockchain. This means that once data is stored on the blockchain, it cannot be removed or changed (McKinsey, 2022). In the following paragraphs we will highlight three of the data subjects´rights to illustrate the compatibility issues with blockchain technology.

Article 16: Right to Rectification

The right to rectification is related to the basic principle of accuracy under Article 5(1d) which states that the controller is responsible for keeping personal data up to date, and rectifying or deleting any incorrect data (Directive 95/46/EC, Article 16). As blockchains are immutable, data on the blockchain can't be changed, but new data can be added. So if a natural person's data was stored on a blockchain and it appeared to be wrong, new and accurate data could be added but the inaccurate data could never be changed.

Article 17: Right to Erasure

Article 17 of the GDPR provides the data subject with the right to erasure of their personal data (Directive 95/46/EC, Article 17). This right only applies on certain grounds, but if a data subject exercises their right to erasure on one of the grounds mentioned under Article 17, the controller would have to delete the entire blockchain to comply with GDPR. However, according to Article 17(1)(b), the legitimate interest or legal grounds of the controller may override the data subject's right to erasure.

Article 18: Right to Restriction of Processing

The data subject has a right to restrict the processing of their personal data if it is inaccurate, processing is unlawful, the controller no longer needs it for the initial purpose of processing, or the data subject objects to processing based on their particular situation (Directive 95/46/EC, Article 18). On a public blockchain, which is the most common, anyone can get a copy of the data and there is no way to control what they do with it from that point. This has obvious implications when it comes to GDPR, and directly violates Article 18 as there is no way to restrict processing.

Permissioned blockchains are blockchain networks where only known and trusted nodes can access the data. By using a permissioned blockchain rather than a public blockchain, it would be possible to comply with Article 18. However, permissioned blockchains are still immutable and do not help to comply with Articles 16 and 17.

Ambiguity in Legislation

When legislation is considered ambiguous, it means that there is a lack of clarity in the language or wording of the laws, which makes it unclear or open to multiple interpretations. This might result in individuals, businesses or legal authorities having a difficult time understanding what the law precisely requires or prohibits in certain situations. As technology generally evolves at a much higher pace than legislation, it is natural that legislation is formulated in a general way to increase its scope. This may however lead to ambiguous language, which again leads to confusion and disputes.

One of the most debated issues revolves around the question of who assumes the roles of data controller and processor in public blockchain networks. However, researchers have been unable to reach a consensus on determining the responsible parties for the data. This is because the information in a blockchain is decentralized across a network of nodes, and the roles become blurred. Consequently, some researchers argue that the privacy and data protection laws may not be applicable in this context (R. Belen-Saglam et al., 2023). The issues of how the legal roles are distributed is a big problem in the blockchain network and creates ambiguities in the legislation.

An article that especially creates ambiguity and makes researchers doubt the laws is Article 17. As mentioned, by removing specific data from the chain, it would contradict one of its foundational principles. This obviously creates a challenge when trying to reconcile with the GDPR and Article 17. Compliance with GDPR could be easier when implementing permissioned blockchain, where a central authority oversees the data being stored, in contrast to the public ones. Still, drawing a conclusion on this issue is difficult. This is just one of several examples where traditional data regulation creates tension with innovative technology, which creates ambiguity in legislation.

"As a technologically-neutral legal framework, the GDPR was designed in such a manner as to enable its application to any technology" (European Parliament, 2019). The neutrality towards technology also implies that applying its requirements to particular instances of personal data processing can be challenging. The decentralized and immutable ledger makes the blockchain unique and complex and therefore more difficult for GDPR to be applicable. By looking through the GDPR, one can also tell that there is a lack of examples involving blockchain. This can contribute to legal uncertainty as business and legal experts navigate the continuous updates and innovations of the blockchain space.

Blockchain Design for GDPR Compliance

When a controller is in the scope of GDPR, storing personal data on a blockchain introduces a number of complications when it comes to GDPR compliance: Determining the controller of a blockchain network is close to impossible for public blockchains, and determining the processors is dependent on who the controller is; even though relevant forms of legal basis can be applied, there are concerns surrounding the validity of consent, and performance of a contract, making it necessary for legitimate interest to apply for blockchain to be a viable means of processing; it is not possible to change incorrect data, it is only possible to delete data if you delete the entire blockchain, and the only way to ensure restriction of processing is by using a permissioned blockchain.

To solve the problem of identifying roles and distributing responsibility, the blockchain should be permissioned. This allows the data controllers to select processors and control the distribution of data, ensuring that processors comply with GDPR and that data is not

transferred to any countries outside of the EU without the appropriate conditions being met. When it comes to the immutability of the blockchain, the most discussed solution in the scientific community is *hashing out* personal data (R. Belen-Saglam et al., 2023). This is when personal data is stored on a separate, secure server *off-chain*[4], and a hash of the data is stored *on-chain*[5]. In this way the integrity of the data is guaranteed by the hash on the blockchain and if the personal data that is stored *off-chain* is deleted, the hash stored *on-chain* cannot reveal the data. However, there is still no consensus around a solution that is completely robust, further underlining the difficulties presented in this synopsis and concluding with the need for blockchain to be a part of future development of the GDPR.

References

Andola, N., Raghav, V. K. Yadav, S. Venkatesan, S. Verma, S. (2021). Anonymity on blockchain based e-cash protocols—A survey. *Computer Science Review*, 40, 100394. https://doi.org/10.1016/j.cosrev.2021.100394

Belen-Saglam, R., Altuncu, E., Lu, Y., & Li, S. (2023). A systematic literature review of the tension between the GDPR and public blockchain systems. *Blockchain: Research and Applications*, *4*(2), 100129. https://doi.org/10.1016/j.bcra.2023.100129

Codecademy (2023). What Is Hashing, and How Does It Work? Codecademy. Retrieved November 14th, 2023 from https://www.codecademy.com/resources/blog/what-is-hashing/

European Parliament. (2019). Study on blockchain and GDPR. Retrieved December 1st, 2023, from https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634445/EPRS_STU(2019)634445(ANN1)_EN.pdf

European Parliament. (2019). Blockchain and the General Data Protection Regulation. Retrieved December 1st, from

[4] *Off-chain* refers to data that is stored or transferred outside of a blockchain network.
[5] *On-chain* refers to data that is stored or transferred on a blockchain network.

https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634445/EPRS_STU(2019)634445_EN.pdf

Felten, E. (2012, April 22). Does hashing make data "anonymous"? Federal Trade Commission. https://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2012/04/does-hashing-make-data-anonymous

G. Srivastava, S. Dhar, A. D. Dwivedi & J. Crichigno. (2019). Blockchain Education. 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), 1–5. https://doi.org/10.1109/CCECE.2019.8861828

Jaccard, G., & Tharin, A. (2018). GDPR & Blockchain: The Swiss take. Jusletter IT. Retrieved December 1, 2023, from https://lawded.ch/wp-content/uploads/2019/07/Jusletter-IT_gdpr-blockchain-t_5aebbf8be4_en.pdf

Khan, S. N., Loukil, F., Ghedira-Guegan, C., Benkhelifa, E., & Bani-Hani, A. (2021). Blockchain smart contracts: Applications, challenges, and future trends. *Peer-to-Peer Networking and Applications, 14*(5), 2901–2925. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8053233/

Li, H., Yu, L., & He, W. (2019). The Impact of GDPR on Global Technology Development. *Journal of Global Information Technology Management*, *22*(1), 1–6. https://doi.org/10.1080/1097198X.2019.1569186

Loo, A. (no date). Hash Function. *Corporate Finance Institute.* Retrieved 01. 12. 2023, from: https://corporatefinanceinstitute.com/resources/cryptocurrency/hash-function/

Liu, M., Wu, K., & Xu, J. J. (2019). How Will Blockchain Technology Impact Auditing and Accounting: Permissionless versus Permissioned Blockchain. Current Issues in Auditing, 13(2), A19–A29. https://doi.org/10.2308/ciia-52540

McKinsey. (2022). What is blockchain? McKinsey. Retrieved November 13th, 2023, from https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-blockchain#/

Nabben, K. (2023). Web3 as "self-infrastructuring": The challenge is how. Big Data & Society, 10(1). https://doi.org/10.1177/20539517231159002

Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC

Rouse, M. (2023) What is a node? Techopedia. Retrieved December 6th, 2023 from https://www.techopedia.com/definition/5307/node