A blockchain-based framework for GDPR-compliant personal data processing

Chiara Braghin, Stelvio Cimato, Marco De Simone
Università degli Studi di Milano

March 31, 2023

Consumers have become increasingly connected and are constantly creating and sharing information online: they are researching, purchasing and using online products and services, via any number of connected devices. All of this customer data is being collected by service providers, device manufacturers, desktop and mobile apps, Internet providers and mobile operators for their own purposes, or to be sold to their business partners. The potential consequences of such a distributed approach to personal information exposure was demonstrated by Facebook's Cambridge Analytica scandal [5], where a third party app saw millions of users' profile data scraped, allegedly to influence the outcome of the 2016 US election. The effect was, on one side, the raising concerns of individuals about their privacy and the protection of their personal data; on the other side, the update of privacy laws and regulations, e.g., the General Data Protection Regulation (GDPR) [3] in Europe.

The European data protection and privacy law introduced several obligations and rights, whose compliance requires undertaking new practices for activities involving consent and personal data. It increased the potential fines organisations face for misusing data, and made it easier for people to discover what information organisations have on them. In other words, the GDPR expects businesses not only to comply with EU data protection requirements, but also to be able to demonstrate their compliance.

After the GDPR coming into effect in 2018, different works addressed the processing of personal data according to the new regulation using blockchain technology [6, 2, 1, 4], but most of them are either only theoretical, or they propose solutions ad hoc for specific scenarios.

In our work, we defined a framework that satisfies the set of rules imposed by the GDPR with respect to data processing by combining a blockchain, used as an access-control moderator, with an off-blockchain classical storage solution. Users are not required to trust any third-party and are always aware of the data that is being collected about them and how it is used. In particular, the blockchain, i.e, a decentralised tamper-proof public ledger, is used to record each interaction between the data subject and the service provider, bringing personal data processing to a level of privacy and security that prioritises data subjects and shared transparency, as required by the GDPR. Moreover, we use smart contracts encoding GDPR legal requirements into the blockchain itself, so that they are enforced automatically.

The smart contract then becomes an access control manager that does not

require a trusted third party, and the blockchain can act as a tamper-proof ledger to record digital interactions: the data subject can now verify directly where his personal data is stored and put to (commercial) use. Since the blockchain is a public ledger, it cannot record all data, thus we combine blockchain and off-blockchain storage to define a personal data processing platform focused on GDPR-compliance.

The feasibility and scalability of the approach is shown by means of a prototype implemented in Ethereum with Solidity smart contracts, with a Graphical User Interface to increase the usability of the application. Some cost considerations were done on the specific implementation. The framework consists of the following components: a server-side servlet application accessing the blockchain and the database, a web page for the user registration phase, a client-side application for accessing user's data, a smart contract and a Java library. The library offers APIs to the functions of the smart contract, thus providing an abstraction from Ethereum. By using the library, developers can implement GDPR-compliant client and server applications since by construction they meet the restrictions imposed by the GDPR and satisfied by the contract.

References

[1] Darine Ameyed, Fehmi Jaafar, Francis Charette-Migneault, and Mohamed Cheriet. Blockchain Based Model for Consent Management and Data Transparency Assurance. In *2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, pages 1050–1059, 2021.

[2] Cristòfol Daudén-Esmel, Jordi Castellà-Roca, Alexandre Viejo, and Josep Domingo-Ferrer. Lightweight Blockchain-based Platform for GDPR-Compliant Personal Data Management. In *2021 IEEE 5th International Conference on Cryptography, Security and Privacy (CSP)*, pages 68–73, 2021.

[3] European Parliament. 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, and repealing Directive 95/46/ec (General Data Protection Regulation). *Official Journal of the European Union*, L 119:1–88, 2016.

[4] Ricardo Neisse, Gary Steri, and Igor Nai-Fovino. A Blockchain-based Approach for Data Accountability and Provenance Tracking. In *Proceedings of the 12th International Conference on Availability, Reliability and Security*, ARES '17, pages 14:1–14:10, New York, NY, USA, 2017. ACM.

[5] The Guardian. The Cambridge Analytica Files - A year-long investigation into Facebook, data, and influencing elections in the digital age. *The Guardian*, 2018.

[6] Nguyen Binh Truong, Kai Sun, Gyu Myoung Lee, and Yike Guo. GDPR-Compliant Personal Data Management: A Blockchain-Based Solution. *IEEE Transactions on Information Forensics and Security*, 15:1746–1761, 2020.