Hogan Lovells

A guide to blockchain and data protection

November 2018

Contents

Authors

Winston Maxwell
Partner, Paris
+33 1 53 67 48 47
winston.maxwell@hoganlovells.com

John Salmon
Partner, London
+44 20 7296 5071
john.salmon@hoganlovells.com

Introduction and executive summary

Since publishing our original guide to blockchain and data protection in September 2017 there has been a great deal of further commentary, some of which suggests that there is an inherent incompatibility between blockchain and data protection law. In our new data protection compatibility section we will put forward our view on those comments.

A good place to start is to look at lessons learnt in the development of cloud computing and how these apply to blockchain projects. In particular, as in cloud computing, there is no one-size-fits-all solution for blockchain, given the huge diversity of architectures and use cases.

The major difference between blockchain and most cloud computing environments is that blockchain systems do not rely on a single provider of storage or computing resources. Each user of the blockchain uses his or her computing resources, on a peer-to-peer basis. Moreover, each user has a complete copy of the distributed ledger on his or her own computer. Consequently, the user of a blockchain system may at the same time be data controller for the data that he or she uploads onto the blockchain, and data processor by virtue of storing the full copy of the blockchain on his or her own computer.

Our guide assumes some level of knowledge about blockchain principles but little knowledge of data protection. We address the key data protection questions that will arise in any blockchain project. These include:

- Does the blockchain process personal data?
- Is a hash personal data or anonymised data?
- What about a public key?
- Who is the data controller and the data processor in a blockchain context?
- What is the applicable law?

The answers to these questions may lead to the conclusion that a given blockchain project's nexus to personal data is so remote that only minimal data governance mechanisms are required. By contrast, some projects will involve high-risk data processing, requiring a full-blown data protection impact assessment.

Data protection basics

GDPR

means the EU General Data Protection Regulation, a new regulation that came into effect on 25 May 2018, replacing the Data Protection Directive (95/46/EC). It is directly applicable in all Member States and the government has confirmed its intention to bring the GDPR into UK law notwithstanding the UK's decision to leave the EU. The GDPR applies only in respect of personal data (as opposed to data generally).

Personal data

means any information relating directly or indirectly to a 'living natural person', whether it actually identifies them or makes them identifiable.

Processing

means any operation or set of operations performed upon personal data, for example, collection, recording, organisation, structuring, storage, adaptation and alteration.

Data controller

means someone who determines the purposes for which and the manner in which any personal data is processed, whereas a data processor is someone who processes personal data on behalf of a data controller. In other words, the data controller determines how and why personal data is processed, and the data processor carries out processing according to the data controller's instructions.

Article 29 Working Party; EDPB

are the names for respectively the Data Protection Working Party established by Article 29 of Directive 95/46/EC and the European Data Protection Board created by the GDPR. The Working Party provided the European Commission with independent advice on data protection matters and helped in the development of harmonised policies for data protection in EU Member States. The EDPB has an expanded role under the GDPR, particularly to ensure coordination and consistency in applying the GDPR principles to cross-border processing.

> The major difference between blockchain and most cloud computing environments is that blockchain systems do not rely on a single provider of storage or computing resources.

What is a blockchain?

A distributed ledger is a replicated, shared, and synchronised digital data structure maintained by consensus algorithm and spread across multiple sites, countries, and/or institutions.

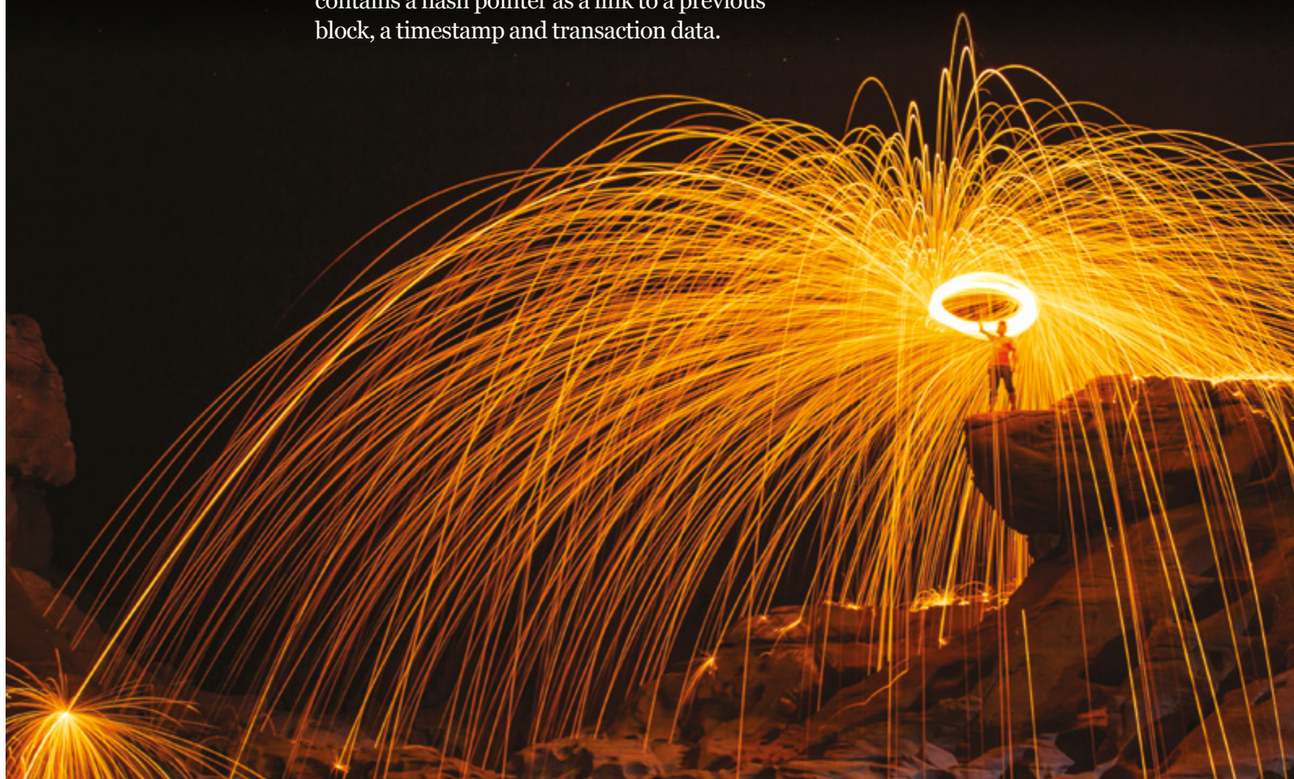
Blockchain is a type of distributed ledger, comprised of digitally recorded data in packages called blocks which are linked together in chronological order in a manner that makes the data very difficult to alter once recorded, without the alteration of all subsequent blocks and a majority of the network colluding together.

Each node on the network (generally) contains a complete copy of the entire ledger, from the first block created—the genesis block—to the most recent one. Each block contains a hash pointer as a link to a previous block, a timestamp and transaction data.

What is a Smart Contract?

Smart contracts use blockchain technology. The term is used to describe computer program code, maintained on the various "nodes" constituting a blockchain network that is capable of facilitating, executing, and enforcing the negotiation or performance of an agreement upon the occurrence of pre-defined conditions.

The smart contract code executes on each node and the resulting output is stored on the blockchain. Where "tokens" of value are involved, the smart contract code can also automatically transfer these tokens (and underlying value), thus effectively enforcing the outcome of the smart contract code.



Do blockchains process personal data?

‘Personal data’ is any information relating directly or indirectly to a ‘living natural person’, whether it actually identifies them or makes them identifiable. To determine whether data protection rules apply, we need to assess whether personal data is being processed when blockchain technology is used.

The nature of the public blockchain means that every transaction taking place will be published and linked to a published public key that represents a particular user. That key is encrypted so that no-one who views the blockchain would be able to directly identify the individual or corporate entity that represents the user.

However, the re-use of the public key enables individuals to be singled out by reference to their public key, even if they cannot be directly identified. Indeed the very purpose of the public key is to single out the authors of a given transaction, to ensure that transactions are attributed to the correct people. The public key, when associated with an individual, will likely qualify as personal data for the purposes of European data protection legislation. Some newer blockchain technologies permit the public key not to be published, which may alter the analysis.

When the public key is visible, it could be possible to attain information that enables an individual to be identified, either because it is held by the service provider or because someone is able to connect a public key to an individual or organisation, (for example, through their IP address or its connection with a website). At that point, all transactions that the relevant individual has made are publicly available.

In 2014, the Article 29 Working Party, provided guidance on the difference between pseudonymised and anonymised data in its Opinion 05/2014 (WP 216). This distinction is important in relation to blockchain as data protection rules do not apply to anonymised data, as such data cannot be traced back to a living individual. However, the threshold for data to qualify as anonymised is very high.

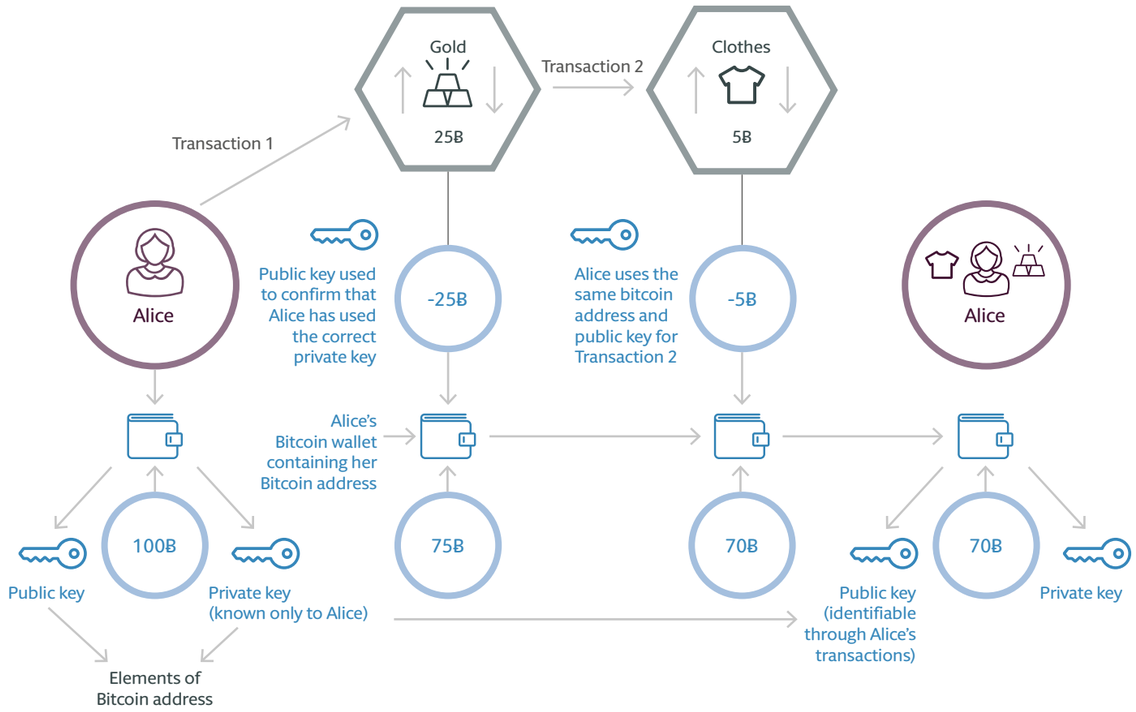
The guidance states that ‘anonymisation results from processing personal data in order to irreversibly prevent identification.’ Data controllers must have regard to all means likely reasonably to be used for identification (either by the controller or any third party). Because hashing permits records to be linked, hashing will generally be considered a pseudonymisation technique, not an anonymisation technique. This high standard continues to apply under the European General Data Protection Regulation 2016/679 (GDPR).

Encrypted personal data can often still be traced back to a person if enough effort is put into it by experts or someone holds the key to decryption. Therefore, encrypted data will often qualify as personal data and not as anonymous data. This means that in most instances the privacy rules will be applicable to at least some of the data involved in blockchain systems.

> “The public key, when associated with an individual, will likely qualify as personal data.”

Public key/Private key on a bitcoin transaction

Ƀ = Bitcoin



> “Data protection rules do not apply to anonymised data, as such data cannot be traced back to a living individual. However, the threshold for data to qualify as anonymised is very high.”

Case law on the concept of personal data

The Court of Justice of the European Union (CJEU) issued its final judgment in Case C-582/14 Patrick Breyer v Bundesrepublik Deutschland on 19 October 2016 relating to dynamic IP addresses.

The court's assessment of what constitutes 'personal data' in this judgment will have a general impact on how to define 'personal data' in a blockchain environment. The GDPR has not changed the definition of personal data, so the conclusions in the Breyer case continue to apply in the context of the GDPR.

The CJEU ruled that dynamic IP addresses (temporary IP addresses assigned to a computing device when it is connected to a network) may constitute 'personal data' even where only a third party (in this case an internet service provider) has the additional data necessary to identify the individual – but only under certain circumstances. The possibility of combining the data with this additional data must constitute a "means likely reasonably to be used to identify" the individual (the court assumed such means for Germany).

Patrick Breyer, a German national, took legal action against the Federal Republic of Germany as the operator of publicly accessible websites on which German public institutions supply topical information. He sought, based on data protection law, a prohibitory injunction against the Federal Republic of Germany, as the website-operator, because it stores IP addresses of visitors to their websites for cyber security reasons.

The German Federal Court of Justice referred the case to the CJEU asking:

a) whether dynamic IP addresses of website visitors constitute personal data for website operators; and

b) whether a specific data protection provision of the German Telemedia Act, that basically precludes a justification based on legitimate interests (Article 7(f) of the Directive), is in line with EU-law.

The CJEU decided that dynamic IP addresses collected by an online media service provider only constitute personal data if the possibility to combine the address with data necessary to identify the user's of a website held by a third party (i.e. the user's internet service provider) constitutes a means "likely reasonably to be used to identify" the individual or by a third party.

The court emphasised, in accordance with the opinion of the Advocate General, that this would not be the case:

"if the identification of the data subject was prohibited by law or 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 CJEU therefore assumed, subject to the final assessment of the referring German Federal Court of Justice, that:

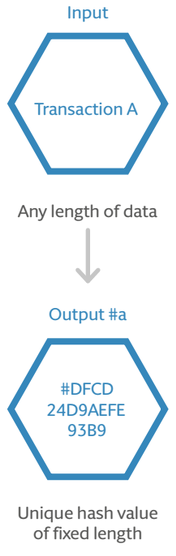
"the online media services provider has the means which may likely reasonably be used in order to identify the data subject, with the assistance of other persons, namely the competent authority and the internet service provider, on the basis of the IP addresses stored."

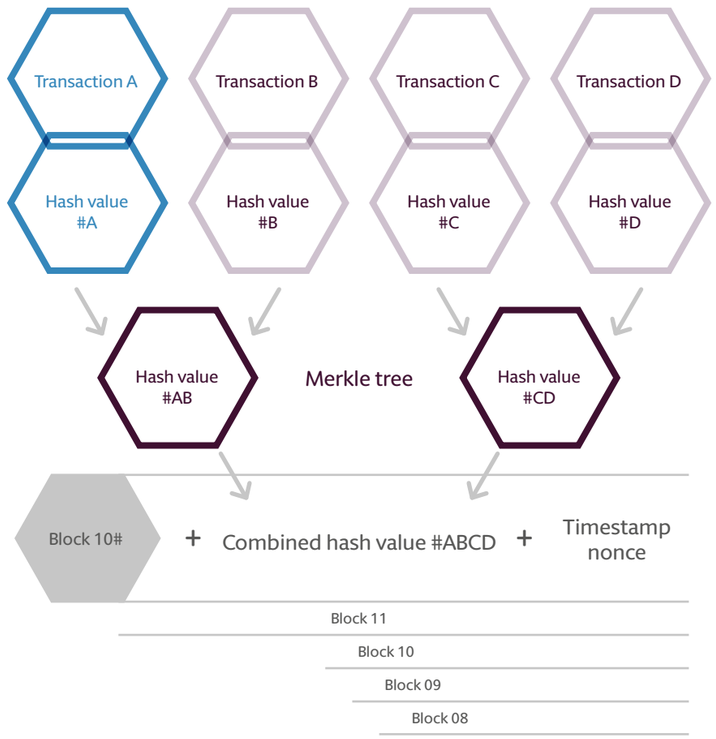
Hashing technology

> "Article 29 Working Party considers in its Opinion 05/2014 that hashing is a technique of pseudonymisation, not anonymisation."

Blockchain technology relies on hashing, which consists of generating a code of a fixed length for a given piece of digital information, regardless of its length. Hashing is important because it permits someone to verify, by recalculating the hash, that a given piece of digital information is identical to the digital information that was originally hashed. This permits document authentication – proof that a given document is the same one as the one that was originally hashed. This is an important feature of many blockchain systems.

A hash cannot be reverse-engineered to discover the original document. The process only works in one direction, from the original document to the hash. Yet in spite of this, the Article 29 Working Party considers in its Opinion 05/2014 that hashing is a technique of pseudonymisation, not anonymisation. According to the Article 29 Working Party, it is sufficient for a hash to permit records to be linked – the working group speaks of "linkability" – for a piece of information to constitute personal data. Consequently a hash that represents a person's ID card or medical record would likely be considered personal data even though the hash itself is impossible to reverse engineer into the original personal information. By contrast, a hash that represents a bill of lading would not be considered personal data, but for reasons linked to the bill of lading, not to the hash, as the bill of lading does not contain personal data.





Who is the data controller?

> More than one party may qualify as controller for one category of processing.

Usually when looking at data protection compliance issues, the first step is to identify the roles of the different parties involved. We need to ask:

a) who are the data controllers (those who determine the purposes and manner of processing, and have primary legal responsibility for data protection compliance)?

b) who are the data processors (those who process on behalf of the data controllers)?

This is challenging in a distributed ledger scenario.

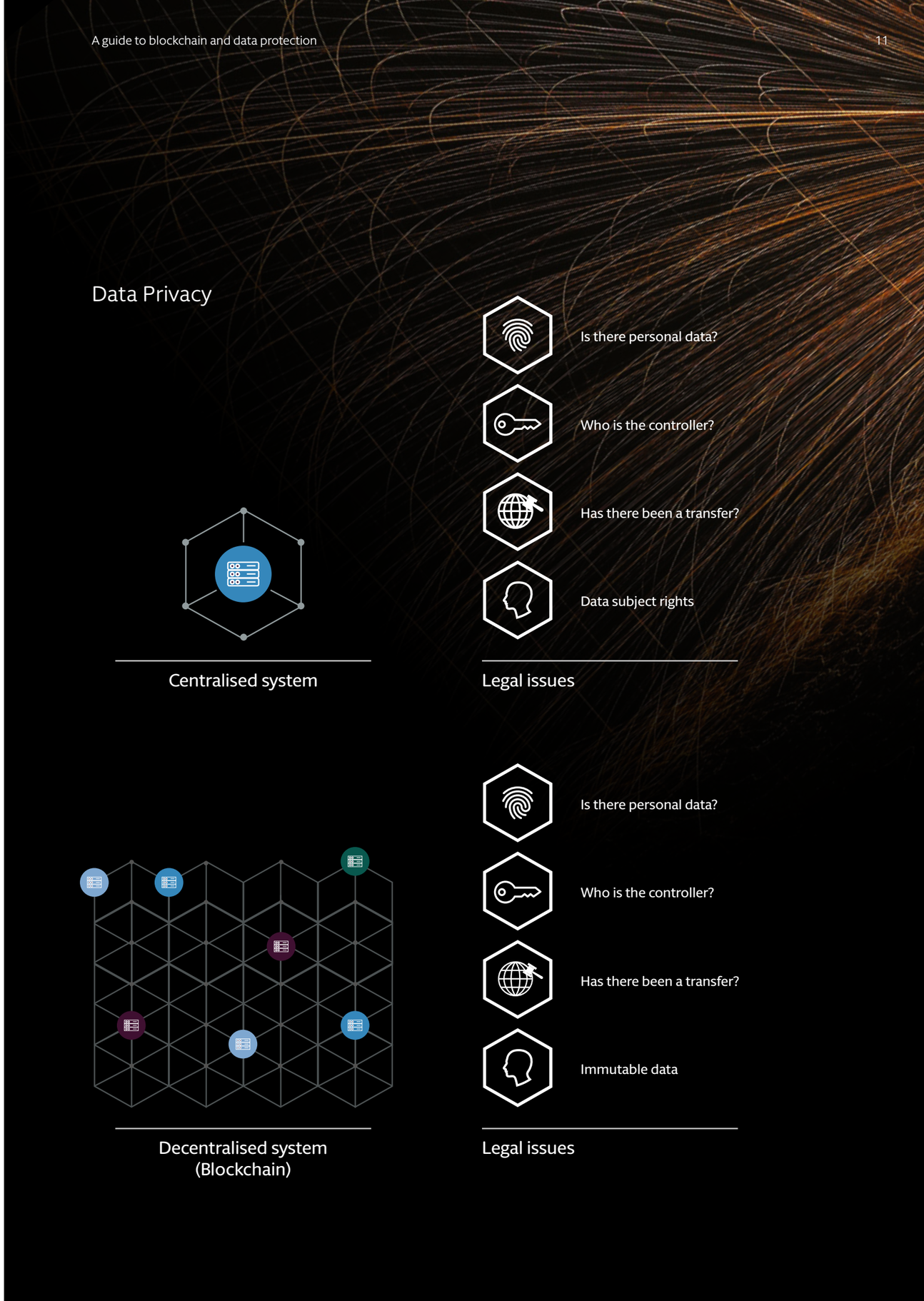
More than one party may qualify as controller for one category of processing. As outlined above, a data controller determines the purposes for which and the manner in which personal data is processed. This means that a data controller exercises overall control over the 'why' and the 'how' of a data processing activity, for example, the purpose or purposes the data are to be used for, whether to disclose the data, and if so to whom. A data processor processes personal data on the data controller's instructions, on behalf of the data controller. Whether a party qualifies as a controller or a processor will therefore depend on the degree of independence that each party has in determining how and in what manner the data is processed, as well as the degree of control over the content of the personal data.

Both data controllers and data processors have explicit but divergent obligations under the GDPR. It is therefore important to determine whether a party qualifies as a controller or a processor in relation to the processing of personal data in a blockchain network. This determination is not straightforward in a blockchain network as there are different types of blockchain systems which operate in different ways and which contain different types of participants carrying out different activities.

A blockchain system may be compared to a decentralised cloud computing system, whereby the operator of the cloud system is the data processor, and those uploading data to the cloud are the data controllers. However, for many blockchain systems, there is no central operator or administrator of the system. The system is operated by all its users in a peer-to-peer network environment. It is therefore necessary to consider to what extent the different participants in the blockchain network are controllers based on their respective activities.

The analysis of controllership will need to be carefully assessed for each blockchain system on its own merits. This assessment will need to differentiate between the participants that determine the purpose of data processing at the application layer as opposed to the processing at the infrastructure layer. Participants who submit the personal data to the blockchain platform are more likely to be considered controllers as they determine the purpose (to execute the transaction) and the technical and organisational details of the processing at the application layer. Whereas, nodes and miners who merely process data on behalf of users at the infrastructure layer, arguably are processors rather than controllers as they simply facilitate the running of the network.

The nodes in a blockchain system might be compared to autonomous systems on the Internet. Each autonomous system receives packets and routes them autonomously to another node and the process repeats itself until the packets reach their destination. How autonomous systems route packets is largely outside the control of users. The kind of processing that blockchain nodes perform is arguably similar. The only purpose of the nodes is to ensure the integrity of the blockchain and validate the addition of supplemental blocks.



Jurisdiction and applicable law

Blockchain systems are usually run on nodes located in different countries. The transnational nature of these systems therefore triggers jurisdictional issues.

The jurisdictional issues have been further compounded by the expansion in geographical scope of the European privacy rules for controllers and processors without an establishment in the EU. The GDPR now applies to a controller or processor:

"not established in the Union, where the processing activities are related to:

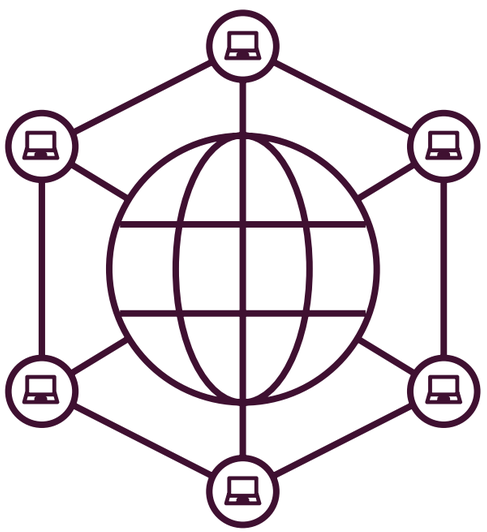
(a) The offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union; or

(b) The monitoring of their behaviour as far as their behaviour takes place within the Union."

A single blockchain system may involve multiple data controllers located around the world, some of whom have no establishment in the EU and do not target EU residents. In a cross-border decentralised blockchain environment, applicable law will likely have to be analysed on a transaction by transaction basis.

Since data protection choice of law rules are different from contract choice of law rules, the data protection law applicable to a transaction may not correspond to the contractual law. Unlike contract law, data protection law cannot be chosen by the parties. The applicable law depends on factors listed in Article 3 of the GDPR.

Given the cross-border nature of blockchain, and the GDPR's broad territorial reach, European data protection rules are likely to apply to many blockchain-based transactions that have little or no connection to Europe.



> "Applicable law and jurisdiction are complicated by the fact that a single blockchain system may involve multiple data controllers located around the world"

> "In a cross-border decentralised blockchain environment, applicable law will likely have to be analysed on a transaction by transaction basis."

> European data protection rules are likely to apply to many blockchain-based transactions that have little or no connection to Europe.

Increased enforcement

Under the GDPR, increased enforcement – fines of up to EUR 20 million or 4% of the worldwide turnover of a company – means the importance of privacy compliance will only grow.

However, it will be difficult to apply the enforcement provisions of the GDPR to public blockchains which are not owned or controlled by any individual person or firm.

Data protection principles

As is the case in many cloud environments, administrators of blockchain will not necessarily know whether personal data is present on the blockchain, let alone whether the data is sensitive. As noted above, the blockchain will show hashes pointing to previous blocks, transaction data that may be encrypted and/or a hash pointing to data stored off the chain.

For example, the MedRec blockchain, a system developed for managing patient medical records that uses the Ethereum blockchain, allows management of sensitive data – patient medical records – but the records themselves continue to be stored in hospital databases, off the chain. In the case of MedRec, the system is designed for medical records, so the designers of the system will not only know that personal data is involved, but also that the data is sensitive.

In many cases, a generic blockchain will be used by participants to register many different kinds of documents and transactions, involving both non-personal data and personal data. Like a social network, a generic blockchain can host any kind of data uploaded by users.

Because of the great variety of uses, data and configurations, generic blockchains will not be able to build in privacy protections adapted to the kind of data processed. At best, governance rules can regulate users of the blockchain to respect data protection laws when they upload personal data onto the blockchain. For special-purpose blockchains such as the MedRec system, governance rules can be much more developed, for example by prohibiting users from uploading actual medical records to the blockchain itself.

Two main features of the blockchain are:

a) information transiting through the blockchain is visible to every node; and

b) information cannot be removed from the blockchain.

These features clearly conflict with the principle of data minimisation and the storage limitation. Indeed, making data visible to every node could be considered excessive while perpetual storage of the data on the blockchain is clearly difficult to reconcile with the storage limitation rules.

Key legal challenges
Privacy and data protection

Right to erasure

One of the design features of blockchain architecture is that transaction records cannot be changed or deleted after-the-fact. A subsequent transaction can always annul the first transaction, but the first transaction will remain in the chain.

The GDPR recognises a right to erasure. The broad principle underpinning this right is to enable an individual to request the deletion or removal of personal data where there is no compelling reason for its continued processing.

When does the right to erasure apply?

The right to erasure does not provide an absolute 'right to be forgotten'. Individuals have a right to have personal data erased and to prevent processing in specific circumstances:

a) Where the personal data is no longer necessary in relation to the purpose for which it was originally collected/processed.

b) When the individual withdraws consent.

c) When the individual objects to the processing and there is no overriding legitimate interest for continuing the processing.

d) The personal data was unlawfully processed (i.e. otherwise in breach of the GDPR).

e) The personal data has to be erased in order to comply with a legal obligation.

f) The personal data is processed in relation to the offer of information society services to a child.

Does erasure mean erasure?

What constitutes "erasure" is still open to debate. Some data protection authorities have found that irreversible encryption constitutes erasure. In a blockchain environment, erasure is technically impossible because the system is designed to prevent it. However, smart contracts will contain mechanisms governing access rights. Therefore the smart contract can be used to revoke all access rights, thereby making the content invisible to others, albeit not erased.

> "Perpetual storage of data is difficult to reconcile with storage limitation rules"

> "The right to erasure does not provide an absolute 'right to be forgotten'."

Variety of blockchain systems

There is no single model for blockchain systems. Unlike the Internet, blockchain has no single set of standards, meaning that the technology can be deployed in an almost infinite variety of configurations. Each project will therefore have to be analysed on its own distinct merits.

Private vs. Public blockchains

From a privacy perspective, it matters greatly whether the blockchain is generally accessible or only accessible to parties that are members of a closed group. For instance, this may influence the assessment of whether data is transferred to countries that do not ensure adequate protection.

On another level, it is possible that each party to the blockchain network only has "access" to part of the information stored via the blockchain. As each party has its own copy of the entire blockchain, restricted access is achieved via encryption. Depending on how this is given substance, it may help to ensure compliance with the relevant privacy requirements.

Similar to debates in the cloud industry, blockchain will raise the questions of whether making a copy of a hash in, for example, Singapore means that data has been "transferred" to Singapore for the purposes of data protection law. In some sense, data put on a public blockchain is similar to data posted to the public internet.

The reasoning in the CJEU's Bodil Lindvist case (C 101/01) may apply to the question of transfer. The CJEU held that it cannot be presumed that the word "transfer", which is not actually defined in the Directive, was intended to cover the loading by an individual of data onto an Internet page.

"Off-Chain"

There have recently been some experiments made on public blockchains by introducing "off-chain" mechanisms to store the confidential information separately on another system with access control restrictions. To protect data and manage storage on the blockchain, some solutions use only a hash of personally identifiable information (PII), which serves as a reference point and link to an off-chain PII database. Storing information "off-chain" provides privacy of the transaction details. The "off-chain" system can be set up to restrict access to the transaction details to authorised parties only.

However, storing information "off-chain" also negates a number of the advantages of using blockchain. The blockchain can no longer be a single, shared source of truth and in most cases both counterparties will be required to maintain their own records.

"Sidechains"

Unlike "off-chain", which generally stores the chosen information on a traditional network, but at the expense of the benefits of using a blockchain, a "sidechain" is a parallel blockchain. It sits alongside the primary blockchain, serving multiple users and generally persisting permanently. The degree of confidentiality and privacy provided for transactions that take place on sidechains depends on what technology the sidechain uses.

These sidechains are independent. If they fail or are hacked, they won't damage other chains. So damage will be limited within that chain. This has allowed people to use sidechains to experiment with pre-release versions of blockchain technologies and sidechains with different permissions to the primary blockchain.

Non-Permissioned vs. Permissioned Blockchains

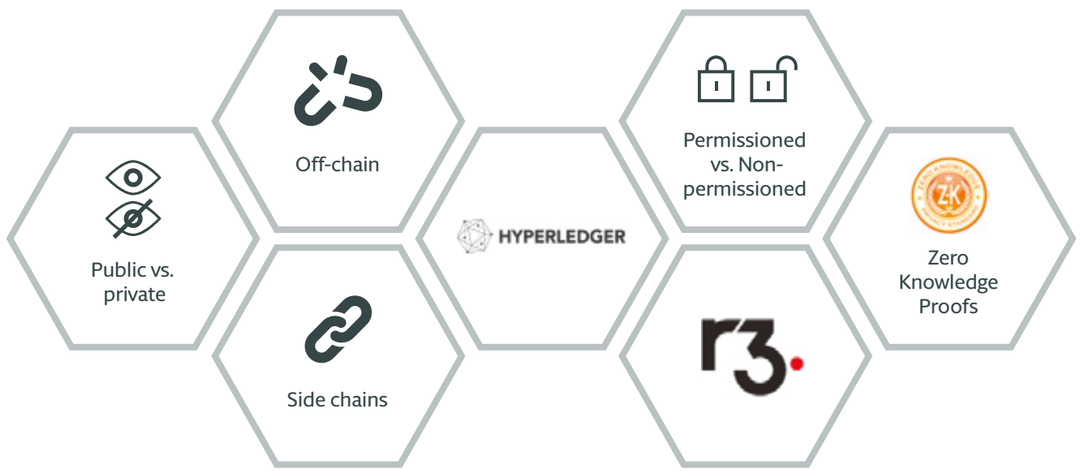
With non-permissioned blockchain applications, all parties are in principle free to add information to the blockchain. With permissioned blockchain, on the other hand, access is restricted. In this way, trusted intermediaries are reintroduced into the system, which impacts the allocation of control over it.

The party that determines the means and the purposes for the processing should ensure that the privacy rules are taken into account, meaning the choice between non-permissioned and permissioned control also influences which parties should comply with what privacy requirements.

Hyperledger

Hyperledger is a hub for open industrial blockchain development; it is not a company, a cryptocurrency, or a blockchain. Hyperledger provides technical knowledge, software frameworks and contacts to industries and developers. The platform aims to "create an enterprise-grade, open source distributed ledger framework and code base" as well as creating, promoting and maintaining an open infrastructure.

Hyperledger incubates and promotes a range of business blockchain technologies, including distributed ledger frameworks, smart contract engines, client libraries, graphical interfaces, utility libraries and sample applications. One of the distributed frameworks is called Hyperledger Fabric ("HLF"), which is an open-source project within the Hyperledger umbrella project. HLF is a modular, general-purpose, permissioned blockchain system, which can also be seen as a distributed operating system for permissioned blockchains. (Source: www.hyperledger.org)



R3

R3 is the largest consortium of global financial institutions working on developing commercial applications for the distributed ledger technology. R3 has its own proprietary ledger that can be used to develop applications, and it also supports an infrastructure network for financial services firms and technology companies wanting to build their own ledger-based applications and services.

The blockchain technology that R3 is currently developing is a distributed ledger platform designed specifically for financial services, called Corda. The Corda network is permissioned, with access controlled by a doorman. Communication between nodes is point-to-point, instead of relying on global broadcasts. Each network has a doorman service that enforces rules regarding the information that nodes must provide and the know-your-customer processes that they must complete before being admitted to the network.

> “Access to Corda network is controlled by a doorman.”

Zero Knowledge Proofs

A zero knowledge proof (“ZKP”) is a cryptographic technique which allows two parties (a prover and a verifier) to prove that a proposition is true, without revealing any information about that thing apart from it being true. A zk-SNARK (zero-knowledge Succinct Non-Interactive Arguments of Knowledge) is a ZKP that proves some computation fact about data without actually revealing the data. Zk-SNARKS are the underlying cryptographic tool used for verifying transactions in Zcash. This is done while still protecting users’ privacy.

Zcash can be described as an encrypted open, permissionless, replicated ledger. It is a cryptographic protocol for putting private data on a public blockchain. Zcash uses zk-SNARKS to encrypt all of the data and only gives decryption keys to authorised parties. Previously this could not be done on a public blockchain because if everything was encrypted it would prevent miners from checking to see if transactions were valid. However ZKPs made this possible by allowing the creator of a transaction to make a proof that the transaction is true without revealing the sender’s address, the receiver’s address and the transaction amount.

ZKPs and blockchains complement each other – a blockchain is used to make sure the entire network can agree on some state that may or may not be encrypted, whereas ZKPs allow you to be certain about some properties in that state.

> “ZKPs permit users to hide the sender’s address, the receiver’s address and the transaction amount.”

Analysis of the different systems

Understanding the data on a traditional (Bitcoin/Ethereum) blockchain.

On each block of the blockchain there are two types of data: 1) a header that includes a date stamp, the identity of the source of the data (an address), and the previous block's header hash, called 'the pointer'; and 2) the payload, which is the data to be stored.

The header is not encrypted, only the payload is normally encrypted. The hash in the header is of earlier blocks to create the immutable chain of blocks. The payload is generally a description of the document (metadata) and the hash representing the actual document.

A smart contract would operate as follows: when X wants to upload a new document description to the blockchain, the smart contract will create a transaction by combining a description of the document and its hash to form a payload and add a header. The header plus payload equals a transaction and a validated transaction equals a block. Upon validation of the block, the smart contract would then send the document itself to the Y database system for storage. We assume that the Y database is off the blockchain and has limited access through the use of passwords which can be time sensitive. The blockchain transaction will be proof that the document was uploaded at a given time, and anyone will be able to verify that the document held in the off-chain database is the same document as the one referred to in the blockchain transaction.

If using blockchain technology similar to Bitcoin or Ethereum, which are both public, open, transparent blockchains, where all transaction details are visible on the blockchain, i.e. you can see the public key, then we could single out an individual by their transactions, assuming that they use the same public key for each transaction.

In addition, after a public key and the associated transactions are singled out as relating to the same individual or entity, there is no way to 'erase' the information as this information is now part of the blockchain and public knowledge. With Bitcoin, the public key must be visible to avoid double spend issues and means that we are able to track transfers (ie. able to see bitcoins coming in and bitcoins going out).

Understanding the data on new blockchain technologies

As mentioned above, some technologies permit a greater degree of anonymity. Whether the degree of anonymity satisfies European standards under the CJEU's Breyer decision and the Article 29 Working Party Opinion is another matter. For example, Dash encrypts public keys, while new blockchain technologies using zero knowledge proofs can verify transactions without details of the transaction itself. However, the work involved in developing a proof is extensive and has significant computation costs. As a result, there are scalability challenges with these tools.

The alternative could be to use a new public key, with the same private key for each transaction. However, this must be done properly, and would only be possible for some public keys. This would involve a key issuing authority (centralised, although there may be a number of these) that generates a different key for each transaction from a core private key.

The mapping between the core key and transaction keys is never revealed to the other participants – they only see the individual transactions keys. This is the approach being developed and used by Hyperledger Fabric.

Blockchain data protection impact assessments

Under the GDPR a data protection impact assessment is required for processing, which is likely to present a high risk to the rights and freedoms of natural persons. Blockchain projects can be roughly divided into three categories:

- Specialised blockchain systems designed to process essentially non-personal data, such as bills of lading, letters of credit, or diamond certificates;
- Specialised blockchain systems designed to process personal data, such as proof of identification, or even sensitive personal data such as medical records;
- Non-specialised blockchain systems that can be used to process any form of data.

A data protection impact assessment is likely to be required for the second category of blockchain system, where processing personal data is the purpose of the system. In that case, regulators will expect the system to build in privacy protections, via data protection by design and default, to ensure that the system does not pose a risk to the rights and liberties of individuals. A data protection impact assessment will be required, particularly where the type of data involved is risky.

As noted above, determining which law applies will be challenging for systems that process data from several continents.

Non-specialised blockchains are likely to put the onus of compliance on the users themselves, through terms of use that:

a) prohibit posting of certain kinds of personal data; and

b) require users to have consent or another legal basis for processing.

The data protection impact assessment will need a robust technical analysis to show that the security of the system is at least as robust, if not more robust, than traditional cloud-based systems.

Commentary and Our Views

Since publishing the original version of this paper in September 2017, there has been considerable further commentary from academics, politicians and practitioners, some of which has proposed that there is inherent incompatibility of blockchain systems with data protection law. Our view is more optimistic.

Reflecting on our experience of applying privacy law to the internet over the last 20 years, we have largely seen a pragmatic approach taken to frictions between the law and technological innovation in the area of privacy law. For example, the European Court of Justice has historically tried to avoid interpretations of the 1995 Data Protection Directive that lead to irreconcilable conflicts between the Directive and new technologies. In the Bodil Lindqvist case[1], the Court found that the posting of information on the Internet could not be considered as a transfer of personal data to every country in the world from which the website was accessible, since that would lead to an unworkable result. Similarly, in the Google v. Spain case[2], the Court found that the operator of a search engine was the data controller but only for certain aspects of its activity and only within the scope of its role and functions: *"inasmuch as the data processing carried out in the context of the activity of a search engine can be distinguished from and is additional to that carried out by publishers of websites and affects the data subject's fundamental rights additionally, the operator of the search engine as the controller in respect of that processing must ensure,* **within the framework of its responsibilities, powers and capabilities**, *that that processing meets the requirements of Directive 95/46."*[3] Consequently, we think that it is likely that the European Court of Justice would seek to find a way of applying GDPR principles to the blockchain that would not put the two on a direct collision course.

In this section, we will look at some key issues that have been raised by academics and practitioners and put forward our views on those comments.

Controllership

Michèle Finck, 'Blockchains and Data Protection in the European Union'

Michèle Finck discusses the issue of controllership in her excellent paper and concludes that:

"Permissionless blockchains are distributed and decentralised peer-to-peer networks that everyone can participate in to interact with unknown or untrusted counterparties. In such a setting either no node qualifies as the data controller in the absence of independent determination of the means and purposes of processing, or, more likely, every node qualifies as a data controller."

We would argue that the conclusion reached by Michèle Finck is overly pessimistic. As we have explained above at p.11, the nodes in a blockchain system can be compared to autonomous systems on the Internet. Each autonomous system receives packets and routes them autonomously to another system and the process repeats itself until the packets reach their destination. The kind of processing that blockchain nodes perform is arguably similar. The only purpose of the nodes is to ensure the integrity of the blockchain and validate the addition of supplemental blocks. Therefore, we believe that the more pragmatic conclusion would be to find the data controller is the person or entity that determines the purpose of data processing at the application layer as opposed to the processing at the infrastructure layer, which is where the nodes operate.

J Bacon and others, 'Blockchain Demystified'

This issue of controllership in permissionless blockchains was also raised by J Bacon and others in their paper on blockchain. They form a similar view to our own and use an insightful analogy to demonstrate their finding that the more autonomy a participant has, the more likely it is that they will qualify as a controller:

"If nodes and miners only process data on behalf of users, they could arguably qualify as processors…If, however, nodes and miners take a more active role with regard to the transaction data, they may also be deemed to be controllers. In that case, nodes and miners could be compared to SWIFT, a service that processes personal data such as the names of the payer and payee. SWIFT presented itself as a processor, relaying messages on behalf of the financial institutions. However, the Article 29 Working Party determined that SWIFT should be considered a controller, since it acted with a significant level of autonomy in respect of the personal data it processed, including by developing, marketing and changing the services it offered, deciding to establish a data centre in the US and to disclose data to the US Treasury".[4]

We would conclude that controllership depends on the particular blockchain use case and the level of autonomy at the application level. For example, with permissioned blockchain applications, where access is restricted, trusted intermediaries are reintroduced into the system through a centralised party, making it much easier to identify the level of control. Whereas, with non-permissioned blockchain applications, all parties are in principle free to add information to the blockchain. So we would argue that participants who submit the personal data to the blockchain platform are more likely to be considered controllers as they determine the purpose (to execute the transaction) and the technical and organisational details of the processing at the application layer. On the other hand, nodes and miners simply facilitate the running of the network.

1 Case C-101/01 Bodil Lindqvist, 6 November 2003 http://curia.europa.eu/juris/document/document.jsf?docid=48382&doclang=en

2 Case C-131/12 Google Spain, 13 May 2014 https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62012CJ0131&from=IT

3 Ibid, paragraph 83 (emphasis added)

4 J Bacon and others, 'Blockchain Demystified' Queen Mary School of Law Legal Studies Research Paper No 268/2017

Data Protection Principles

Jan Philipp Albrecht

Jan Philipp Albrecht, a former member of the European Parliament, who played a key role in the finalisation of the GDPR, was quoted as saying that:

"Certain technologies will not be compatible with the GDPR if they don't provide for [the exercising of data subjects' rights] based on their architectural design. This does not mean that blockchain technology, in general, has to adapt to the GDPR, it just means that it probably can't be used to process personal data."[5]

This statement seems to overlook the different blockchain systems available and the technical solutions that are being designed to provide relief to the frictions between data subjects' rights and blockchain technology. We would argue that each blockchain system should be analysed on its own merits including an in depth analysis of what personal data is processed on the blockchain and who the data controller is. Once this tailored analysis has been completed you can then look at how data protection law will apply to that system's data protection design issues. There are already solutions available to accommodate data subjects' rights on blockchain systems. For example, the right to erasure does not provide an absolute right to be forgotten and what constitutes 'erasure' is still open to debate. Some data protection authorities have found that irreversible encryption constitutes erasure. So a smart contract could be used to revoke all access rights, making all content invisible to others.

Although the rights of data minimisation and storage limitation requirements are more difficult to accommodate within an immutable blockchain system, we disagree with the implication that blockchain technologies are inherently incompatible with the GDPR. Instead we are confident that technical solutions, if not already available, will be developed to address the compatibility issues between the GDPR and blockchain technology.

The European Union Blockchain Observatory and Forum

The European Union Blockchain Observatory and Forum ("EUBOF") has come to a similar conclusion as our own in its report[6], where it held that:

"GDPR compliance is not about the technology, it is about how the technology is used. Just like there is no GDPR-compliant Internet, or GDPR-compliant artificial intelligence algorithm, there is no such thing as a GDPR -compliant blockchain technology. There are only GDPR-compliant use cases and applications."

Until regulators settle these issues the EUBOF suggest that developers consider the following four principles:

1. *Start with the big picture: how is user value created, how is data used and do you really need blockchain?*
2. *Avoid storing personal data on a blockchain. Make full use of data obfuscation, encryption and aggregation techniques in order to anonymise data.*
3. *Collect personal data off-chain or, if the blockchain can't be avoided, on private, permissioned blockchain networks. Consider personal data carefully when connecting private blockchains with public ones.*
4. *Continue to innovate, and be as clear and transparent as possible with users."*

5 D Meyer, 'Blockchain Technology is on a Collision Course with EU Privacy Law' IAPP Privacy Advisor

6 The European Blockchain Observatory and Forum, 'Blockchain and the GDPR', 16 October 2018

The French Data Protection Authority (the CNIL) guidelines

The CNIL published guidelines[7] on how to make a GDPR-compliant blockchain

After receiving queries from actors (public and private) from the healthcare and financial services, the CNIL published guidelines on responsible use of blockchain and concrete solutions to use blockchain in conjunction with personal data:

- **Data controller:** for the CNIL, the participants, having writing rights and deciding to submit data for validation to the miners, should be regarded as controllers. Miners and individuals acting under the household exception (Article 2 of the GDPR) should not be controllers. In some cases, all the participants may have joint-controllership, but the CNIL recommends the creation of an entity regrouping them.
- **Processor:** the processors may be the smart contract developer or the miners validating the recording of personal data in a blockchain. The CNIL is currently conducting a study on public blockchain systems and encourages the development of solutions that provide a framework for contractual relationships between participants/controllers and miners.
- **Privacy by design:** The CNIL highly recommends conducting a data protection impact assessment in order to assess the necessity and proportionality of blockchain technology and the potential threats to personal data. The choice of whether to use a blockchain or what kind of blockchain to use can have a great impact on threats to fundamental rights caused by processing personal data (and the risk of infringing GDPR). Therefore, the CNIL asks controllers to choose carefully, and to favour blockchains based on permission, as they enable a better control on personal data, especially with regards to transfers outside the EU.
- **Data subjects' rights:** blockchains can be useful to provide evidence of consent or of data processing operations and also to allow certain rights to be exercised effectively, such as the right to be informed, access rights and data portability rights. However, there are some obvious obstacles regarding the exercise of other rights, such as the right to rectification, right to erasure, or right to object. If the data entered in the chain is an imprint from a hash key function or an encrypted code, the controller may, technically, make the data almost unreachable and, as a consequence, come close to GDPR requirements regarding full erasure of data. The same technique can be used for rectification rights, where the rectified data must be recorded in a new block of the chain and the old data must be made unreachable.

Conclusion

Commentary still seems divided between those who see blockchain as fundamentally incompatible with the GDPR, and those who see a middle ground world where GDPR and blockchain can coexist. We see many parallels between today's blockchain and the early days of the Internet and cloud computing. Both the Internet and cloud computing are based on decentralized processing, and both raised existential concerns regarding their compatibility with European data protection law. In the end, courts and regulators took a case-by-case approach, focusing more on the particular use case rather than the underlying technology used. We believe a similar approach will emerge for blockchain. The French data protection authority's initial guidelines seem to go in this direction. However, we can imagine at some point a national court referring a question to the CJEU regarding the role of blockchain nodes and how article 28 of the GDPR should apply in the event the nodes are considered processors.

7 Commission Nationale Informatique & Libertès, 'Premiers éléments d'analyse de la CNIL: Blockchain', 24 September 2018

Blockchain contacts

Gregory C. Lisa
Partner, Washington
+1 202 637 3647
gregory.lisa@hoganlovells.com

Christian Mammen
Partner, San Francisco
+1 415 374 2325
chris.mammen@hoganlovells.com

Winston Maxwell
Partner, Paris
+33 1 53 67 48 47
winston.maxwell@hoganlovells.com

Theodore Mlynar
Partner, New York
+1 212 918 3272
ted.mlynar@hoganlovells.com

Patrice Navarro
Counsel, Paris
+33 1 53 67 47 47
patrice.navarro@hoganlovells.com

Mark Parsons
Partner, Hong Kong
+852 2840 5033
mark.parsons@hoganlovells.com

John Salmon
Partner, London
+44 20 7296 5071
john.salmon@hoganlovells.com

Hogan Lovells Engage: blockchain tool

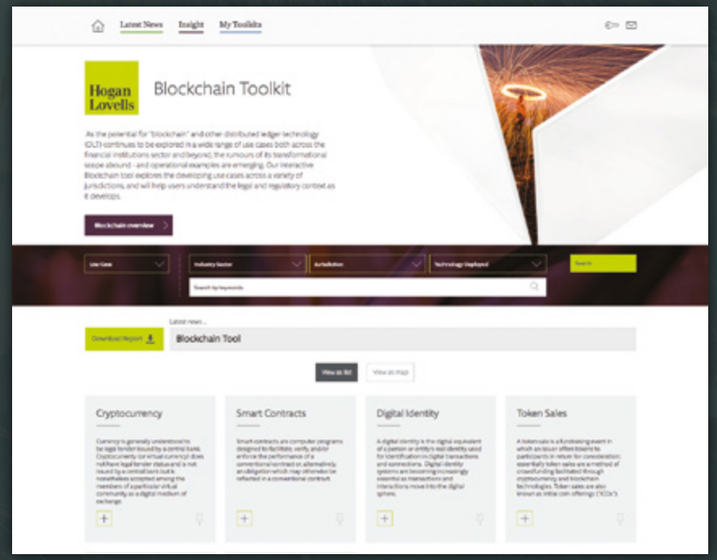
Take advantage of blockchain's huge potential and disruptive impact, while avoiding falling foul of ever-developing regulatory and legal requirements.

The Hogan Lovells Engage: Blockchain Toolkit lets you:

- investigate the different ways blockchain can be used
- see where the new technology is shaking up industries
- track unfolding legal and regulatory approaches across jurisdictions
- use interactive functionality to download reports and share information

Get started now by registering on:

hlengage.com/blockchain



Alicante
Amsterdam
Baltimore
Beijing
Birmingham
Boston
Brussels
Budapest
Colorado Springs
Denver
Dubai
Dusseldorf
Frankfurt
Hamburg
Hanoi
Ho Chi Minh City
Hong Kong
Houston
Jakarta
Johannesburg
London
Los Angeles
Louisville
Luxembourg
Madrid
Mexico City
Miami
Milan
Minneapolis
Monterrey
Moscow
Munich
New York
Northern Virginia
Paris
Perth
Philadelphia
Rio de Janeiro
Rome
San Francisco
São Paulo
Shanghai
Shanghai FTZ
Silicon Valley
Singapore
Sydney
Tokyo
Ulaanbaatar
Warsaw
Washington, D.C.
Zagreb

Our offices
Associated offices

www.hoganlovells.com