Group Briefing

February 2019

Rethinking GDPR, blockchain and data-protection-by-design: An opportunity to innovate?

In this briefing, Colin Rooney and Hugh McCarthy examine a question at the intersection of blockchain technology and the GDPR – does the ledger contain personal data? – and consider the importance of applying the GDPR's data protection-by-design principle at an early stage.

KEY TAKE-AWAYS

1. At least some of the data on the blockchain ledger will likely constitute personal data depending on: (i) the encryption and hashing techniques applied; and (ii) the architecture of the framework;
2. The architecture of a particular blockchain framework is a key part of ensuring GDPR compliance. To meet the GDPR's data protection-by-design principle, a data protection risk / opportunity analysis should be conducted as early as possible in the design stage of the blockchain framework; and
3. Blockchain presents certain challenges to GDPR compliance but innovative approaches can also present potential solutions to many of the GDPR's most challenging requirements.

PERSONAL DATA IN A BLOCKCHAIN CONTEXT

While blockchain technology potentially presents a range of innovative technical solutions to many GDPR compliance issues, the important threshold question is whether data on a blockchain ledger constitutes personal data.

The concept of personal data is central to any consideration of the application of the General Data Protection Regulation (EU Regulation 679/2016) (the "**GDPR**"). A fundamental question arising in the context of blockchain is whether data contained in the blockchain constitutes personal data for GDPR purposes.

The focus of the GDPR is on personal data of natural persons and accordingly, it does not in principle "*cover the*

processing of personal data which concerns legal persons and in particular the undertakings established as legal persons, including the name and the form of the legal person and the contact details" (Recital 14 GDPR). In addition, the GDPR should not apply to "*[F]iles or sets of files, as well as their cover pages, which are not structured according to specific criteria*" (Recital 15 GDPR). Hence where a legal person such as a corporation or partnership makes use of a blockchain solution for the purposes of a transaction, to the extent that this does not involve personal data relating to a natural person, such data will generally be outside the scope of the GDPR.

Personal data is defined broadly in Article 4(1) GDPR as "*any information relating to an identifiable or identified 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*" (Art. 4(1) GDPR).

When assessing whether a natural person is "*identifiable*" account is to be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. In particular, the GDPR provides that to ascertain "*whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments*". Accordingly, the ever-evolving state of the art in terms of computing power and decryption techniques are relevant to this analysis.

Where information is anonymous, "*namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable*", the GDPR does not apply in respect of such data.

IDENTIFIABILITY - REVERSAL AND LINKABILITY RISKS

The debate on the application of the GDPR to blockchain solutions focuses on the identifiability element of the test for personal data. In particular, where hashed data is recorded on the blockchain ledger the question arises as to whether such data can be considered to be sufficiently anonymised such that a living individual can no longer be identified. However, if the hashed data is considered to be merely pseudonymised personal data (rather than fully anonymised data) then the GDPR will apply.

The GDPR defines pseudonymisation as "*processing 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*" (Art. 4(5) GDPR). An example of where this might potentially arise in the context of blockchain is if an individual repeatedly makes use of a key to make transactions. Although the data that key represents is not discernible, the fact that individual has repeatedly used the key in the same context may make it possible, with the help of other information, for a prudent third party observer of the ledger to link it back to a particular data subject through pattern recognition.

It follows that the legal question of whether personal data is being processed is based on whether the encryption techniques applied to hash and obscure the personal data constitute anonymization or pseudonymisation. This in turn involves detailed technical analysis of the encryption and hashing techniques applied to convert the personal data (e.g. personal records or financial information) into digital signatures that are cryptographically linked to the original data (e.g. user keys) with a view to ascertaining the likelihood – taking account of the technology available, the effort and cost – of a third party being able to use the digital signature to identify a data subject, whether through accessing the data the signature represents or through pattern recognition. The EU Blockchain Observatory refers to these (respectively) as the: (i) the "reversal risk"; and (ii) the "linkability risk" associated with the data (Blockchain and the GDPR, 16 October 2018). The EU Blockchain Observatory works under the European Commission to monitor blockchain initiatives in the EU, and to publish guidance on the regulation of blockchain technology.

The reversal risk refers to the likelihood of a determined party being able to reconstitute the original data (e.g. using brute force decryption techniques), whereas the linkability risk is the possibility of a party being able to link encrypted data to an individual by examining patterns of usage and context or by comparison to other pieces of information. The lower each of these risks are – by virtue of the strength of the encryption or other method of disassociating personal data with a data subject (e.g. data aggregation) – the more likely it is that personal data can be considered irreversibly anonymised, with the consequent effect that the GDPR does not apply, at least to the data in question (although the GDPR may still apply to other aspects of the framework where personal data is processed).

CASE LAW

In **Breyer v Bundesrepublik Deutschland**, (2016) Case C-582/14 the Court of Justice of the EU ("**CJEU**") stated that personal data will be considered anonymised where identifying a data subject would be "*practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant.*" The Article 29 Working Party has taken the view in previous pre-GDPR guidance that in order for data to be considered anonymous, the anonymous technique applied must be irreversible (Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques).

The analysis above is consistent with the CJEU's position in the Breyer case in which the Court made clear that where data is transformed to the degree that it cannot reasonably be used to single out

a particular data subject (i.e. it cannot be reverse-engineered or does not link to a data subject) then it will be considered anonymised data.

While the CJEU's Breyer decision pre-dates the introduction of the GDPR, the definition of personal data considered in Breyer is materially similar to the GDPR's definition of personal data. Accordingly, the CJEU's analysis in Breyer remains relevant. The CJEU held that a dynamic IP address (i.e. one which was temporarily assigned to a particular internet user for a particular session) in the hands of a website operator, together with the date and time of the internet usage, did not alone constitute personal data because "*such an address does not directly reveal the identity of the natural person who owns the computer from which a website was accessed*". However, to ascertain whether the IP address constituted personal data the CJEU focused on the word "*indirectly*" which forms part of the definition of "*personal data*" both under the Data Protection Directive and the GDPR. In addition, the Court made reference to the term "any other person", which is also retained in the GDPR definition (albeit in a recital rather than a substantive article), to support its view that "*it is not required that all the information enabling the identification of the data subject must be in the hands of one person*" for such information to be considered personal data.

On this basis, the CJEU considered that the additional information necessary for the website operator to link the IP address to a particular individual could potentially be obtained from a third party (namely the internet service provider via an application to a third party regulatory authority) and consequently that it was "*reasonably likely*" that the relevant website would be legally able to obtain such additional information to indirectly identify the particular internet user. In particular, the CJEU considered that it would not require "*disproportionate effort in terms of time, cost and manpower*" for the website operator to obtain the additional information necessary to link the IP address to the individual and accordingly that the dynamic IP address was to be considered personal data in that instance.

EU BLOCKCHAIN OBSERVATORY AND FORUM – BLOCKCHAIN AND THE GDPR

On 16 October 2018 the EU Blockchain Observatory and Forum published its second thematic report addressing the GDPR compliance issues arising in the context of blockchain technology. In respect of the processing of personal data the Observatory noted that:

- GDPR compliance is not about the technology, but rather about how the technology is used and accordingly that there is no such thing as a GDPR compliant blockchain technology but rather that there are only GDPR compliant use cases and applications;
- reversibly encrypted personal data remains personal data but in respect of hashed data this question will largely depend on the sophistication of the hashing techniques used;
- storage of personal data on public blockchains should be avoided where possible and range of obfuscation, encryption, aggregation and other techniques should be implemented to anonymise such data stored on public blockchains; and
- personal data should be collected and, to the extent possible stored off-chain or if this cannot be avoided, in private permissioned blockchains networks.

DATA PROTECTION-BY-DESIGN

Both at the time of determining the means of processing personal data and when undertaking that data processing, the GDPR requires that "*appropriate technical and organisational measures be put in place… in an effective manner and to integrate the necessary safeguards*" to ensure GDPR compliance and in particular "*to protect the rights of data subjects*" (Art. 25(1) GDPR). In practical terms this means that data protection risks and compliance solutions to those risks need to be factored at the design stage long before such processing goes "live". In short, design with data protection as an afterthought will fall foul of the data protection-by-design principle and increase the risks of non-compliance with the GDPR. By contrast, by considering data protection at early design stage blockchain technology can potentially be deployed to achieve innovative solutions to some of the GDPR's more challenging compliance issues, including:

- the roles of controllers, processors and/or joint controllers;
- data minimisation and accuracy; and
- data retention and the data subject rights to rectification and erasure (the right to be forgotten).

CONCLUSION

The question of whether data recorded on the blockchain in a hashed format constitutes personal data is ultimately a technical one, but the prudent approach is to consider that such information does constitute personal data where there exists even a possibility, together with other information, of identifying a living individual from the data. Notwithstanding the significant technical difficulty required to reverse engineer that hashed data and to link it back to a living individual within the blockchain ecosystem, unless data can be irreversibly anonymised it will be deemed to be pseudonymous data (rather than fully anonymised data) such that the GDPR will apply as if it were personal data.

While this briefing focuses on the basic threshold question of whether the GDPR applies to blockchain platforms, blockchain technology presents many novel challenges but also potential solutions to many of the GDPR's other requirements. By engaging with and addressing these issues at an early stage in blockchain design, blockchain may present less of an obstacle and more of opportunity to craft innovative solutions to GDPR compliance issues.

KEY CONTACTS

COLIN ROONEY
PARTNER
+353 1 920 1194
colin.rooney@arthurcox.com

HUGH MCCARTHY
ASSOCIATE
+353 1 920 1324
hugh.mccarthy@arthurcox.com