The GDPR and the blockchain technology

Ministry of Digital Affairs, DLT and Blockchain Working Group

The document expresses the views of experts participating in the work of sub-team for eID – GDPR – AML, thus is not the official position of the Minister of Digital Affairs.

This document was prepared within the framework of the work of the working group for DLT and blockchain, operating within the Distributed Ledgers stream, established by the Decision no. 7 of the Chairman of the Committee of the Council of Ministers for Digital Affairs of 10 October 2018, changing the decision on the establishment of the "From Paper to Digital Poland" Task Force.

This document was prepared and edited by a team of authors:

Jan Byrski, Law Firm, Traple Konarski Podrecki & Partners

Agnieszka Hołownia-Niedzielska, ProtectAuth

Marta Kownacka, Law Firm LawIT

Janusz Łaski, ING Bank Śląski SA, Council of Custodian Banks at the Polish Bank Association

Piotr Rutkowski, NASK PIB, Ministry of Digital Affairs

Michał Starczewski, Law Firm BWHS

Michał Tuszyński, Law Firm GWW Legal

Marcin Zarakowski, Lisk Foundation

Introduction

Although the GDPR was prepared as a technologically neutral legal act, its underlying assumption are centralist databases. In such databases, there is two-way communication between the entity that manages the database, and its users. In principle, identifying the data controller does not pose any major difficulties. Such an assumption does not correspond to the architecture of distributed databases such as a blockchain. This difference leads to certain difficulties in determining the conformity of some blockchain-based solutions with the GDPR.

Despite the fact that the blockchain technology and, in particular, some of its characteristics (unchangeability of records, the open nature of ledgers) seem to be in conflict with the assumptions of the GDPR, further development of solutions based on this technology may contribute to increasing the protection of personal data and to improving their free flow. Thanks to the use of the blockchain technology and distributed ledgers, entities may regain control over their own personal data and may – in a way that is safe to them – decide about the scope, recipients, and time of sharing. A blockchain network, as a durable electronic medium, guarantees the unchangeability of a record, which means that it is impossible to use data and then modify or delete them. Because of this, the blockchain technology may be an excellent way to prevent unauthorized use of data.

The present study addresses situations where personal data are to be recorded in a blockchain network. It is worth mentioning that, in the case of recording or transmitting data concerning legal persons enabling identification of legal persons or their activities, there is no need to meet the requirements arising from the GDPR because identification of a legal person does not mean identification of a natural person. Therefore, any doubts related to the GDPR will not apply to information stored in blockchain networks whose participants are legal persons, in particular those identified by means of an electronic seal or a similar solution. An electronic seal may also be a partial answer to the needs of other blockchain-based technological solutions, where the issues concerning the GDPR would only be considered in the context of customers or institutions that do not have a legal personality.

Conformity with the GDPR is not so much a matter of a technical solution itself, as of the way in which it is used. This is why, in the end, the legitimacy of the applied solutions should be assessed on a case-by-case basis.

It is worth noting that this document presents two ways of considering the issue. The first part concerns the basic concepts and their reference to the blockchain technology and to the ways in which the technology can meet the requirements of the GDPR. The second part discusses the legitimacy of statutorily restricting some rights under the GDPR due to the great importance of the blockchain technology. Both concepts are worth exploring, and the observations that they present allow to fully understand the complexity of relating the blockchain technology to the issue of personal data protection as well as the ways in which a clear position in this respect can be reached.

The Working Group may present advice on how to interpret the existing laws, and may provide technological guidelines that enhance the security of data in networks. It should be noted, however, that the institution competent to perform supervisory functions with regard to the processing of personal data is the President of the Personal Data Protection Office.

Key terms

For the avoidance of doubt, this chapter lists definitions of the terms used in the document.

Technology-related terms

- **Blockchain** – a distributed database, which is shared in a network of computers, and which operates according to the rules specified for the given network. The name means a chain of blocks, i.e., the way in which information is organized by recording it one piece after another, together with an indicator of the previous message. Both the indicator and the contents of the block are cryptographically secured[1]. Blockchain networks consist of nodes, which are participants in the network, and which store data.

[1] based on The Polish Chamber of Information Technology and Telecommunications, "Blockchain w Polsce możliwości i zastosowanie", Warsaw 2018

- **Hash function** – a function that, from any string of data, generates a string of characters of a specific length (see hash)[2]. The data for which the function is executed can be any strings of characters, lists of information, computer files or programs. An important characteristic of the hash function is collision resistance, which means minimization of the risk of obtaining the same result from two different pieces of information.
- **Hash** (index, indicator, irreversible abbreviation) – the effect of the operation of the hash function – a sequence of characters attributed to specific data that have an electronic form. This document refers to indicators that are created as a result of a one-way function, i.e., a function that allows to obtain the indicator from data, but not vice versa. Such functions, due to their collision resistance [cf. hash function], serve as cryptographic protections.
- **Miner** – a person or institution that uses the computing power of a computer (or computers) to create further blocks of information in a blockchain network. This process is called 'mining'. In return for creating another block, the miner receives a reward in the form of cryptocurrency.[3]
- **Peppered hash** – a cryptographically safer hash, which is created by introducing an additional, random, secret component (known only to the hasher) into the information subjected to the hash function.[4] The additional component is the same for all the data in a specific section, e.g., for the passwords in a single application.
- **Proof-of-Stake (PoS)** – a way of adding blocks to the chain (consensus method), which may be compared to property rights attached to shares in a capital company. The reward for the validating of a block can be allocated based on the number of held assets of a specific type. As a result, there is no need to connect as much computing power to the network as in the case of Proof-of-Work,[5]
- **Proof-of-Work (PoW)** – a way of adding blocks to the chain (consensus method) that has been used, e.g., in the Bitcoin network. In this type of network, all the nodes have the ability to allocate another block of data, and the likelihood of a node adding another block depends on the computing power used by the node.[6] In networks, PoW is executed by miners.
- **Salted hash** – a cryptographically safer hash, which is created by introducing an additional, individual, random, secret component into the information subjected to the hash function.[7] The additional component is different for each execution of the hash function, e.g., for each of the passwords in a single application.
- **Private network** – a blockchain network, which can be joined only after the requirements of specific institutions have been met, e.g.: an invitation, signing a contract, having a specific legal status, etc. Only after consent has been obtained is it

[2] based on Rajeev Sobti1, G.Geetha, "Cryptographic Hash Functions: A Review", International Journal of Computer Science Issues Vol.9 2012

[3] based on Jakub A. Bartoszewski, "Blockchain Compass 2018", The Startup Poland Foundation, Warsaw 2018

[4] based on https://en.wikipedia.org/wiki/Pepper_(cryptography), retrieved on 14/10/2019

[5] quoted from: The Polish Chamber of Information Technology and Telecommunications, "Blockchain w Polsce możliwości i zastosowanie", Warsaw 2018

[6] quoted from: The Polish Chamber of Information Technology and Telecommunications, "Blockchain w Polsce możliwości i zastosowanie", Warsaw 2018

[7] based on https://en.wikipedia.org/wiki/Salt_(cryptography), retrieved on 14/10/2019

possible to become a node of the network, download data from it, or share them with other verified entities.[8]

- **Public network** – a blockchain network that can be joined by any person or institution thus becoming a network node, downloading data and sharing them with other entities.[9] No consent of any entity is needed in order for a new participant to become a part of the network.
- **Permissioned network** – a blockchain network, where adding new information and validating its authenticity is available only to entities that have obtained permission to join the network. The vast majority of permissioned networks are private networks, which is why these terms are often considered to be synonymous.
- **Permissionless network** – a blockchain network, where adding new information and validating its authenticity is available to everyone and does not require permission of any entity or person. The vast majority of permissionless networks are public networks, which is why these terms are often considered to be synonymous.

Legal terms

- **Personal data controller** – 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.
- **Anonymization of data** – a term not directly defined in legal acts, which means irreversible processing of data in such a way that the data subjects cannot be identified.
- **Personal data** – any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.
- **Processor** – an entity that processes personal data on behalf of the Controller and at the Controller's request; the Processor does not independently determine the purposes or method of data processing.
- **Pseudonymization** – the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person; the pseudonymization process is reversible (unlike the anonymization process) and the personal data processed in such a way can be attributed to a natural person again.
- **Data processing** – any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration,

[8] prof. Krzysztof Piech (ed.), "Leksykon Pojęć na temat Technologii Blockchain i Kryptowalut", 2016

[9] prof. Krzysztof Piech (ed.), "Leksykon Pojęć na temat Technologii Blockchain i Kryptowalut", 2016

retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

- **The GDPR** – 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), [OJ L 119/1, 2016].
- **Joint controller** – where two or more controllers jointly determine the purposes and means of personal data processing, they are joint controllers;

Main identified problems and needs

The decentralized nature of a blockchain network, and the permanence of the information recorded in it, give rise to two types of difficulties with regard to personal data processing by means of this technology.

The first type of difficulty is related to identification of the roles defined by the GDPR. Who is the personal data controller? Is it possible to identify the data processor? Depending on the answers to these questions, the obligations and liabilities of participants in a blockchain network will vary.

The second type of difficulty arises from the obligation to ensure a number of rights of data subjects. In particular, the right to know who, how, and for how long processes the data. Also, the right to demand that the data be rectified, or that their processing be ceased. The permanence of information recorded in a blockchain makes it difficult (if not impossible) to fulfil all these rights.

Indication of needs, based on the example of a business case

A clear indication of the relationship between the obligations under the GDPR and the blockchain technology is crucial for future applications of the technology. The uncertainty with regard to being able to store or transmit data via private or public networks leaves the market in a position where it has to decide not only to develop new products, but also to take certain risks, or to withdraw from solutions, which are needed and sought-after on the market, due to doubts concerning legal aspects. The issue of personal data is present in so many projects, products, and services that many companies working with new technologies have already come across it or will come across it in the near future. Discussed below is an example of a business solution, for which issues relating to blockchain networks and personal data are of fundamental importance.

Validation and history of a customer's activities in a public network

A public blockchain is characterized by the fact that it is independent and accessible to everyone on equal terms – to both an individual customer and a huge corporation. Because

of this characteristic, the information recorded in the network is equally available to both parties, and is validated by an independent technology. If the customer has performed some important activity in the vendor's system, then it is in the interest of both parties (which interest is often required by consumer law) to record this activity on a durable medium allowing access for both parties. An example of such activities is the concluding of an insurance contract, or the placing of an individual order. In such a case, the basic parameters of the activity should get recorded in the network, including who performed it and who approved it. In the described case, the institution accepting the order records its data in the network, and the customer (via a website or application) confirms the data, thereby confirming his/her commitment – in the case of an individual order, or the time of conclusion of the policy – in the case of insurance. If there are any doubts or conflicts, both parties are able to obtain a clear confirmation of the subject matter and of the parties involved in the process, without the possibility of modifying the data. At the same time, an order history of the specific customer is created in the network, and this history is available to the customer even after the end of cooperation with the institution. Such a solution has full business functionality only when the network contains personal data of the customer, the vendor's employee (if such an employee participates in the process), and the involved persons, e.g., the insured.

The personal data controller in projects based on the blockchain technology

In the case of permissioned networks, the personal data controllers are usually the network operators, who often act as joint controllers (Article 26 of the GDPR).

In the case of permissionless networks, provided that they are kept properly decentralized, the controllers should not be considered to be the entities participating in the consensus mechanism – the so-called network nodes (e.g. miners for Proof-of-Work, and validators for Proof-of-Stake). They do not decide about the purposes of data processing, but they merely perform technical activities abstracted from the contents of data contained in the processed transactions.

This does not mean that in the context of permissionless blockchains there will never be data controllers. These might be the entities that use a decentralized infrastructure for their business activities. In such a case, however, there may arise the problem of fulfilling the controller's obligations towards data subjects, and sometimes it will be possible to apply Article 11 of the GDPR.

The principle of minimizing the recording of personal data on the blockchain

Recording of personal data in blockchain-based ledgers entails significant risks for the natural persons who are the subjects of such data, as well as the inability to fully fulfil the rights and obligations under the GDPR. Therefore, one should refrain from or limit to a minimum the recording of personal data in blockchain-based ledgers.

Off-chain recording of personal data

If possible, personal data should be processed off-chain, i.e., outside the blockchain. The blockchain may contain links (e.g. hash-pointers) that allow to verify the correctness of the data. Such solutions are a kind of pseudonymization of data. Off-chain processing of all the personal data allows to avoid the difficulties connected with the use of distributed databases according to the GDPR. All the solutions that effectively prevent access from the blockchain to information about personal data for people who do not know the relevant password, will be sufficient from the perspective of the GDPR.

Personal data processed off-chain are stored in a centralized database, where it is easy to identify the personal data controller responsible for fulfilling all the obligations such as, first and foremost, the obligation to provide information and the obligation to ensure the rights of the persons whose data the controller is processing. It should be borne in mind, however, that a change to the personal data (e.g. their updating or supplementing) may result in a discrepancy between the link stored in the blockchain and the data processed off-chain.
If the pseudonymization requirements are fulfilled, conformity with the GDPR should be attributed to the solutions that store personal data outside the blockchain (off-chain) and only leave links to such data in the blockchain-based ledger, e.g. in the form of hash-pointers. Such solutions may be attractive mainly to private or permissioned networks, but not to public or permissionless networks, because they involve a restriction on the decentralization of the blockchain-based ledger as well as the introduction of a kind of 'trusted third party' that takes care of the personal data recorded outside the ledger (off-chain). In such a case, the data controller is the entity that stores the data off-chain, and the ledger that this entity keeps must conform to the requirements of the GDPR. Individual nodes of the network do not act as joint controllers or processors.

In order to ensure that the manner of recording the link to off-chain-stored data makes it impossible to identify the person, it is helpful to use a technology that prevents identification of the person even if someone is in possession of that person's data. This applies to a situation where the indicators in the blockchain network would be recorded by company X in the form of a standard hash function executed, e.g., on names and surnames. Having a sufficiently large database of names and surnames, it would be technically possible to determine which persons the data indicators recorded in the network correspond to. Such actions are prevented by the latest technical and technological means, such as the use of an irreversible, properly complex hash function that contains additional information not related to the data being subjected to the function's operation (salt and pepper hashes).

Pseudonymization and anonymization techniques vs. blockchain

The GDPR does not indicate any specific techniques (technological neutrality). Data controllers will be held accountable for the effects – whether or not the data processed by them have been anonymized or pseudonymized. Presented below are examples of techniques that may lead to effective anonymization or pseudonymization of data. It cannot be precluded that solutions based on adequate cryptographic protection of personal data

stored in blockchain-based ledgers will provide a sufficient, GDPR-conformant level of security for data subjects. The data protection standard required by the GDPR may also be met (provided that relevant conditions are fulfilled) in the case of the use of technologies that make it impossible or very difficult to identify persons thanks to the inability to decrypt the data recorded on the blockchain. Every solution that increases data security requires an individual analysis in the context of the adequacy of application as well as the possibility of identification based on separating out entries about a specific person or based on a specific record (e.g. a single record about a specific attribute, e.g., a characteristic large amount). When considering the available solutions, the following technological possibilities are worth being taken into account:

- encryption with one-time keys,
- hash functions resistant to attacks using quantum computers,
- salted hashes,
- peppered hashes,
- stealth addresses – one-time addresses generated for the purpose of the recording of data or the transmitting of transactions,
- ring signatures – a scheme for the signing and encrypting of messages, which is used by a group of people and allows to verify the authenticity of a signature but, at the same time, does not allow to identify the person who used it,
- ring confidential transactions – a method for the transmitting of transactions in a blockchain network, where the transmitted value is unrecognizable to persons other than the participants in the transactions.

The selected solution should also be analysed in terms of:

- technological development that will increase the chances of increase the level of security of the solution, e.g., quantum computers,
- the time for which the recorded data will be valuable to someone who will gain access to them (e.g. data concerning the purchase of real estate vs. data concerning purchases at a personal cosmetics shop),
- the amount and content of information that will be added to the network and that concerns a single person – the possibility of linking the information,
- authorized persons gaining and, if necessary, regaining access to the information in the case of, e.g, loss of the device holding the encryption key, or deletion of the data needed to read the information.

Possibility of statutorily restricting certain rights of data subjects under the GDPR in the case of personal data processing by means of the blockchain technology

In the light of Article 23 of the GDPR, it is possible to restrict (based on both the law of the European Union and the law of a Member State, including Polish law) the scope of the obligations and rights provided for in Articles 12 to 22 and Article 34, as well as Article 5 of

the GDPR in so far as its provisions correspond to the rights and obligations provided for in Articles 12 to 22.
Such a restriction is possible only if it respects the essence of the fundamental rights and freedoms and is a necessary and proportionate measure in a democratic society, serving the objectives set out in Article 23(1)(a)–(j) of the GDPR.

As the literature rightly points out: "Article 23 is therefore the most serious exemption in the process of creating a uniform standard for the protection of personal data. On its basis, the EU legislator allows Member States to adopt laws restricting the rights and obligations of controllers and processors provided for in the Regulation if, in the opinion of the given state, such a restriction is necessary to fulfil certain premises, in particular the public interest of this state or of the entire EU."[10]

It should be borne in mind that the legal act introducing the restriction must contain the detailed provisions referred to in Article 23(2)(a) to (h) of the GDPR. Thus, national laws must contain elements to ensure a certain level of protection of personal data (e.g. as regards the purposes or categories of processing, the categories of personal data, the scope of the introduced restrictions, the establishment of professional secrecy, etc.).

According to Article 23(1)(e) of the GDPR, it is possible to introduce a restriction that serves **other important objectives of general public interest of the Union or of a Member State, in particular an important economic or financial interest** of the Union or **of a Member State,** including monetary, budgetary and taxation matters, public health and social security.

As an example of public interest, recital 73 (of the preamble) of the GDPR indicates, inter alia, **an important economic** or financial **interest** of the Union or of a Member State. In this regard, it should be noted that Article 23 of the GDPR does not stipulate that only public authorities may take advantage of a restriction on the rights of data subjects connected with public interest. Therefore, national regulations in this respect may also apply to private sector entities (and such regulations have already been introduced in the sectoral act, which entered into force on 4 May 2019). Nevertheless, these restrictions should be in line with the requirements of the Charter of Fundamental Rights and the European Convention for the Protection of Human Rights and Fundamental Freedoms.

Due to the very broad possibilities when it comes to application of the blockchain technology in both the private sector (including, inter alia, the financial sector) and the public sector, as well as the numerous advantages (including resistance to IT system failures, resistance to cyberattacks, transparency, low costs, high efficiency and scalability, etc.), it can be reasonably assumed that **the use of this technology is of general public interest being an important economic or financial interest of a Member State**. It seems that, in this respect, the important economic interest may also be related to enabling (in particular by creating an appropriate legal framework) the lawful development of innovative and secure digital services.

[10] *M. Sakowska-Baryła* (ed.), General Data Protection Regulation. Commentary, Warsaw 2018

Undoubtedly, due to the characteristics of the blockchain technology (including its decentralized and distributed nature), it is not possible to effectively or entirely fulfil all the rights under the GDPR (e.g. the rights to rectify data, object to data processing, or erase data [the right to be forgotten]). Problems arising from this have also been reported on the official website of the supervisory authority.[11]

When restricting fundamental rights and freedoms, particular attention would have to be paid to the proportionality and necessity of such action, i.e., it would have to be ensured that **the restriction applies only to those rights that must be restricted in order to enable effective use of the blockchain technology.** Neither could the restrictions introduced in order to fulfil the provisions of Article 23(1)(e) of the GDPR violate the essence of fundamental rights or freedoms.

For example, the introduction of restrictions may be necessary in the context of the fulfilment of the right to be forgotten (Article 17 of the GDPR). Due to the nature of a blockchain, **in my opinion, it would not be possible to conclude that such a right can be effectively fulfilled with regard to personal data recorded in a blockchain**. In this scope, however, appropriate technical and organizational measures should be ensured in order to adequately secure the data being processed, e.g., measures that would allow to limit access to the data.

Similar restrictions would also have to apply, for example, to the right to rectify data (Article 16 of the GDPR) or to the right to object to the processing of data on the grounds of a specific situation (Article 21(1) of the GDPR), provided that the basis for the processing of data would be the legitimate interest of the data controller.

[11] https://techinfo.uodo.gov.pl/technologia-blockchain-a-dane-osobowe/ retrieved on 24/10/2019