Practitioner's Corner

Blockchain Technology and the GDPR – How to Reconcile Privacy and Distributed Ledgers?

*Matthias Berberich and Malgorzata Steiner**

Blockchain technology raises a multitude of legal questions that are still in early stages of discussion. While legal research is so far mainly focused on financial regulation of cryptocurrencies, distributed ledger technology in general brings considerable privacy implications. While the EU Commission expects the new General Data Protection Regulation and its technological neutrality to enable 'innovation to continue to thrive under the new rules,'[1] *some features of Blockchain pose questions under EU data protection principles.*

I. Technical Core Features and Use Cases of the Blockchain Technology

Blockchain (BC) is widely depicted as the most disruptive technology since the advent of the Internet. While this technology was used at first for the virtual currency Bitcoin, its current applications go far beyond cryptocurrencies, and its potential is compared with the TCP/IP protocol that forms the backbone of today's Internet. A vast scope of business models can be built upon BC, ranging from FinTech products like cryptosecurities or 'smart bonds' over 'smart property' registers, to so called 'smart contracts' with transactional protocols which assume the formation, performance and execution of a fully electronic contract.

In essence, BC is a distributed ledger technology with three constituting elements.[2] *Firstly*, BC is a continuously amended and *persistent* ledger, which means that all transactions effected on BC are perpetually stored. New transaction information is added in a new block and connected to the previous block in the chain, so that a BC will grow over time and include all transactions ever made. *Secondly*, BC is a *distributed* peer-to-peer ledger, stored on every node of the system as a complete BC copy. If new transactions are effected, the majority of nodes must verify the legitimacy of the effected transaction and, if confirmed, every BC copy is updated accordingly. With this distributed authentication process, the core feature of BC is the lack of a central entity or intermediary. And *thirdly*, BC is asymmetrically *encrypted* and requires private and public keys to effect transactions.

The recent enactment of the General Data Protection Regulation 2016/679 (GDPR), which shores up the level of data protection throughout the EU and expands its territorial scope (II.), might bring a tension between general data protection principles and the core features of BC technology[3]: If encrypted, BC transactions involve personal data and are not fully anonymous (III.), the decentralised BC nature without a central entity as data controller will pose a challenge for regulators (IV.). Moreover, the persistence

[*] Dr Matthias Berberich, LL.M. (Cambridge), is Attorney-at-Law and visiting lecturer at the Humboldt University Berlin; Malgorzata Steiner, MPP (Harvard), is former Head of Department in the Ministry of Administration and Digitisation in Poland and Senior Advisor at the *Stiftung Neue Verantwortung* in Berlin.

[1] European Commission, 'Questions and Answers - Data protection reform' (Press release, 21 December 2015) <http://europa.eu/rapid/press-release_MEMO-15-6385_en.htm> accessed 13 August 2016.

[2] For technical details cf Jeni Tennison, 'How might we use blockchains outside cryptocurrencies?' (21 May 2015) <http://www.jenitennison.com/2015/05/21/blockchain.html> accessed 13 August 2016; especially on BITCOIN cf Satoshi Nakamoto, 'Bitcoin: a peer-to-peer electronic cash system' (*Bitcoin*, 2012) <https://bitcoin.org/bitcoin.pdf> accessed 13 August 2016; Michael Nielsen, 'How the bitcoin actually works' (*DDI*, 6 December 2013) <http://www.michaelnielsen.org/ddi/how-the-bitcoin-protocol-actually-works> accessed 13 August 2016.

[3] For privacy as a general issue for blockchains see also Markus Kaulartz, 'Die Blockchain-Technologie, Hintergründe zur Distributed Ledger Technology und zu Blockchains' (2016) Computer und Recht 474, 479; Donald B Johnston, 'More on the Law of the Blockchain' (*Lexology*, 11 April 2016) http://www.lexology.com/library/detail.aspx?g=ee10a0f-3447-4088-81dd-7521244fecb3 accessed 13 August 2016; Gregory Brandman and Samuel Thampapillai, 'Blockchain- Considering the Regulatory Horizon' (*University of Oxford*, 7 July 2016) https://www.law.ox.ac.uk/business-law-blog/blog/2016/07/blockchain-%E2%80%93-considering-regulatory-horizon accessed 13 August 2016.

of all transactions may raise tensions with the privacy by design principle (V.) and the right to be forgotten (VI.) enshrined in the GDPR.

II. Territorial Scope of the GDPR – A Long Arm on Distributed Ledgers

Expanding the territorial scope of application of EU data protection laws, the GDPR applies either where a data controller or processor has a place of establishment in the EU, regardless of where the actual data processing takes place.[4] Alternatively, it applies to the processing of personal data of data subjects in the EU by a controller or processor not established there, where processing activities are related to the offering of goods or services (irrespective of payments) or to tracking the behaviour of data subjects within the EU.[5]

This broad scope of application will most likely provide a wide catcher also for distributed ledgers like BC. Under the establishment principle the GDPR is likely to apply to operating entities of a private BC in the EU, or, in the case of a public BC, to every node in the EU, if it would be considered data controller. This blurred question of who, if anyone, can be considered data controller in a public distributed ledger (see IV.) will thus arise even for determining the territorial application of the GDPR. Under the extraterritorial effect[6] of the market principle, BC data processing will fall within the ambit of the GDPR, if the offering of the respective BC (most probably: transaction-) services from abroad is envisaged to address data subjects in the EU,[7] or should include behaviour-tracking elements.

III. Personal Data on the Blockchain – Within or Outside the GDPR?

For the application of EU data protection law, data stored on the Blockchain must furthermore be *personal* data under the definition of Article 4(1) GDPR, which requires the information in question to be related to an identified or identifiable natural person. With this approach being broadly in line with the former Article 2(a) Data Protection Directive 95/46 (DPD),[8] it can be envisaged that the discussions, under what circumstances a data subject is still identifiable or not, will continue under the new regime.[9] Under the DPD and GDPR alike, a data subject is considered identifiable if there are means reasonably likely to be used for identification by the data controller or any other person,[10] which requires an assessment of, inter alia, the involved costs, time and technological efforts. In the pending case *Breyer v Germany*, Advocate General Sánchez-Bordona opined that it would suffice for dynamic IP-addresses to be personal data if there is a known third party holding additional information necessary to identify an individual.[11] If, however, data were rendered *anonymous* in such a way that the data subject is no longer identifiable, the GDPR would not apply.[12]

Whether the use of BC must comply with the GDPR, will first and foremost depend on whether personal data is stored on BC under the respective business model. Most of currently discussed use cases involve transactions of all sorts, be it financial assets, property registers or smart contracts, which all usually involve specific information in relation to a specific person as a rightholder, owner or contract party. Albeit this information is normally encrypted and can only be accessed with the correct keys, encryption of the data as such – ie giving access only to authorised parties – will normally not take such data out of the scope of the GDPR and may even be required[13]. Even if personal information only entails reference ID numbers, such identifiers are typically unique to a specific person. While in all such cases

[4] art 3 para 1 GDPR. The specific legal form of establishment, ie a branch or subsidiary with a legal personality is no determining factor, see recital 22.

[5] art 3 para 1 GDPR.

[6] Tim Wybitul, 'EU-Datenschutz-Grundverordnung in der Praxis – Was ändert sich durch das neue Datenschutzrecht? ' (2016) Betriebsberater 1077, 1079.

[7] See recital 23 GDPR for determining factors.

[8] Gerald Spindler, 'Die neue EU-Datenschutz-Grundverordnung' (2016) Der Betrieb 937.

[9] Especially in Germany, the threshold for data subjects not to be 'identifiable' anymore was previously understood as quite high by some commentators, namely where it is absolutely impossible to identify the data subject, see on that discussion Thilo Weichert, in Däubler et al (eds.), *BDSG* (5th ed, Bund Verlag 2016) §3 no 13.

[10] Recital 26 DPD; recital 26 GDPR.

[11] Case C-582/14 *Breyer v Germany,* Opinion of AG Manuel Campus Sanchez-Bordona (12 May 2016), no 52–78.

[12] Recital 26 GDPR; Spindler (n 8) 938.

[13] See only art 32 GDPR, with its risk-based approach setting out encryption measures as a factor to determine the adequate level of data security, or art 6(4)(e) GDPR, allowing changes in the purpose of processing under certain circumstances.

additional information may be necessary to attribute information to the data subject, such information would be merely pseudonymised and count as personal information.[14] A connection between pseudonymised data and the data subject will usually (and necessarily) arise in BC transactions effected for off-chain goods, eg conversion into real money payments, purchase of goods or services, registration data, where the transaction parties must be known.[15] Against that background, there is a strong case for arguing that individual-related information on BC is personal data.[16]

IV. Data Controllers in Multi-Node Systems – De-Centralisation as Accountability Gap?

Distributed ledgers pose a challenge for regulatory approaches that hinge on central intermediaries. Just as financial regulation struggles with handling virtual currencies[17] the EU data protection regime also relies on the notion of a data controller as central intermediary who is responsible for compliance with data protection legislation and implementation of privacy protective measures. In its absence, the GDPR's data protection framework, including the individuals' rights to data deletion, access and portability, security breach notifications, system development and setup rules, and especially enforced sanctions would practically lose effectiveness.

In that regard, the distinction between 'private BC' and 'public BC' becomes crucial. While both cases are technically distributed ledgers, the 'public BC' technology was initially developed as an open peer-to-peer-system for everyone to participate and to effect 'trusted' transactions with unknown counterparties. Yet, the use of BC is also tested in closed groups of 'trusted' entities, eg a consortium of banks, which do not only benefit from higher security, but may also see a chance of setting system standards for the years to come. In the latter case of 'private BC', one can easily imagine regulators to focus on either a technical system operator (if any, eg a joint venture setup) or consider the group of participating entities as joint controllers.

In a public BC, however, every node of the system will process all BC data simultaneously. To take an extreme view, if the notion of data controller implies any actual control over the information, either no node would qualify as such, as there is no *individual* control over the distributed BC, or, every node where BC copies are *technically* processed. Both outcomes would hardly bring meaningful results; even a concept of joint control responsibilities would remain opaque and fail to meet the standards of Article 26(1) GDPR, which require transparent and clear allocation of responsibilities.[18] Which stance regulators would take is hard to predict, but one can imagine that any 'gatekeepers' to BC may then find themselves as focal points of regulation.

V. Privacy by Design and Blockchain Core Features – Implementation or Clash of Principles?

While BC may ensure some privacy protection by applying advanced pseudonimisation and encryption techniques, other core features, like the absence of central intermediaries or persistent, distributed transaction storage, pursue different policy objectives such as system security, trusted transactions and intermediary independence. The fact that this (current) architecture runs counter to data minimisation, storage limitations and a clearly determined data controller, may raise the question whether it is in line with what is termed 'Privacy by Design' (PbD).

[14] In some discussed use cases, information is not stored directly on the BC, but off-chain-storage on a central server is in discussion for sensitive information to which BC only provides a pointer, without which such data could not be accessed, see IV. below. Here it is the combination of the BC pointer and server storage that give rise to dta being qualified as personal data.

[15] In pure on-chain-transaction models, however, encryption might make identification difficult. Without off-chain transactions, it is conceivable that encryption, which would make it impossible to attribute data to any individual, would not count as personal data, cf Steffen Kroschwald, 'Verschlüsseltes Cloud Computing - Auswirkung der Kryptografie auf den Personenbezug in der Cloud' (2014) ZD 75, for cases of fully-encrypted cloud computing.

[16] For pseudonymised data as personal data cf Spindler (n 8) 938; for technical details on pseudonymisation of Bitcoin-transaction data cf Franziska Boehm and Paulina Pesch, 'Bitcoins: Rechtliche Herausforderungen einer virtuellen Währung, Eine erste juristische Einordnung' (2014) Multimedia und Recht 75.

[17] See only Sarah Gruber, 'Trust, Identity, and Disclosure: Are Bitcoin Exchanges the next virtual havens for money laundering and tax evasion?' (2013) 32 Quinnipiac L Rev 135, 194; Gerald Spindler and Martin Bille, 'Rechtsprobleme von Bitcoins als virtuelle Währung' (2014) Wertpapiermitteilungen 1357, 1363. Among the challenges regulators face are not only anti-money-laundering rules and know-your-customer principles, but also issues of exchange controls and monetary policy. The trend towards disintermediation also runs counter to regulatory shifts towards central counterparties and central clearing mechanisms.

[18] See recital 79 GDPR.

This concept, which originated in the 1990s as a set of organisational and technological guiding principles,[19] is now explicitly enshrined in Article 25 GDPR. The idea behind PbD is to address privacy risks not only as a legal restriction for processing personal data, but to meet privacy concerns already at the early stage of IT architecture design.[20] Hence, Article 25 GDPR requires data controllers, at both the time of determining the means for processing and the processing itself, to implement appropriate technical and organisational measures in line with data-protection principles (eg data minimisation, pseudonimisation).

Whether the aforementioned BC features are incompatible with Article 25 GDPR is doubtful. From a legal perspective that sets out the PbD framework, firstly, views already diverge on whether Article 25 GDPR really brings additional 'hard' obligations[21] or whether it just restates the general obligation of compliance with GDPR principles.[22] Secondly, even if the incorporation of PbD as a legal principle stresses that technological solutions will matter for assessing GDPR compliance, Article 25 GDPR does not set out absolute requirements. Implementing PbD will take account of state-of-the-art technology, implementation costs, nature, scope, context and purposes of data processing as well as the likelihood of privacy risks. It does not strictly rule out that a balance may be struck between legitimate policy objectives. This relative view is, thirdly, even strengthened if such objectives can be enshrined in fundamental rights[23].

An examination of technical BC features within that PbD framework also mandates a differentiated view. At the current speed of technological development, an exhaustive analysis of distributed ledger technology in general would hardly bring certain results, as they will depend on the individual business model and technological implementation in any individual case. But in principle, BC embodies an inherent tension: At one hand, some of its features like perpetual distributed storage and the lack of central entities (in public BC models) could be seen as not completely in line with data minimisation and accountability. In respect of the latter, it could therefore make a difference whether public or private BC models are used. At the other hand, a strong BC encryption and data security would be in line with PbD. Against this background, an approach to mitigate this tension and to tip BC more towards the safe zones of data protection may lie in implementing additional PbD compatible technology. This may include eliminating ways that would allow tracking back pseudonyms to individual users, or adding 'noise' to BC data so that transactions are mixed up.[24] Another promising idea is to combine on-chain and off-chain storage for more sensible data and using BC only as a 'pointer' to centrally stored data, as it is the case in the MIT 'ENIGMA' project, which uses modified distributed hashtables for storing secret-shared data in combination with an external BC for network, access and identity control.[25] If necessary in sensible cases, even the use of private BC models could help. In any case, the BC ecosystem is evolving fast and some new models to handle privacy concerns can be expected soon. Yet, the 'benchmark' for PdB implementation may also rise, as this will be the then state-of-the-art technology.[26]

[19] It has first been proposed by Ontario's Information and Privacy Commissioner, see <https://www.ipc.on.ca/english/privacy/introduction-to-pbd/> accessed 13 August 2016.

[20] It can be assumed that art 25 GDPR does not introduce new *standards* of data protection, but expands compliance to the early phase of selecting the means of processing, ie product design and coding. On PbD see Article 29 Working Party, 'The Future of Privacy' (1 December 2009) Working Paper 168 (02356/09/EN), no 41-56; Dag Wiese Schartum, 'Making privacy by design operative' (Summer 2016) 24 (2) Int J Law Info Tech 151-175; for specific guidance on how to translate this policy into process design see ENISA, 'Privacy and Data Protection by Design – from policy to engineering' (2014) and 'Privacy by design in big data – An overview of privacy enhancing technologies in the era of big data analytics' (2015), both available at https://www.enisa.europa.eu/topics/data-protection/privacy-by-design accessed 13 August 2016.

[21] Cf Sibylle Gierschmann, 'Was „bringt" deutschen Unternehmen die DS-GVO? Mehr Pflichten, aber die Rechtsunsicherheit bleibt' (2016) ZD 51, 53.

[22] Cf Niko Härting, 'Art. 23 Abs. 1 DS-GVO (Privacy by Design): Cupcake ohne Rezept' (2015) Privacy in Germany 193.

[23] For balancing the right to data protection against other fundamental rights see Case C-70/10 *Scarlet Extended* (CJEU, 24 November 2011) ECLI:EU:C:2011:771, no 50 seq. It is worth noting that the German Constitutional Court (judgment of 27 February 2008, 1 BvR 370/07 and 1 BvR 595/07) has created a constitutional right in the confidentiality and integrity of information technology systems, which in the case of BC might exceptionally be balanced against privacy, if acknowledged on the EU level.

[24] See for example Cynthia Dwork, 'Differential Privacy: A Cryptographic Approach to Private Data Analysis' in Lane et al, *Privacy, Big Data and the Public Good* (Cambridge University Press 2014).

[25] See for example the MIT 'ENIGMA' project at Guy Zyskind, Oz Nathan and Alex 'Sandy' Pentland, 'Enigma: Decentralized Computation Platform with Guaranteed Privacy' http://enigma.media.mit.edu/enigma_full.pdf accessed 13 August 2016, which uses modified distributed hashtables for storing secret-shared data in combination with an external BC for network, access and identity control.

[26] As in general for implementation measures under the GDPR, see recital 78.

VI. The Right to be Forgotten and Blockchain Persistence

Finally, BC raises the question how to implement a Right to be Forgotten (RtbF), as recognized by the CJEU in the *Google Spain* case.[27] Article 17 GDPR strengthens that right to claim erasure of personal data if, eg, these data are no longer necessary for the purposes for which they were collected or processed, or if the data subject withdraws its consent and there is no other legal ground for processing.

Distributed ledger technology poses two challenges here. At first, it would be even harder to identify a data controller in a public BC who could fulfil data subjects' claims. But secondly, and more importantly, the distributed, persistent BC architecture may technologically even preclude a simple deletion upon request by the data subject, at least in public BC models: As set out above, one of the unique BC features is that data stored on-chain cannot be altered without acceptance of other nodes, which requires at least cooperation of more than half of all nodes for every transaction. All transactions ever effected on a BC are also perpetually stored on that BC so that a BC will be build up over time. If old transactions were to be removed retroactively, under current BC models, the majority of all P2P connected nodes would have to verify again the legitimacy of every effected transaction backwards, unbuild the entire BC block by block and then rebuild it afterwards, with every such transaction step to be distributed block-wise to all existing nodes. This would require extreme computational power and the cooperation of all nodes to de-construct a BC back to the point of change and to re-construct it afterwards. As the BC would be completely 'blocked' before it can even resume its function of adding new transactions,[28] it appears at the moment almost unfeasible from an operational and technical perspective to change BC content subsequently in practical operation. This applies in principle also to private BC models.

Yet, for some cases of deletions, it is conceivable that the wording of the GDPR may allow to take into account the functioning principle of BC. Under Article 17(1)(a) it might be argued that the personal data are still necessary for the processing purpose, as BC by design requires a persistent and continuously written chain. For Article 17(1)(b) it remains to be seen whether the core functioning principle of a technology could represent another 'legal ground for the processing' consisting, eg, in the performance of a contract or legitimate interests of the controller.[29] When balancing the interests of the controller and the data subject,[30] the aforementioned technological solutions may be of additional help, if full deletion is not possible.[31]

VII. Summary

BC is a groundbreaking technology and comes as a test for the technological neutrality of the GDPR. It remains to be seen whether EU data protection laws can embrace this development (and even steer its direction) in a meaningful way. The worst result, which would run counter to the intention of the legislator, would certainly be a formalistic approach that regards BC incompatible with GDPR. The strength of BC is creating trust in the authenticity of information and the safety of transactions. These objectives should be balanced with privacy concerns. While PbD can be achieved by mitigation measures, the lack of a central intermediary will probably pose a bigger challenge for regulators.

[27] Case C-131/12 *Google Spain und Google* (CJEU, 13 May 2014) ECLI:EU:C:2014:317, holding for art 12 DPD that, under certain conditions, individuals have the right to ask search engines to remove search results linking to pages containing personal information.

[28] Jeni Tennison, 'What is the impact of blockchains on privacy?' (*Open Data Institute*, 12 November 2015) <https://theodi.org/blog/impact-of-blockchains-on-privacy> accessed 13 August 2016.

[29] art 6(1)(b) and (f) GDPR.

[30] According to recital 69 GDPR, the controller must demonstrate a compelling legitimate interest that overrides the fundamental right to privacy of the data subject. It can be envisaged that this will require again a balancing of policy objectives and fundamental rights, see (n 23).

[31] For a pragmatic differentiating approach to the right to be forgotten see also Kieron O'Hara and Nigel Shadbolt, 'The Right to be Forgotten: Its Potential Role in a Coherent Privacy Regime' (2015) 3 EDPL 178.