Part 2: Impacts on the Wider Landscape

Section 10

Data Protection

Section 10: Data Protection

PART A: Data Protection

Anne Rose and Jon Baines (Mishcon de Reya LLP)

Introduction

The EU GDPR became binding on 25 May 2018 and is based, in large part, and at least in big-picture, thematic terms, on the 1995 Data Protection Directive, which it replaced.[200] Since the 2020 guidance the UK has now left the EU and the UK GDPR applies in the UK, along with the Data Protection Act 2018 **(DPA 2018)**.

As a result of the UK's exit from the European Union, GDPR no longer directly applies. However, it was in large part retained, in slightly amended form, and became the "UK GDPR". It is to be noted, however, that since Brexit, UK governments have indicated a willingness more fully to reform the domestic data protection laws, and it will be important to monitor developments in this respect, not least because there have been suggestions that the definition of personal data itself might be altered, and this could have significant implications for the legal and regulatory aspects of blockchain.

UK GDPR's objective is essentially two-fold. On the one hand, it establishes a framework of fundamental rights in respect of the handling of personal data, with various measures based on the right to privacy (Article 8 of the Charter of Fundamental Rights), and on the other hand, it seeks to facilitate the free movement of personal data (see Article 1, UK GDPR).

Dual Regimes

In light of the amendments to data protection law since the 2020 guidance, if a controller/processor is carrying out processing activities or targeting/monitoring individuals in both the UK and the EU, there is now the added risk of dual enforcement by both the ICO and the EU Data Protection Authorities, as they will be subject to both UK and EU GDPR, since both have extra-territorial effect under Article 3 UK/EU GDPR. If activity is limited to the UK only, controllers/processors will now only be subject to UK GDPR.

For ease, this guidance refers to UK GDPR only and assumes that organisations are not subject to dual regimes. The 2020 guidance considered EU GDPR. For the avoidance of doubt, this guidance can also be applied to UK GDPR. The legal framework creates a number of obligations on data controllers, which are the entities determining the means and purposes of data processing. It also allocates a number of rights to data subjects – the natural persons to whom personal data relates – that can be enforced against data controllers. Blockchains, however, are distributed databases that seek to achieve decentralisation by replacing a unitary actor with many different players. The lack of consensus as to how (joint-) controllership ought to be defined, and how it impacts upon accepted (or, even contested) meanings within UK GDPR, hampers the allocation of responsibility and accountability. Moreover, UK GDPR is based on the assumption that data can be modified or erased where necessary to comply with legal requirements, such as Article 16 (personal data must be amended) and Article 17 (personal data must be erased). Blockchains, however, intentionally make the unilateral modification of data onerous (if not impossible) in order to ensure data integrity and to increase trust in the network.

For the 2020 guidance, the Group focused on the definition of "personal data" under EU GDPR and noted that depending on context, the same data point can be personal or non-personal and therefore subject to EU GDPR or not. In addition, the Group considered the impact of changes in technology that could increase the tension between blockchain and EU GDPR, as well as the possibility that blockchain could support EU GDPR. The Group did not go into detail on all the various issues, as these are discussed widely elsewhere.[201]

200 Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1

201 For example, Michèle Finck, Blockchain Regulation and Governance in Europe (Cambridge University Press 2018)

Experts and evidence

The Group heard from a number of experts for the First Guidance, including Peter Brown (Group Manager (Technology Policy), Technology Policy & Innovation Executive Directorate, ICO, UK); and Adi Ben-Ari, (Founder & CEO, Applied Blockchain).

Further, the Group liaised with Dr Michèle Finck, Senior Research Fellow at the Max Planck Institute for Innovation and Competition who has provided her perspective on certain elements in blockchain and the EU GDPR, which was produced at the request of the Panel for the Future of Science and Technology (STOA) and managed by the Scientific Foresight Unit, within the Directorate-General for Parliamentary Research Services (EPRS) of the Secretariat of the European Parliament.[202] Dr Finck has written widely on the points of tension between blockchain and EU GDPR – including questions of when and under which circumstances on-chain data qualifies as personal data.[203]

Anne Rose, Solicitor at the law firm, Mishcon de Reya LLP, has also considered the tensions at play between blockchain and EU GDPR in an interactive entertainment context.[204]

What is Personal Data?

Article 4(1) UK GDPR defines personal data as:

> *"**any information relating to an identified or identifiable natural person ('data subject');** an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (**bold** for emphasis).*

This underlines the fact that the concept of personal data is to be interpreted broadly, and could include anything from a picture to a post code or an IP address of a living individual.

It is also clear that an item of data may be personal data (for example, a name: Michael), or non-personal data (for example, information which was never personal in the first place: a pencil case), but there are also circumstances where it may be unclear or may even change (for example, an IP address or a hash where the linkage between the natural person and the hash has been removed – or, in simpler terms, Michael's pencil case). To assess whether data is personal, pseudonymous (personal data which can no longer be attributed to a specific data subject without the use of additional information) or anonymous (data which cannot be attributed to a specific data subject, including with the application of additional information) involves considering Article 4(5) UK GDPR and Recital 26 UK GDPR:

Article 4(5) UK GDPR (defining pseudonymous data) provides as follows:

> *"processing of **personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information**, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person" (emphasis added).*

202 Panel for the Future of Science and Technology, 'Blockchain and the General Data Protection Regulation: Can Distributed Ledgers be Squared with European Data Protection Law?' (European Parliamentary Research Service, July 2019) <https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634445/EPRS_STU(2019)634445_EN.pdf> Accessed 13 April 2020

203 See, for example, Michèle Finck, Blockchain Regulation and Governance in Europe (Cambridge University Press 2018)

204 Anne Rose, 'GDPR challenges for blockchain technology', (2019) 2 IELR 35

Recital 26, UK GDPR (which sets the background to Article 4(5)) states:

> *"…To determine whether a natural person is identifiable,* ***account should be taken of all the means reasonably likely to be used…****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..." (emphasis added).*

Recital 26 UK GDPR assumes a risk-based approach to assessing whether or not information is personal data, which the ICO has also adopted. The ICO notes that "the risk of re-identification through data linkage is essentially unpredictable because it can never be assessed with certainty what data is already available or what data may be released in the future".[205] In contrast, the Article 29 Working Party (now renamed as the European Data Protection Board, or EDPB) seems to suggest that a risk-based approach is not appropriate and that "anonymisation results [only] from processing personal data in order to irreversibly prevent identification".[206] This uncertain standard of identifiability and the elements which also need to be taken into account (costs, time required for identification and available technology) require further guidance from data protection authorities and bodies.

The Group considers this to be particularly important in times where personal data is dynamic and technical developments and advances make anonymisation (if defined as permanent erasure) near-impossible. Further, it is possible that anonymous data today becomes personal data in the future, once further data is generated or acquired allowing for identification by the controller or by another person. On the basis of this, it could result in the uncomfortable conclusion that personal data can only ever be pseudonymised, but never anonymised.[207]

This definitional issue needs to be constantly monitored by data controllers. As noted by the former Article 29 Working Party: "One relevant factor…for assessing 'all the means likely reasonably to be used' to identify the persons will in fact be the purpose pursued by the data controller in the data processing."[208] The French supervisory authority (the **CNIL**) determined that the accumulation of data held by Google, which enables it to individually identify persons using personal data, is "[the] sole objective pursued by the company is to gather a maximum of details about individualised persons in an effort to boost the value of their profiles for advertising purposes".[209] In line with this reasoning, public keys or other sorts of identifiers used to identify a natural person constitute personal data.

The next section looks at various technical approaches to re-identification using a number of practical examples and considers the issues that arise.

Technical measures for re-identification – pseudonymous or anonymous?

Actors interested in using DLT and worried about UK GDPR compliance will seek to avoid the processing of personal data to start with. However, as noted below, this is far from straightforward as much of the data conventionally assumed to be non-personal qualifies as personal data as a matter of fact.

205 Information Commissioner's Office, Anonymisation: Managing Data Protection Risk Code of Practice (November 2012) 16 <https://ico.org. uk/media/1061/anonymisation-code.pdf> Accessed 13 April 2020. Other data protection authorities have reached different conclusions but we have not considered them here.

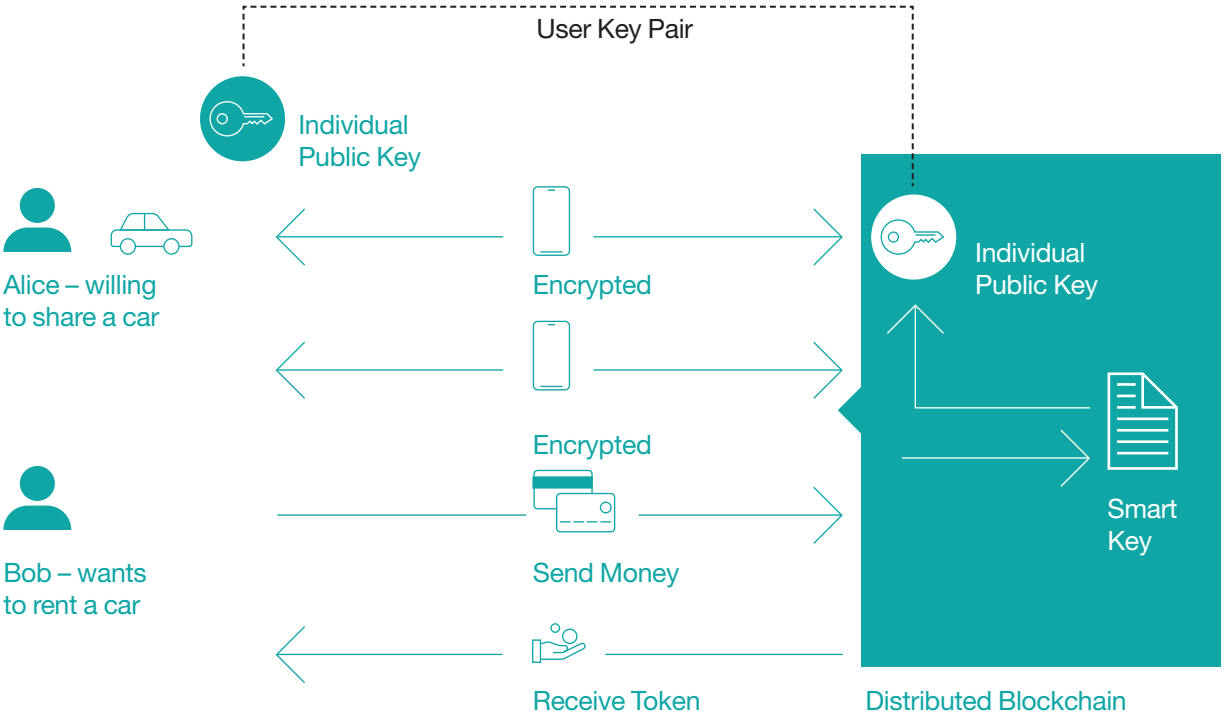
206 Article 29 Working Party, Opinion 05/2014 on Anonymisation Techniques (2014) WP 216 0829/14/EN, 3 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf> Accessed 13 April 2020

207 Michèle Finck, Frank Palas, 'They who must not be identified – distinguishing personal from non-personal data under the GDPR', (2020) 10(1) IDPL 11, 26 <https://academic.oup.com/idpl/advance-article/doi/10.1093/idpl/ipz026/5802594> Accessed 13 April 2020

208 Article 29 Working Party, Opinion 04/2007 on the Concept of Personal Data (2007) WP 136 01248/07/EN, 16 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf> Accessed 13 April 2020

209 Commission Nationale de l'Informatique et des Libertés, 'Deliberation No. 2013-420' (Sanctions Committee of CNIL, 3 January 2014) < https://www.legifrance.gouv.fr/affichCnil.do?oldAction=rechExpCnil&id=CNILTEXT-T000028450267&fastReqId=1727095961&fastPos=1ff> Accessed 13 April 2020

Scenario:



In this scenario, Alice is willing to rent her car to Bob. In order to do this, both Alice and Bob will install an app on their personal device (e.g. a smart phone) and verify their respective digital identities (using a driver's licence or other form of ID). This will need to be verified by a third party. Once the verification process is complete, Bob will need to agree to all applicable terms and conditions in respect of price, rental duration, insurance policies and more. Once approved, Bob can proceed with verification on the smart contract. Payments will be made by reducing the balance in Bob's wallet and sending it to Alice's wallet. After payment, Bob will receive a unique car token with which to enter the car.

Is transactional data 'personal data'?

In order for the payment from Bob to Alice to work, Bob and Alice will create and manage their addresses in wallets (here, a wallet app on their smart phones). The address is a public key belonging to a private-public key pair randomly generated by a particular user. Bob will therefore transfer money from his address, 'A', to the address key of Alice, 'B', and sign the transaction with the private key responding to A. Where a blockchain uses proof of work, miners validate the transaction based on the public key A and the publicly known balance. While the transactional data is not explicitly related to a natural person, it is related to an identifier (the address) which is pseudonymous data and may be classified as 'personal data' if you are able to single out the individual; by linking records to the individual and inferring information concerning the individual, the address may become personal data.[210]
Steps to take to prevent identification?

To prevent re-identification of a natural person, there are a few approaches that one can take. Though by no means exhaustive, these include:

— Use hash-based pseudonyms instead of clear-text identifiers. These are irreversible or one-way functions;[211]

210 Article 29 Working Party, Opinion 05/2014 (n 25) 14
211 In October 2019, the European Data Protection Supervisor (EDPS), in conjunction with the Spanish data protection authority, has also issued a joint paper on the hash function as personal data pseudonymisaton technique: https://edps.europa.eu/sites/edp/files/publication/19-10-30_aepd-edps_paper_hash_final_en.pdf (accessed 9 August 2021).

— Consider 'salting' and 'peppering' the hash to prevent re-identification. In both cases, additional data is added to the clear-text data before the hash function is applied, but the added data differs between contexts so that the resulting hashes also differ. There is, however, some argument that these methods can make the system more vulnerable, as each next validation relies on the validation of the previous hash, so if wrong once, the error could cascade through the system;

— Keep details of each party's identity off-chain to enable it to be modified and deleted;

— Consider the implementation of ring signatures and ZKP. Ring signatures hide transactions within other transactions by tying a single transaction to multiple private keys even though only one of them initiated the transaction. The signature proves that the signer has a private key corresponding to one of a specific set of public keys, without revealing which one. By using ZKP techniques, an individual (e.g. Bob) could prove to the owner of the car that he or she meets the rental requirements (e.g. a valid driver's license, insurance coverage, and bank account to cover costs) without actually passing any personal data, such as driver's license number, home address, and insurer, to the owner of the car (Alice). Where ZKP is used, the blockchain only shows that a transaction has happened, not which public key (Bob, as sender) transferred what amount to the recipient (Alice). For further details on ZKP see Part B on data security measures. This would also help with compliance with data protection principles, such as the purpose limitation and data minimisation principles.[212]

While these steps all assist in preventing transactional data being classified as 'personal data' under the UK GDPR, there is at present no legal certainty for developers wishing to handle public keys in a UK GDPR compliant matter and the Group considers that further guidance is needed from data protection authorities in respect of this.

The benefits of blockchain as a means to achieve UK GDPR's objective

Blockchain technologies are a data governance tool that support alternative forms of data management and distribution and provide benefits compared with other contemporary solutions. Blockchains can be designed to enable data sharing without the need for a central trusted intermediary. They also offer transparency as to who has accessed data, and blockchain-based smart contracts can automate the sharing of data, which has the additional benefit of reducing transaction costs. These features may assist the contemporary data economy more widely, such as where they serve to support data marketplaces by facilitating the inter-institutional sharing of data. Furthermore, they could provide data subjects with more control over the personal data that directly or indirectly relates to them. This would accord with the right of access (Article 15 UK GDPR) and the right to data portability (Article 20 UK GDPR), that provide data subjects with control over what others do with their personal data and what they can do with that personal data themselves.

Further guidance and support by regulatory authorities is required before these projects can become more mainstream.

On the basis of the Group's discussions and evidence examined, the Group believes that some of the questions to be addressed by the ICO and other data authorities should include the following:

— What does "all means reasonably likely to be used" mean under Recital 26 UK GDPR? Does this require an objective or subjective approach?

212 Under the UK GDPR one is expected to comply with the purpose limitation which means that data is only collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes and the data minimisation principle which means that data ought to be 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed' (see the UK GDPR, Article 5(1)(b) and (c)).

— Does the use of a blockchain automatically trigger an obligation to carry out a data protection impact assessment?

— Does the continued processing of data on blockchains satisfy the compelling legitimate ground criterion under Article 21 UK GDPR?

— How should "erasure" be interpreted for the purposes of Article 17 UK GDPR in the context of blockchain technologies?

— How should Article 18 UK GDPR regarding the restriction of processing be interpreted in the context of blockchain technologies?

— What is the status of anonymity solutions such as ZKP under UK GDPR?

— Should the anonymisation of data be evaluated from the controller's perspective, or also from the perspective of other parties?

— What is the status of the on-chain hash where transactional data is stored off-chain and subsequently erased?

— Can a data subject be a data controller in relation to personal data that relates to them?

— What is the relationship between the first and third paragraph of Article 26 UK GDPR? Is there a need for a nexus between responsibility and control?

— How should the principle of data minimisation be interpreted in relation to blockchains?

— Is the provision of a supplementary statement sufficient to comply with Article 16 UK GDPR?

Dr. Finck outlines other questions to be addressed in *Blockchain and the General Data Protection Regulation: Can distributed ledgers be squared with European data protection law?* [213]

None of these questions has been formally addressed since the publication of the 2020 guidance.

PART B: Data Security Enhancing Measures

Adi Ben-Ari (Applied Blockchain)

Introduction – Zero Knowledge Proofs

ZKPs are cryptographic outputs that can be shared and used by one party to prove to another that it is in possession of data with certain properties, without revealing anything else about that data.

In order for a cryptographic scheme to be considered a ZKP, it must demonstrate the following properties:

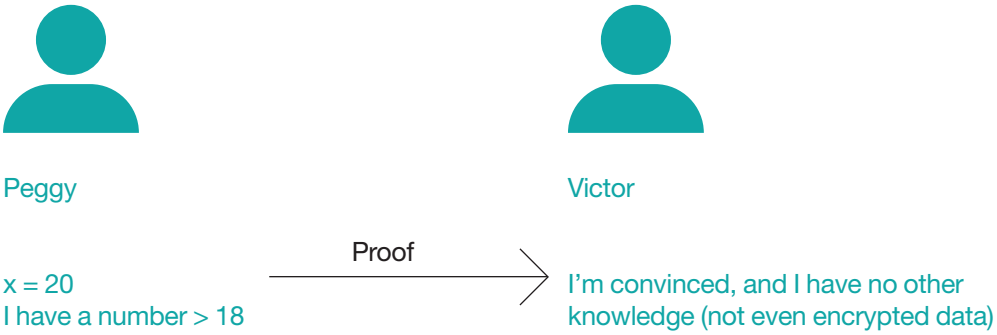
— **Completeness:** If the statement is true, an honest verifier will be convinced of this fact by the honest prover. That is, the algorithm must work in the sense that the party verifying the proof is satisfied that the proving party is in possession of the underlying data.

213 Michele Finck, 'Blockchain and the General Data Protection Regulation: Can Distributed Ledgers Be Squared With European Data Protection Law?' (STOA: Panel for the Future of Science and Technology, 2019) 97-98 <https://www.europarl.europa.eu/RegData/etudes/STUD/2019/634445/EPRS_STU(2019)634445_EN.pdf> Accessed 28 December 2019

— **Soundness:** If the statement is false, no cheating prover can convince the honest verifier that it is true, except with some small probability.
— **Zero knowledge:** If the prover's statement is true, no verifier learns anything that was intended by the prover to be protected, other than the fact that the prover's statement is true.

Proof of age example

An oft-cited example is proof of age. There are many situations in life, including in the digital world, where a person might be required to prove that they are over 18 years of age, including access to age appropriate content, purchase of goods that may only be sold to persons over 18, and signing agreements that require the consent of an adult.



However, a person's age can constitute personal data for the purposes of data protection law, and many individuals would prefer not to share such information with a third party unless it is absolutely required. In fact, an important principle of the UK GDPR regulation is minimisation, where data processing should only use as much data as is required to successfully accomplish a given task.

Using ZKP, an individual possessing an item of data on their device expressing their age may now generate and provide a zero-knowledge cryptographic proof that they are over 18 without revealing their actual age. This would, in theory, allow them to satisfy the requirement of a third party by proving that they are over the age of 18, while at the same time protecting their data and implementing the UK GDPR minimisation by not revealing or sharing their actual age (or any other personal data) with the third party.

There are two potential flaws in this approach, and they illustrate how this technology should be considered in practice:

1. the prover could simply issue a statement that they are over 18, without the need for sophisticated cryptography; and

2. if the data the prover holds is incorrect, then a ZKP will be of little value to the third party verifier.

Simply issuing a statement

If a prover was to simply issue or sign a statement that they are over the age of 18, they would be making an assertion without providing any proof of that assertion. In other words, the prover could lie. This presents a risk to a third party who needs to be satisfied as to the prover's age, and often they will ask for proof in the form of a government issued document (e.g. driving license or passport). If the prover were to present such a document, they would be handing over their personal data (typically more than just their age), and be exposing themselves to the risk that their data may be used inappropriately or fraudulently, and may even be stolen or sold for commercial gain. The verifying organisation may also be non-compliant with the UK GDPR minimisation principle, as it is collecting more personal data than is required to satisfy the age check requirement.

Proving the information correct

If the verifier receives proof that a prover's dataset shows that they are over the age of 18, but doesn't trust the dataset itself (whether because the wrong data was mistakenly or deliberately inputted to the prover's dataset by the prover or another party), then further verification is required. In the proof of age example, the verifier would likely revert to government issued identification as a secondary verification step.

A ZKP system might therefore also include a third-party signature verifying the accuracy of a prover's dataset. The verifier can then be satisfied that not only does the prover's dataset asserts that they are at least aged 18, but that such dataset (and therefore the assertion) has been signed by and verified by a third party such as a government entity. In other words, the requirement of the verifier to be satisfied that the prover is over the age of 18 is now achieved through the sharing of a cryptographic proof without receiving the precise age of the individual, nor the government documentation.

Types of provable knowledge

The first generation of ZKP enable proof of the following:

— **Range proofs:** a prover is in possession of a number within a range (e.g. age).

— **Location within a geofence:** a prover is located in a region (e.g. London), without revealing the prover's exact location (e.g. a specific road in a specific borough of London).

— **Set membership/non-membership:** a prover holds a value that is a member or not a member of a particular set of values (e.g. AML checks on sanction lists).

— **Anonymous provenance to a cryptographic identity:** a prover owns an asset, together with properties of the asset's history, without revealing the history of the prover or historic parties.

This is not an exhaustive list but illustrates the type of data properties that ZKP systems can prove for data in a prover's possession.

State of technology

ZKP technology is very much in its infancy and new, more secure, more efficient algorithms are regularly announced. Government entities that sanction use of cryptography algorithms for government and industry (e.g. NIST) are yet to make their official recommendations, which we look forward to in due course.

Everything described thus far in this section can be achieved without a blockchain. The added value of a blockchain-based ZKP is twofold:

1. **Immutability.** An activity can be recorded, ordered, time-stamped and then jointly secured by a group of parties, which is potentially more secure than relying on the ordering and time stamps set and stored by an individual party who may modify or even destroy records. This can improve the verifier's confidence in the integrity of a prover's dataset.

2. **Double spend prevention.** In the case of assets, blockchain-based ZKP can provide assurance to verifiers that a single copy of an asset is available to all parties, avoiding duplicate records, as well as removing the need to trust a single party to hold and manage all of the records.

These additional attributes may or may not be required for a particular use case of ZKPs.

ZKP and blockchain

One of the myths surrounding blockchains is that the data stored on them is automatically encrypted. In some blockchains (e.g. the Bitcoin blockchain) cryptography is primarily used to sign messages and ensure that historical transactions confirming asset ownership can be secured by a group. Nevertheless, the data showing the wallet holdings and transfers between wallets is publicly available.

There was a conflict between the need for transaction and data privacy on the one hand, and the need for transparency and verifiability on the other. Prior to ZKP, privacy was achieved in enterprise blockchains by separating the parties into "mini" blockchains, also known as private channels. The issue with this approach is that the number of validating parties for private activity, and therefore overall security and integrity assurance of the blockchain, is greatly reduced. These issues motivated research into advanced cryptographic techniques that would eventually lead to the first practical implementations of ZKPs.

ZKPs enable the solving of both data privacy and verifiability issues at the same time. This is because, rather than storing the assets and data openly on a blockchain, ZKPs of their existence and consistency are stored. A transaction, such as transferring an asset to a different account, will only be permitted if ZKPs are available to verify the asset ownership. A new node in the blockchain can download a copy of all of the proofs and validate the consistency and historical correctness of the data without seeing any of the actual data.

ZKP and blockchain privacy

The first practical implementation of such a blockchain was zCash, launched in late 2016. zCash implemented a ZKP called a succinct, non-interactive argument of knowledge (zkSNARK). A succinct proof reduces the volume of data required to be stored on a blockchain network (thereby improving its performance), and a non-interactive protocol allows for one time generation of proofs that are stored indefinitely on a distributed ledger which multiple parties can verify, without each verifying party requiring interaction with the prover.

There are three stages in the life of a typical ZKP. These are:

1. Circuit production

2. Proof generation

3. Proof verification

A circuit expresses the mathematical logic that the proof will implement (e.g. prove a person is over 18). This will vary depending on the use case, and there are a number of initiatives to create multi-purpose generic circuits currently in development. The circuit acts as a template for producing a certain type of proof. The circuit need only be created once, and can then be used by multiple parties to generate proofs.

A more complex area of research and development is ZKP for privacy in blockchain-based smart contracts, where there exists a much broader range of functionality that would need to be expressed privately. A number of protocols are in development for smart contracts in Ethereum (Baseline, AZTEC) and Hyperledger Fabric (ZKAT), or both (Applied Blockchain's K0).

ZKP and blockchain scalability

ZKPs offer two approaches to improving the scalability of a blockchain platform. These are:

1. Rollups

2. Flat blockchains

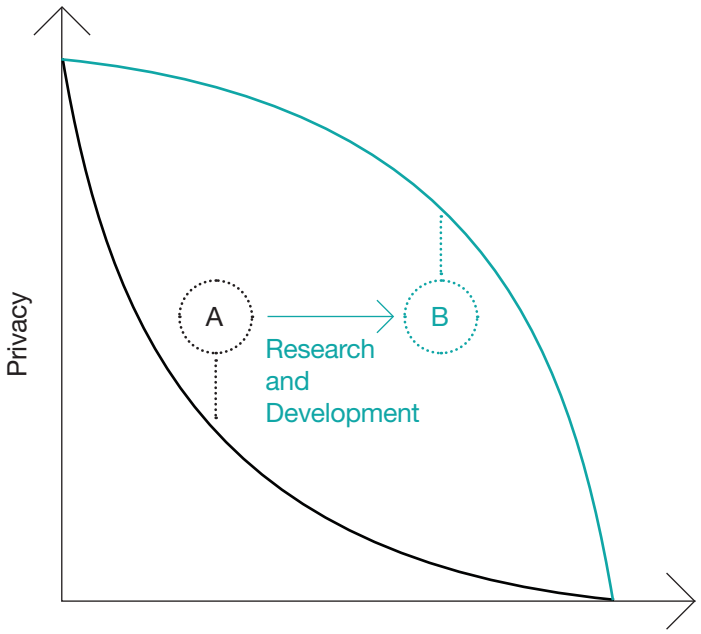
Rollups are designed to reduce the number of transactions on a blockchain by executing batches of transactions off-chain, rolling these up into a proof of the outcome of the transactions, and then posting only the proof to the blockchain. This greatly reduces the load on a blockchain, as it is no longer required to execute all of the transactions on-chain.

Succinct blockchains are even more compact and never grow. Rather than maintaining a full and growing history of transactions in each node, a flat blockchain will only ever contain a single row. This single row is a ZKP of the current state of the accounts on the blockchain. Any party can verify the proof and be satisfied with the integrity of the blockchain despite the fact that they have no access to the underlying data and transactions. Each time a new block of transactions is generated, a ZKP is created to prove the changes to the blockchain taking into account the previous proof. The technique is known as recursive zkSNARKs, and the result is that transactions are compressed to the point where the blockchain hardly grows.

As has been illustrated, ZKP technology is having a profound impact on the structure and implementation of blockchains. The capabilities described in this section were not available two or three years ago, when the popular enterprise platforms in use today were designed and conceived.

Other Privacy Enhancing Technologies (PETs)

Another example of a PET is Homomorphic Encryption (HE), and the closely related Somewhat Homomorphic Encryption (SHE) and Fully Homomorphic Encryption (FHE). These cryptography schemas enable data to be encrypted in a way that allows third parties to run calculations on the encrypted data without having the ability to decrypt and see the data. This may be particularly useful where data processing is outsourced to cloud computing services, but the data is of a sensitive nature and the data owner wishes to keep the data hidden from the cloud data processor. It may also enable analytics companies to perform analytics on data that is not shared with them.



These technologies are part of a greater trend to increase data privacy by sharing less, while enabling increasing utility from privately held data. This is in direct contrast to the proliferation of data sharing in recent decades when both individuals and companies shared vast quantities of data with third parties in return for utility.

Hardware Secure Enclaves

An additional emerging technology for preserving data privacy is the hardware secure enclave (HSE). This is an area of a computer chip that is isolated by hardware and prevents other areas of the computer from having access to data inside. This means that even the system administrator of a device or someone with physical access to the machine would not have access to the data inside the HSE.

A common use of HSEs is to store private keys. A private key and public key pair is generated inside a hardware enclave. The public key is shared, but the private key never leaves the enclave. Data can be sent to the enclave for signing by the private key, but the key itself is never revealed. An example of hardware secure key storage is Apple Pay, where the private key to initiate payments is stored in an enclave on the phone, and the key itself cannot be shared with Apple or any apps. Instead, the key can sign transactions proving that they came from the device (in this case, use of the enclave is also tied to the biometrics tests conducted on the device).

HSEs have many more uses beyond key storage. In fact, any data can be sent to an enclave, and any private processing can occur in the enclave. Unlike ZKPs and other software-based cryptography methods, hardware enclaves run at almost the same speed as regular tasks that run on the processor. This means that performance and scalability issues associated with software-based cryptography do not apply in a hardware secure enclave environment.

Intel's SGX (secure guard extensions) is an example of a relatively mature hardware secure environment that enables complex privacy-preserving applications.