АКТУАЛНО
CURRENT TOPICS

NAVIGATING PRIVACY IN CRYPTO: CURRENT CHALLENGES AND (FUTURE) SOLUTIONS

Vyara Savova
University of Library Studies and Information Technologies
https://doi.org/10.70300/NAFD2837

Abstract: *This article examines the evolving relationship between privacy and blockchain technology, examining the challenges and innovations in the context of decentralized finance (DeFi) and smart contracts. It explores the complex interplay between blockchain's transparency and the necessity for user privacy, drawing insights from previous research, including the works of Michèle Finck and Primavera De Filippi, who analyze the impact of GDPR and the dichotomy of decentralization and privacy. The article further delves into recent proposals, notably Vitalik Buterin et al.'s 'Privacy Pools', addressing the balance between privacy and regulatory compliance through the use of cryptography. It concludes by acknowledging ongoing privacy challenges while highlighting promising developments, suggesting a possibility of blockchain technology reaching a harmonious co-habitation with privacy principles.*
Keywords: *Blockchain, Privacy, GDPR, Zero-Knowledge Proofs, Privacy Pools*

INTRODUCTION

Blockchain technology can be defined as a "decentralized ledger (or state machine) that relies on cryptographic algorithms and economic incentives in order to ensure the integrity and legitimacy of every transaction (or state change)". A fundamental aspect of the technology is that a copy of the blockchain is shared amongst all nodes (or computers) connected to the network, comprising the history of all valid transactions. Then, "each transaction is recorded into a 'block', which is appended sequentially to the previous block of transactions" (De Filippi 2017). The technology has proven to be among the groundbreaking innovations of the past few decades due to its potential to revolutionize numerous sectors with its unique technical and social attributes, notably decentralization, transparency, and immutability.

Originally conceptualized for digital currency transactions, its application has since expanded far beyond, permeating various fields such as finance, art, healthcare, and governance. However, this rapid integration brings forth significant privacy concerns. As blockchain networks often operate on public and permissionless ledgers, the balance between transparency and user privacy becomes a complex issue, particularly in the context of stringent data protection regulations such as the EU's General Data Protection Regulation (GDPR 2020). Generally, to better understand the difference between permissioned / permissionless and public/private blockchains, by using the Bitcoin blockchain as an example, see Michèle Finck's article Blockchains and Data Protection in the European Union, where the author explains how "[t]he original Bitcoin blockchain is a public and unpermissioned (or 'permissionless') blockchain, which means that it is open-source and open-access so that anyone can create a Bitcoin address and download or design software to run nodes" (Finck 2018).

Moreover, the rise of decentralized finance (or DeFi, refers to a "set of newly emerging financial products and services that operate on decentralized platforms using blockchains to record and share data". Furthermore, "DeFi products and services are conducted without a trusted central intermediary" such as a bank, and they include various services like payments, lending and borrowing, trading and investments, capital raising (crowdfunding), and insurance, among others (Carapella et al. 2022). and the implementation of smart contracts have further complicated the privacy landscape.

Smart contracts, although not legal contracts but instead self-executing lines of code, already pose unique legal challenges, including data protection and user privacy. While there isn't a unified agreement reached around the definition of a smart contract, for the purposes of this article, we may understand smart contracts as 'automated software agents hosted on blockchains that are capable of autonomously executing transactions on the triggering of certain conditions' (Goldenfein and Leiter 2018). While they offer automation and efficiency, they also raise questions about the control and confidentiality of personal data. An intriguing aspect of the broader discussion around blockchain and privacy is the status of smart contracts in the context of the limitations around automated decision-making and solely automated data processing under the GDPR. See, for example, Michèle Finck's article Smart contracts as a form of solely automated processing under the GDPR, in which the author argues that "[s]mart contracts indeed appear to qualify as a form of solely automated data processing under Article 22(1) GDPR" (Finck 2019).

In addressing these challenges, this article will delve into the intricate relationship between blockchain technology and privacy. Drawing from the works of experts such as Michèle Finck and Primavera De Filippi and considering the latest proposals from Vitalik Buterin et al., we will explore the current state of privacy in blockchain, the ongoing challenges, and the promising solutions on the horizon. The objective is to provide a comprehensive understanding of where blockchain stands in terms of privacy and what the future may hold for this rapidly evolving technology.

OVERVIEW OF PAST DEVELOPMENTS

Even though the introduction of blockchain technology predates that of the GDPR, the fragile relationship between an infrastructure that is by design transparent and the concerns around personal data protection has been ongoing since the very genesis of the technology. The Bitcoin whitepaper, authored by the pseudonymous Satoshi Nakamoto, addresses the issue of privacy on-chain, however, primarily in the context of the anonymity of transactions. It further emphasizes that while the transaction flow is public on the blockchain, the identities of the parties involved in transactions are not directly linked to their public keys. According to the author(s) of Bitcoin's whitepaper, this design offers a level of privacy by keeping users' identities separate from their transaction history, although it is not completely anonymous. The whitepaper discusses methods to increase privacy, such as using new addresses for each transaction. However, it also acknowledges that linking public keys to real-life identities can potentially reduce this privacy (Nakamoto 2008). Furthermore, the reason for this fragile relationship can be traced back to the design of both blockchain and GDPR, as "[w]hereas the GDPR was fashioned for a world where data is centrally collected, stored, and processed, blockchains decentralize each of these processes" (Finck 2018).

Beyond that, the evolution of the concerns around privacy in blockchain technology can further be traced through seminal works of research conducted in the past decade, such as the publications of Michèle Finck and Primavera De Filippi – although far not the only researchers in the fields, those two authors have been chosen to be highlighted here as their seminal research not only lay at the intersection of technical and legal analysis, without focusing disproportionately on one of the two but also covers and analyses a significant number of other publications.

In a nutshell, Michèle Finck's studies examine the intersection of blockchain technology with data protection laws, particularly focusing on the European Union's GDPR. In her work, Finck highlights the complexities of categorizing blockchain within existing legal frameworks, emphasizing the challenges posed by the decentralized and immutable nature of blockchain data. As described above, she also delves into the concept of smart contracts, exploring their designation as a form of solely automated data processing under the GDPR and discussing the implications for user privacy and data protection rights.

Below are listed some of the conclusions of her work that are considered crucial for the current understanding of the conflict between privacy and blockchain and, therefore, also fundamental for finding solutions to this conflict.

According to Finck, "[m]ost DLTs [short for distributed ledger technologies] contain two types of data: (i) the header which includes the timestamp, the identity of the data's source such as an address and the previous block hash, whereas (ii) the block content (or payload) contains the data to be stored (on the

Bitcoin blockchain this would be the relevant transactions as well as the coinbase transaction). Whereas the header is usually not encrypted, the payload normally is" (Finck 2018).

It is important to mention here that according to Opinion 04/2014 on Anonymization Techniques, 0829/14/EN, by Article 29 Working Party (the organization is a precursor of the currently active European Data Protection Board), encryption is considered a pseudonymization technique under EU data protection regime given that the data subject can still be indirectly identified (Article 29 Data Protection Working Party 2014). Therefore, the use of encryption would not preclude the application of GDPR but instead would deem all encrypted personal data as pseudonymized. Nevertheless, considering that the European Data Protection Board has included the drafting of new Guidelines on Anonimyzation in its work program for 2023–2024, the topic is worth revisiting once those guidelines are published (European Data Protection Board 2023).

In terms of potential solutions, Finck stresses on multiple occasions that personal data could (and should) be stored off-chain and merely linked to the blockchain "through a hash pointer" (without going into too many technical details, hash pointers can be defined as a "data structure containing the previous block's hash value and a pointer to that block", basically meaning that the data is chronologically ordered in a manner that makes it difficult to tamper with information without altering subsequent blocks on the blockchain. The term "hash" refers to a cryptographic hash function, such as SHA256). "In such a scenario," Finck explains, "personal data is recorded in a referenced encrypted and modifiable database and not on the blockchain," solving, to a large extent, the fundamental issue of blockchain and data protection, as no personal data will be stored on-chain and would therefore be visible to practically anyone with an internet connection. Nevertheless, Finck is careful to point out that "[d]evelopers working on such solutions must, however, be careful to ensure that metadata is also treated appropriately as it can reveal personal information even where personal data is not directly stored on-chain".

Furthermore, Finck concludes, "[b]lockchains are a technology that might in the future achieve some of the objectives inherent to the GDPR through technological means, although through mechanisms distinct from those envisaged by the legal framework itself". The possibility of such future achievements is at the very heart of the current analysis, as it traces the earlier concerns expressed by academia and technical development alike and finds them within the solutions currently developed in the blockchain space (see Section 3 below).

Therefore, Primavera De Filippi's research deserves a special mention in this article, as it takes a broader view, analysing the interplay between decentralisation and privacy in blockchain technologies. In a nutshell, in her article titled "The interplay between decentralization and privacy: the case of blockchain technologies", De Filippi underscores the tension between the inherent transparency of blockchain and the need for user anonymity, particularly in the context of decentralized finance. Her work brings to light the challenges in maintaining privacy in decentralized systems and the potential trade-offs between decentralization, security, and privacy.

De Filippi concludes that "in spite of the obvious benefits they provide when it comes to data sovereignty, decentralized architectures also present certain characteristics that – if not properly accounted for – might ultimately impinge upon users' privacy," further adding that while such architectures are capable of preserving the confidentiality of data, they cannot "easily protect themselves against the analysis of metadata". Therefore, if not properly designed, "decentralized infrastructures intended to promote individual privacy and autonomy might turn out to be much more vulnerable to governmental or corporate surveillance than their centralized counterparts" (De Filippi 2017).

Further, De Filippi explains the fundamental differences between centralized and decentralized architectures, paving in the process also our current understanding of the dichotomy between privacy and blockchain as a conflict between a framework aimed at minimizing centralized control over data and a technical architecture created to make central control impossible:

"[D]ecentralised systems are much more difficult to implement than centralized platforms. In order to allow for an effective coordination amongst a distributed network of peers, decentralized architectures generally rely on the disclosure of everyone's interactions. Hence, if the price of centralization is trust (as users need to trust centralized operators with their data), decentralization comes at the price of transpar-

ency (as everyone's interactions are made visible to all network's nodes)." Echoing Finck's delineation between blockchain's two types of data, De Filippi further adds that "while decentralized architectures can provide more privacy at the content layer (to the extent that content has been encrypted), they cannot protect themselves against the third parties' analysis of data (or metadata) which are publicly disclosed on a decentralized network". Therefore, the possible solution would be to implement additional technical means to protect the confidentiality of online communications, such as advanced cryptographic techniques, leading to a future where "in spite of the apparent dichotomy between transparency and privacy, there is no real conflict between the two."

The analysis of these foundational works is crucial, as, to a large extent, they set the stage for understanding the current blockchain privacy challenges that must be addressed for its future development. They are, in a sense, also a blueprint for finding solutions.

LATEST PROPOSALS ON PRIVACY AND BLOCKCHAIN

Apart from the theoretical concerns, it is no less important to highlight the significant developments taking part in the technical side of blockchain applications, occurring since the articles highlighted above were published.

The co-creator of Ethereum, Vitalik Buterin, has been directly involved in fixing some of the issues around the Ethereum blockchain, key among which was the migration to a more ecologically friendly consensus mechanism through the adoption of Proof-of-Stake (Ethereum 2024). However, he has also been focusing on addressing concerns about privacy in blockchain. In the "Blockchain Privacy and Regulatory Compliance: Towards a Practical Equilibrium" article, authored by Buterin and a collective of technical experts (further referenced as Buterin et al.), the authors call for creating a balance between maintaining user privacy and fulfilling regulatory compliance through the development of "Privacy Pools", citing the currently insufficient privacy on-chain as the rationale behind their work (Buterin 2022).

These pools utilise advanced cryptographic techniques, such as zero-knowledge proofs (a zero-knowledge proof, or ZKP, is a cryptographic method which makes it possible to prove the validity of a statement without revealing the statement itself. In the ZKP context, the 'prover' is the party trying to prove a claim, while the 'verifier' is responsible for validating the claim (Ethereum May 2024) and, in particular, the ZK-SNARKs (the so-called ZK-SNARKs, a General-purpose zero-knowledge proofs that allow a prover to prove mathematical claims about some combination of public data and private data that the prover holds in such a way that satisfies two key properties: "Zero-Knowledge: nothing about the private data is revealed, aside from the fact that the private data satisfies the claim that is being proven" and "[s]uccinctness: the proof is short (in bytes), and can be verified very quickly, even if the underlying claim being proven involves a heavy computation that takes a very long time to run."), to ensure transactional privacy while adhering to regulatory requirements (such as, potentially, to the GDPR by the creation of custom privacy pools that would align with the regulation's requirements without exposing any personal data). As described by the authors themselves, the core idea of the proposal is to "allow users to publish a zero-knowledge proof, demonstrating that their funds (do not) originate from known (un-) lawful sources, without publicly revealing their entire transaction graph". The above can be achieved by membership in custom association sets that are able to satisfy certain properties required by regulation — for example, GDPR — as well as social consensus and agreements, with privacy pools merely giving "additional options by extending the users' action set. [The users] can still provide more detailed proofs to specific counterparties, if needed" (Buterin 2022). Therefore, participation in such a custom association should not preclude the rights of the data subjects to exercise control over the data they share with other parties. Instead, it would allow for the execution of on-chain transactions without exposing its content or metadata.

Furthermore, this approach would represent significant strides in addressing other privacy challenges in blockchain, particularly in the context of decentralized finance and the use of smart contracts. Its basis is the introduction of two sets of smart-contract-based "proofs", or association sets that can be used interchangeably: membership proofs ("I prove that my withdrawal comes from one of these deposits") or exclusion proofs ("I prove that my withdrawal does not come from one of these deposits") which are

then used "to reach a separating equilibrium between honest and dishonest protocol users" (Buterin 2022). It is important to note that the privacy pools as described exist beyond the theoretical ideation and are already available to the Ethereum blockchain users – with more information provided on the website privacypools.com and by accessing the documentation required for their execution on-chain from a GitHub repository – https://github.com/ameensol/privacy-pools.

In more legal terms, their proposal will likely still fall within the current scope of the GDPR, as zero-knowledge proofs and their sub-category ZK-SNARKs would still be considered an encryption mechanism. Nevertheless, the "Privacy Pools" concept provides an interesting solution to the privacy on blockchain conundrum – one that is based on data subjects exercising control over their data and actively participating in the process of its placing in a specific pool by also following its rules and incentive mechanisms. What is more, the de-linking between the pooled deposit and the original deposit provides a solution to Finck's on-chain data storage issue, as no personal data will be shared publicly due to the use of validation mechanisms that don't rely on data exposure. However, as is the case with many such proposals, despite the positive outlook, time and adoption will show the exact benefits of this solution in practice.

CONCLUSION

In conclusion, while there are significant challenges regarding privacy in the realms of blockchain, decentralized finance, and smart contracts, there is also substantial and promising work underway to address these issues. The insights and proposals from experts like Michèle Finck, Primavera De Filippi, and the work by the likes of Vitalik Buterin highlight a dynamic field where technological innovation is continually being balanced against privacy concerns and regulatory requirements. As the technology evolves, these ongoing efforts are crucial in shaping a blockchain ecosystem that upholds privacy without compromising on the decentralized principles that grant it with the potential of being a truly transformative technology. It is, therefore, also crucial that the topic is often revisited in order to interpret the legal and privacy implications of the latest technical developments, ensuring that the blockchain ecosystem not only evolves but does so in a manner that aligns with core principles of user privacy and data protection.

REFERENCES

Article 29 Data Protection Working Party (2014). Opinion 04/2014 on Anonymisation Techniques, 0829/14/EN,.

Buterin, V. et al. (2022). Blockchain Privacy and Regulatory Compliance: Towards a Practical Equilibrium. In: *Crypto Valley Conference on Blockchain Technology (CVCBT 2022),* LNCS, vol. 12345, pp. 567–585. Springer, Cham.

Carapella, F., E. **Dumas,** J. **Gerszten,** N. **Swem,** L. **Wall** (2022). *Decentralized Finance (DeFi): Transformative Potential & Associated Risks.* Federal Reserve Bank of Boston.

De Filippi, P. (2017). *The Interplay Between Decentralization and Privacy:* The Case of Blockchain Technologies. Journal of Peer Production (9), 75–96.

European Data Protection Board. EDPB Work Programme 2023–2024. [Online]. Available: https://edpb.europa.eu/system/files/2023-02/edpb_work_programme_2023-2024_en.pdf.

Ethereum. Consensus Mechanisms – Proof of Stake (PoS). Available at: https://ethereum.org/en/developers/docs/consensus-mechanisms/pos/.

Ethereum. Zero-Knowledge Proofs. Available at: https://ethereum.org/en/zero-knowledge-proofs/.

Finck, M. (2018). Blockchains and Data Protection in the European Union. *European Data Protection Law Review 4(1),* 17–35.

Finck, M. (2019). Smart Contracts as a Form of Solely Automated Processing Under the GDPR. *Journal of Information, Communication and Ethics in Society 17(2),* 177–194.

GDPR (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union L 119,* 1–88.

Goldenfein, J., A. **Leiter** (2018). Legal Engineering on the Blockchain: "Smart Contracts" as Legal Conduct. Law and Critique. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3176363.

Nakamoto, S. *Bitcoin: A Peer-to-Peer Electronic Cash System.* https://bitcoin.org/bitcoin.pdf.

ТЕМАТА ЗА ПОВЕРИТЕЛНОСТТА В КРИПТО: ТЕКУЩИ ПРЕДИЗВИКАТЕЛСТВА И (БЪДЕЩИ) РЕШЕНИЯ

Резюме: Тази статия проследява развиващата се връзка между защитата на личните данни и блокчейн технологията, като разглежда предизвикателствата и иновациите в контекста на децентрализираните финанси (DeFi) и интелигентните договори. Изследва се сложното взаимодействие между прозрачността на блокчейна и необходимостта от защита на личните данни на потребителя, стъпвайки върху изводите от предишни изследвания, включително трудовете на Мишел Финк и Примавера Де Филипи, които анализират въздействието на GDPR и дихотомията на децентрализацията и поверителността. Статията допълнително се задълбочава в скорошни предложения, по-специално „Privacy pools“ на Виталик Бутерин и др., насочени към баланса между поверителността и съответствието с нормативните изисквания чрез използването на криптография. Анализът включва обзор на продължаващите предизвикателства, свързани с поверителността, като същевременно подчертава обещаващите развития, които дават възможност на блокчейн технологията да постигне хармонично съжителство с принципите за защита на личните данни.

Ключови думи: блокчейн, поверителност, GDPR, доказателства с нулево знание, Privacy pools

Вяра Савова, докторант
Университет по библиотекознание и информационни технологии
E-mail: v.savova@unibit.bg