Proof of concept
Blockchain and the right to erasure

November 2024

Executive summary

Blockchain is a technology that makes it possible to create infrastructures for storing and exchanging information. Different processes can use these infrastructures, among other systems, to implement services. The management of the infrastructure itself requires the execution of specific processes. Both sets of processes may involve processing of personal data.

Blockchain infrastructure designs have generally not applied data protection by design. In many cases, the development of infrastructures has used components that have not been thoroughly documented, nor have they been analysed by their developers and managers. In most Blockchain infrastructures, de facto governance measures have been applied, often due to unforeseen problems, and also poorly documented.

The Proof of Concept presented in this document demonstrates the feasibility of building GDPR-compliant Blockchain infrastructures.

This Proof of Concept analyses and documents the components of a Blockchain infrastructure, widely used in the market, in compliance with the accountability principle. It also discusses real-world cases of change implementation and governance management common to such an infrastructure. Next, policies, including organisational and technical measures, are developed to implement the right to erasure in a Blockchain infrastructure by means of inconsistency management. Finally, they are practically applied in a use case on a real Blockchain infrastructure.

While previous work exists to manage the deletion of information in a Blockchain infrastructure, this Proof of Concept is a fully functional and documented demonstrator that is specifically geared towards GDPR compliance. Furthermore, it contemplates the management of personal information stored in the entire Blockchain, i.e. not only the information in block's transactions, but also other information such as that recorded in transaction receipts.

The Proof of Concept demonstrates that compliance with the GDPR is possible in Blockchain infrastructures, specifically in relation to compliance with the right to erasure, without claiming to be a commercial solution for direct market application. It is important to stress that, although the Proof of Concept has been developed on an infrastructure that has not implemented data protection by design, it is not intended to validate such a way of proceeding. On the contrary, it is intended to promote among designers, authorities and organisations that have a role in the design or development of such infrastructures, the adoption of data protection by design and by default strategies.

This document aims to serve as a bridge between the data protection professional and the Blockchain technologies professional, therefore, it will analyse the terminology used and clarify the concepts to avoid misunderstandings hidden behind imprecise descriptions. Furthermore, it is complemented with additional technical information and demonstrative videos that are accessible on the AEPD website, in the Innovation and Technology section.

INDEX

I. INTRODUCTION

The deployment of new personal data processing on Internet services, such as digital identity management, digital currencies, access to certificates or the medical record itself, has been proposed using different technologies. Some of the technological options incorporate features of decentralised availability of databases, exchanges through P2P protocols[1] and management of the integrity of the stored digital information. One such technology is the so-called Blockchain.

The term "Blockchain" is sometimes used in isolation to refer to the technology as it is ideally defined, sometimes to a specific materialisation in an infrastructure, and sometimes to the data set that is managed in a specific infrastructure. In this text we will differentiate the three meanings with the terms Blockchain technology, Blockchain infrastructure, and "Blockchain/table" respectively.

Blockchain technology makes it possible to implement data management infrastructures in a distributed and decentralised manner. Using this technology, and adapting it to different use cases, Blockchain infrastructures can be created as a concrete materialisation. Different Blockchain infrastructures will adjust the principles of Blockchain technology and complement it with additional functions or systems. In this way, we can speak of Bitcoin or Ethereum infrastructures, among others, which are different ways of bringing the possibilities of Blockchain technology to reality, but incompatible with each other, and with different functionalities and objectives.

Blockchain technology, and many of the infrastructures in which it materialised, were initially designed to operate in a deregulated environment, i.e. an environment in which the controls established by the rule of law were ignored. Moreover, they incorporated almost fully automated governance models and management procedures, without recourse to trusted third parties or supervisory bodies. In this way, it was possible to implement an information storage and sharing infrastructure in which all participants were equal in powers and functionalities (peers), eliminating the assignment of responsibilities to a centralised body and distributing responsibility to those persons (natural or legal) who decided to participate in such an infrastructure. In principle, and without contracts or legal acts, the participants did not acquire commitments such as, for example, those guaranteeing a level of quality of service or compliance with the various regulations[2].

Bitcoin was the first infrastructure created using Blockchain technology. Since then, different infrastructures have been created that have innovated, developed and modified the principles of this technology. Therefore, although we can speak of the properties of Blockchain technology in a generic way, the specific implementation in different infrastructures can substantially alter its characteristics.

Different data processing may use a Blockchain infrastructure to implement some of their operations, such as data storage. Such processing will need to include other elements in addition to the Blockchain infrastructure to implement other operations[3].

But, in addition, some specific operations are needed to interact with the infrastructure. This set of additional operations or systems, when integrated with the Blockchain infrastructure, is what in the field of this technology has been called the Blockchain ecosystem. These can be digital wallets, issuers of credentials or attributes, developers of programmes and applications in the infrastructure, browsers or data viewers, additional

[1] P2P (Peer to Peer): Network between peers. The connected nodes behave as equals to each other. It allows the direct exchange of information between the interconnected nodes. Examples for file sharing are eMule and BitTorrent but they have other applications (https://en.wikipedia.org/wiki/Peer-to-peer).

[2] Hence the conflicts with copyright or intellectual property regulations that have given rise to the aforementioned P2P networks.

[3] Such as cloud storage and processing systems, AI systems, etc.

governance mechanisms, etc. All these elements introduce operations, strengths and weaknesses that alter the theoretical properties that Blockchain enables. For example, although some implementations have been built as an anonymous means of exchange between individuals, reality (and the ecosystem) has shown that such anonymity does not exist.

One of the problems faced by processing built on Blockchain infrastructures is that the governance model employed has not considered compliance with any type of regulation as an objective. In particular, requirements to ensure and be able to demonstrate, compliance with the GDPR have not been taken into account. Therefore, they have been designed without taking data protection into account from the design stage.

To the extent that a process implemented on Blockchain does not involve personal data, this is not an impediment to its use from a GDPR perspective, without prejudice to other regulatory limitations. Similarly, if Blockchain is used in certain personal data processing operations to implement functions that do not involve personal data, for example, maintaining a repository of entity certificate data that supports personal data operations, the use of Blockchain would also have no impact from a GDPR perspective.

However, where the use of such technology is the means chosen to implement personal data processing operations, then, in accordance with Recital 15 of the GDPR, "*In order to prevent creating a serious risk of circumvention, the protection of natural persons should be technologically neutral and should not depend on the techniques used*", compliance with the GDPR must be ensured.

The ultimate purpose of a processing is different from the means selected to implement it. Article 25 of the GDPR states that the controller is obliged to implement appropriate technical and organisational measures in order to ensure and demonstrate that the processing is in compliance with the GDPR. In this regard, the European Data Protection Board has stated that technical impossibility cannot be invoked to justify non-compliance with the requirements of the GDPR, especially given that Article 25(1) of the GDPR provides that data protection by design shall be taken into account at the time of the determination of the means of processing and at the time of the processing itself[4].

The use of processing based on Blockchain infrastructures has grown significantly in various sectors, with numerous applications beyond the financial or cryptocurrency sphere, such as supply chain management, asset *tokenisation*, asset traceability, digital identity management, voting, land registries, metaverse, etc. A controller's choice of a particular Blockchain infrastructure as an element of its processing could lead to specific breaches and risks to data subjects' rights and freedoms. One of its key aspects is to implement data protection principles and to enable the exercise of data subjects' rights, in particular the accuracy principle, the storage limitation principle and the rights to rectification and erasure. In case of non-compliance due to the choice of a particular Blockchain infrastructure, the design of the Blockchain infrastructure or the processing has to be changed.

No technological option is immutable, least of all those based on digital systems. Regulatory compliance requires system adaptations, and all system components must be properly documented to ensure and demonstrate compliance.

In order to demonstrate that it is possible to develop Blockchain infrastructures that enable data controllers to comply with the GDPR, the AEPD has developed this Proof of Concept. The Proof of Concept includes the definition of governance measures, policies and the implementation of the necessary technical modifications to facilitate compliance with the

[4] EDPB, Report of the work undertaken by the ChatGPT Taskforce of May 2024, paragraph 7: "*In particular, technical impossibility cannot be invoked to justify non-compliance with these requirements, especially considering that the principle of data protection by design set out in Article 25(1) GDPR shall be taken into account at the time of the determination of the means for processing and at the time of the processing itself*".

GDPR in the aforementioned aspects. To this end, after a study of the source code of the necessary components to build a Blockchain infrastructure, several use cases have been executed on the official Ethereum implementation, configured with the consensus mechanism called Proof of Authority[5].

The use cases include a set of cryptocurrency transfer transactions between several accounts (users), as well as the interaction of some of them with two *Smart Contracts* programmed for the Proof of Concept. The Proof of Concept makes it clear that the right to erasure of one of the accounts participating in the transactions is possible.

This Proof of Concept is aimed at controllers and processors of data processing that are built on Blockchain infrastructures, controllers and processors of data processing that support the operation of the infrastructure, and those authorities and organisations that support the development of projects based on Blockchain technologies. This document does not assess the assignment of the roles of data controller or data processor in each of the processing operations of specific Blockchain infrastructures or the lawfulness of the processing operations.

[5] Proof of authority is a reputation-based consensus algorithm rather than a participation-based mechanism such as PoS. Proof of authority requires relying on a set of authorised signers that are established in the genesis block (some implementations allow new signers to be included during the operation of the Blockchain). In most current implementations, all authorised signers retain equal power and privileges in determining the consensus of the chain. The idea behind reputation *staking* is that all authorised validators are known to all through mechanisms such as know your customer (KYC) or having a known organisation as the sole validator; this way, if a validator does something wrong, their identity is known. Source: https://ethereum.org/en/developers/docs/consensus-mechanisms/poa/

II. BASIC CONCEPTS

When analysing the impact that certain technological options may have from a data protection point of view, it is essential to define the terminology used rigorously and to make sure that its actual and technical meaning is accurately understood. This set of descriptions aims to provide clarity and understanding of this technology[6].

a) **Account or address (of an account)**: Unique identifier obtained from the user's public key that allows interaction and transactions on the Blockchain infrastructure. The address of a *Smart Contract* is its unique identifier on the Blockchain, but it is not based on a public key.

b) **Block/record**: A fundamental unit of information containing a set of validated transactions and other relevant data, linked by a hash with previous blocks to manage the integrity of the order of transactions.

c) **Blockchain/table**: A data structure representing a distributed general ledger, with sequentially organised, confirmed blocks/records, which only allows to add blocks/records linked by hash links[7]. Hence the name Blockchain. However, this is not the only data structure in the various Blockchain infrastructures. They need to use several additional structures, which allow their efficient functioning and operation.

d) **Blockchain Ecosystem**: Set of technologies, platforms, networks, services, applications, systems, databases, organisations, developers, nodes, users and processes that interact around a Blockchain infrastructure and enable the execution of data processing.

e) **Blockchain Infrastructure or System[8]**: A concrete instance of a peer-to-peer information exchange network using Blockchain technologies. In many cases, the technologies for building a Blockchain are confused with a specific instance of the use of these technologies, more or less adapted to specific use cases. A CBDC and Bitcoin may use Blockchain technologies, but they are two radically different infrastructures.

f) **Blockchain Technology**: Technology that enables the operation and use of Blockchain systems[9]. It is the technical solution that defines the general conditions for independent and equal parties (peers) on the Internet to exchange information, with a mechanism for agreeing what information is to be stored, and a mechanism for managing the integrity of both the information and the order in which it is stored.

g) **Consensus:** Agreement reached between nodes that a transaction is valid. Also agreement that a set of valid transactions is stored in a consistent order. There are several consensus mechanisms for determining which node adds a new block to the chain. The best known are the *Proof of Work* (PoW) used in Bitcoin, the *Proof of Stake* (PoS) used in Ethereum, and the *Proof of Authority* (PoA) used by Ethereum in private and test networks[10].

h) **Consistency**: The degree to which data are free of contradiction and consistent with other data in a specific context of use. It can be analysed in data that refer to both one and several comparable entities[11]. In Blockchain technology, measures are implemented so that the information shared between nodes is the same and

[6] ISO 22739 "Blockchain and distributed log technologies - Vocabulary" contains definitions related to Blockchain. A web version is available in English: https://www.iso.org/obp/ui#iso:std:iso:22739:ed-2:v1:en

[7] ISO 22739.

[8] ISO 22739 System implementing a distributed ledger.

[9] ISO 22739

[10] Ethereum has stopped supporting PoA consensus since v1.14.0 on 24 April 2024, however there are several Ethereum-based implementations that do support it, such as Hyperledger Besu.

[11] ISO 25012.

there are no contradictions, i.e. all nodes in a given Blockchain infrastructure share the same state at a given point in time.

i) **Cryptocurrency**: A digital asset that is used as a medium of exchange or value. Some Blockchain infrastructures use cryptocurrencies as their native token, which is integrated into their protocol and fundamental to their operation (it is not created by a *Smart Contract*). It is not a currency, unlike CBDC *(Central Bank Digital Currency)*. Not all Blockchain infrastructures implement their own or native cryptocurrencies. However, cryptocurrencies represent the most successful use case of Blockchain technology[12].

j) **Data traceability**: The ability to know the entire data lifecycle: the exact date and time of extraction, when it was transformed, and when it was loaded from one source environment to another destination. This process is known as *Data Linage*[13].

k) ***Distributed Ledger Technology*** **(DLT)**: Distributed Ledger Technology (DLT) allows data to be stored in different participants in a network, synchronised through consensus mechanisms. Blockchain is a particular case of DLT.

l) ***Exchange***: Private platform that facilitates and allows users to buy, sell, *trade* and exchange cryptocurrencies and other cryptoassets (*Non Fungible Token* or NFT, etc.).

m) ***Fork***: In the context of Blockchain, this is the main mechanism by which software updates are implemented. A split in the network occurs, and it happens due to modifications in the code, in the protocol, or due to decisions made in certain situations (a *hard fork* causes separate and incompatible chains).

n) **Governance**: The exercise of authority and control to decide the objectives of an organisation; to prioritise and balance objectives; to make decisions, based on assets, resources, context and risks, to achieve those objectives; and to continuously monitor that progress on each action taken is on track.

o) **Hash**: A function that from the input of a set of characters of variable length, generates as output another string of fixed length, which satisfies the following properties: there is no computationally feasible procedure to obtain the input from the output, or to find two inputs that have the same output[14].

p) **Integrity**: The property of accuracy and completeness of data[15]. In this paper the term is used in the sense that there are measures in place to detect whether information has been unilaterally modified or altered. In the Blockchain environment, integrity management has been confused with the requirement of immutability[16] on a given Blockchain.

q) **Node**: (see Participant) A device or process that participates in a Blockchain infrastructure and stores a full or partial replica of the Blockchain/table[17]. Node is a technical concept. Depending on their functionality, different classes of nodes can be defined. Validator or miner nodes are those that validate and add transactions to the block and add new blocks to the data recorded in the Blockchain infrastructure through the corresponding consensus mechanism. They receive

[12] Although its greatest success is because it has been and is used for dubious or criminal purposes such as black market payments (*silkworm, darkweb, etc.*), *ransomware*, money laundering, circumventing bans, etc. in the manner of a digital "tax haven".

[13] https://datos.gob.es/en/blog/importance-data-cataloguing

[14] Although such properties may have vulnerabilities in some cases. See Introduction to the hash function as a personal data pseudonymisation technique.

[15] ISO 27000:2018

[16] Immutability is defined in ISO 22739:2024 as the property that data cannot be modified or deleted once it has been added to a distributed ledger. Immutability thus refers to a desirable requirement or objective in a given Blockchain infrastructure, rather than a property of that technology.

[17] ISO 22739

rewards and commissions in the form of the Blockchain cryptocurrency for this work. Private networks can be set up without these incentives.

r) ***Non Fungible Token*** **(NFT)**: Non Fungible Token (NFT) means that something is unique, has different properties from another NFT and cannot be replaced. In contrast, tokens representing cryptocurrencies are identical and have the same properties, i.e. they are fungible[18]. Examples of NFTs include artwork, comics, sports collectibles, trading cards, games and more.

s) **Participant**: (see Node) Natural or legal persons that set up, operate and maintain management procedures, commitments, potential outsourcing, devices and programmes that implement one or more nodes in a Blockchain infrastructure. They may act on their own behalf or on behalf of others. The participant is the person who will be the obliged subject of the various regulations.

t) ***Peer-to-Peer*** **(P2P)**: An information sharing infrastructure in which there is no hierarchical relationship between the participants, and each is independent in making decisions about how to participate in it.

u) ***Proof of Concept*** **(PoC)**: The realisation of a particular idea, method or principle in order to demonstrate its feasibility, or a demonstration in order to verify that some concept or theory has practical potential. A Proof of Concept is usually limited in scope and may not be complete.

v) **Recipient** (of a transaction): An account or address that is the recipient of the transaction. A *Smart Contract* account or address is the recipient of the transaction when calls are made to its functions and procedures.

w) **Sender** (of a transaction): Account or address of the user initiating the transaction. A *Smart Contract* account or address cannot issue transactions, as it is not in possession of a private key to be able to sign them.

x) **Smart Contract**[19]*:* It is a program that is stored in the Blockchain infrastructure and executes automated decisions when certain conditions programmed in it are met when they are invoked by a transaction. The result of transactions carried out with them is reflected in a change of the state of the information stored in the Blockchain infrastructure, which in turn is automatically recorded in the Blockchain. These changes can cause other *Smart Contracts* to be executed in cascade. Their name can be misleading, as they are neither contracts (in the legal sense of the term) nor are they smart (they are conditional automatic execution programmes). The continued use of this term is due more to its historical roots in the field of Blockchain technology than to a precise description of its functionality[20].

y) **Token**: Digital value representing physical or digital assets. It is created and managed on a Blockchain infrastructure using *Smart Contracts*. Tokenisation is the process of converting a physical or digital asset into a token that can be registered, transferred and managed on a Blockchain. This token represents a fraction, ownership or right to the underlying asset. Some tokens are used as cryptocurrencies.

z) **Transaction**: An operation that is recorded on the Blockchain and modifies the information on it, transferring data or value between the sender and the receiver. The transaction is digitally signed by the sender and must be validated by the nodes for its inclusion in a block/record on the Blockchain.

[18] https://ethereum.org/en/nft/#what-are-nfts

[19] Term coined by cryptographer Nick Szabo, circa 1993.

[20] Article 2(39) of the Data Act defines a smart contract in a broader sense as *"a computer program used for the automated execution of an agreement or part thereof, using a sequence of electronic data records and ensuring their integrity and the accuracy of their chronological ordering".*

aa) **User**: Individual or entity that interacts, in a broad sense, with the Blockchain infrastructure (carrying out transactions, accessing services, managing nodes, obtaining information, etc.).

III. TECHNOLOGICAL AND LEGAL FRAMEWORK

Blockchain is a technology that emerged in 2009 and materialised with the virtual currency or cryptocurrency *Bitcoin,* an infrastructure based on this technology that made it possible to implement a service with guarantees of transparency and integrity of the payments made. Blockchain technology is a particular case of *distributed ledger technology* or DLT, a term that refers to distributed databases, managed through consensus mechanisms by multiple participants.

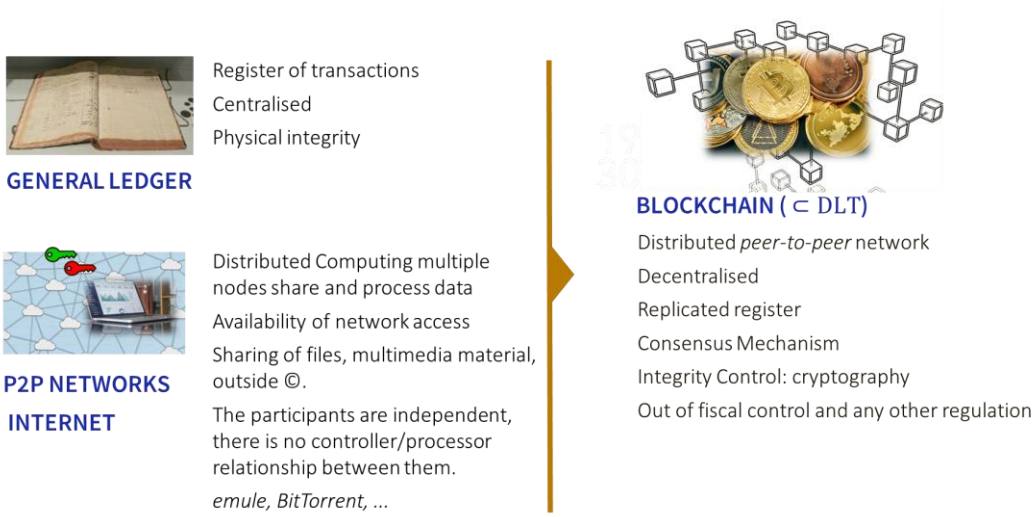


Figure 1. Blockchain is not a completely new technology, it emerges as an evolution of the ledger combined with the use of peer-to-peer (P2P) networks and no intermediaries.

A. DESCRIPTION OF BLOCKCHAIN TECHNOLOGY AND ITS APPLICATION

It is important to differentiate Blockchain technology from the different infrastructures created with this technology, which can be very different from each other. This section will describe the properties of the theoretical implementation of the Blockchain technology[21] as it was originally envisaged. Practical implementations have modified some (or many) of its fundamental characteristics[22].

In a simple way, an infrastructure created with Blockchain technology can be defined as a network of participants (natural or legal persons called *peers* or nodes) sharing a data set in a distributed way, where it is noted who owns what (assets in the form of data), where it is traded with whom these assets are exchanged (transactions) and with measures to manage the consistency and integrity of the data.

Participants/nodes can become part of a particular Blockchain infrastructure (e.g. Bitcoin, Ethereum, a CBDC or other) in their own interest or on behalf of other entities[23]. The terminology participant/node will be used in the text, as a participant is the natural or legal

[21] It has to be taken into account that in any implementation using technology, there is a big gap between generic or theoretical models and the reality of practical and commercial solutions. These not only adapt the concepts to a concrete implementation, but are complemented by additional components, functions and services that affect the ideal properties of the original concept.

[22] This text does not explain Layer 2 solutions, which are technologies that run on top of a blockchain protocol that improves the speed and efficiency of the underlying blockchain with lower fee costs. Layer 1 refers to the distributed database itself, the network that brings together all the nodes of the blockchain into one system with its underlying consensus mechanisms. For example, layer 1 of Bitcoin is the Bitcoin network. Layer 2, however, is an overlay network that sits on top of the blockchain. For example, the Lightning Network is a layer 2 solution for Bitcoin. https://crypto.com/university/what-are-layer-2-scaling-solutions. On Ethereum, there are numerous layer 2 solutions, such as Optimism, Arbitrum or ZKSync. https://ethereum.org/en/layer-2/.

[23] For example, a Blockchain infrastructure can be created with nodes under the direct control of different financial institutions, although some of them may outsource such processing to other legal entities to act on their behalf in managing nodes in the same infrastructure.

person that will manage one or more nodes in the infrastructure, on its own behalf or on behalf of third parties and will be the obliged subject in the regulation. On the other hand, a node is a device or process that is part of a particular infrastructure, i.e. a technical concept. Nodes in the same infrastructure use compatible, but not necessarily the same, means.

Ideally, each participant/node in the Blockchain infrastructure will store a copy of the Blockchain/ table as depicted in the Figure 4 plus any additional data structures implemented by the specific Blockchain infrastructure. It will be a decision of the participant[24], if not acting on explicit instructions from a third party[25], how he stores the information (some kind of database, another kind of dataset, etc.), what resources he dedicates to processing the information, where he stores the information (in one of the participant's systems, in the cloud, etc.), how many blocks/records he will store[26], and on what criteria or when access to such information is given, among others.

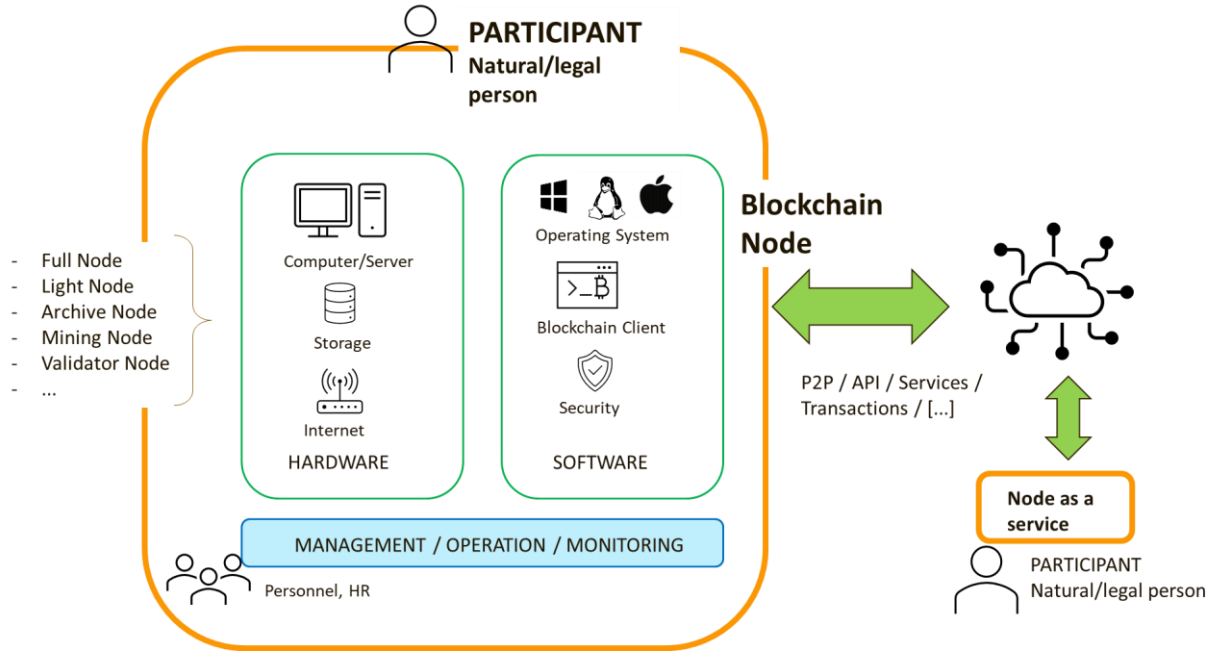


Figure 2 A node of a Blockchain infrastructure is a device for connecting to it. A participant is the person who manages the node and will be subject to regulatory obligations. The participant will be the decision on which Blockchain infrastructure to participate in and what/how much/how many physical and human resources to invest in the node for its operation and functioning.

In contrast to centralised storage systems, in a Blockchain infrastructure ideally all participants/nodes maintain a copy of the data set.

The name "Blockchain" comes from how the replicated data structure is organised in each of the participants/nodes. The data is stored in registers in which transactions are stored. The records are referred to as blocks in Blockchain terminology, but we will keep the name blocks/records in this text to make it easier to understand the explanations.

[24] The choice of means that a participant decides (except that we are in the case of the note below) includes not only the decision of which Blockchain infrastructure to participate in, but also the choice of software that is compatible with that infrastructure and how to implement it (as in any business sector), the machines on which it will run, how to store the blockchain and other systems, facilities and subcontractors. It should be borne in mind that in infrastructure such as Blockchain, and also as in any business sector, the optimisation of these resources is vital to obtain benefits.

[25] In the case of a processing of personal data, of a controller who comes to specify such details.

[26] In many infrastructures, for the operation of the infrastructure to comply with use cases, it is necessary to generate blocks/records continuously, so the number of blocks/records is very high, and many participants/nodes choose to store only the last blocks generated, avoiding not only the storage, but also the analysis of the integrity of the entire chain. This practice is called pruning.

The blocks/records are linked chronologically by storing in each block/record the hash of the previous block/record. This is a technique to manage the integrity of the information, since, if the information in a block/record changes, it would be detected by the inconsistency of the stored hash values.

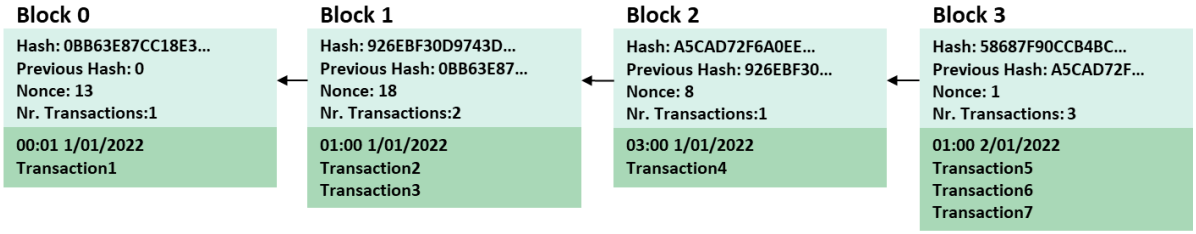


Figure 3. Didactic representation of a set of blocks/records as a chain.

The collection of all blocks/records is referred to as chain in Blockchain terminology. This name can lead to confusion about how and where the blocks are located. The chain is actually stored at each node as a data structure in which the blocks/records are organised. This structure is a database, which in some cases will be relational but most commonly is a key-value database[27] . In Figure 4 the chain is represented as a table, which would be replicated in the participants/nodes of the Blockchain infrastructure. Therefore, the term Blockchain/table will be used in this text.

Nr. Block	Hash	Previous block Hash	Nonce	Nr. Transc	Date	Transactions
0	0BB63E87CC18E3...	0	13	1	01/01/2022 0:01	Transaction1
1	926EBF30D9743D...	0BB63E87...	18	2	01/01/2022 1:00	Transaction2 Transaction3
2	A5CAD72F6A0EE...	926EBF30...	8	1	01/01/2022 3:00	Transaction4
3	58687F90CCB4BC...	A5CAD72F...	1	3	02/01/2022 1:00	Transaction5 Transaction6 Transaction7
4	.....					

Figure 4. Representation of the storage of the above blocks in a node, in the form of records in a table of a relational database. Usually, the blocks/records will be stored in a database or other systems for efficient exploitation. This table does not reflect the additional data structures that need to be stored in the node to manage other information.

In addition, participants/nodes need to store several additional data structures[28] other than the Blockchain/table, which enable its efficient functioning and operation, such as access to information without the need to process the entire chain from the initial or genesis block. Some of these additional data sets are state of accounts and balances; storage of *Smart Contracts*; a temporary record of pending transactions; transaction receipts, which provide information about the outcome of the transaction, including events and logs issued by a *Smart Contract*; information about other nodes to maintain active connections and manage communication and synchronisation; some network and node configuration parameters; intermediate states and copies of the chain state at specific times to facilitate synchronisation; etc. The exact implementation may vary depending on the specific Blockchain infrastructure and the optimisation performed by each participant/node.

[27] https://aws.amazon.com/nosql/key-value/

[28] https://ethereum.org/en/developers/docs/data-structures-and-encoding/patricia-merkle-trie/, https://github.com/ethereum/go-ethereum/blob/master/core/rawdb/schema.go, https://geth.ethereum.org/docs/fundamentals/command-line-options

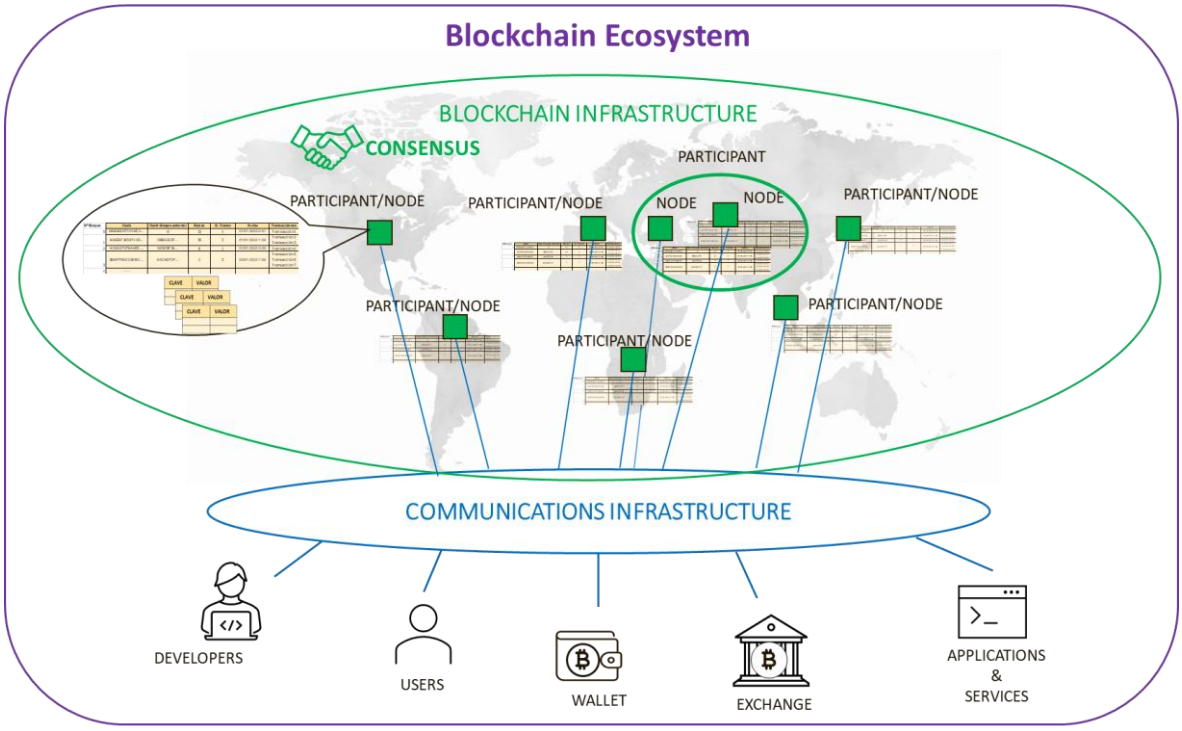


Figure 5. The Blockchain infrastructure is made up of a number of participants/nodes, each storing the blocks/records in a data structure called a Blockchain/table.

In general, Blockchain infrastructures assume that there is no central entity to verify and validate transactions. It is therefore necessary to appeal to consensus mechanisms among participants to make decisions, update data and maintain the consistency of the information stored. In this way, transactions are stored in time-ordered records that are linked to each other, forming an apparently "immutable" and transparent chain, which can even be updated "automatically" by means of smart contracts hosted on the chain itself.

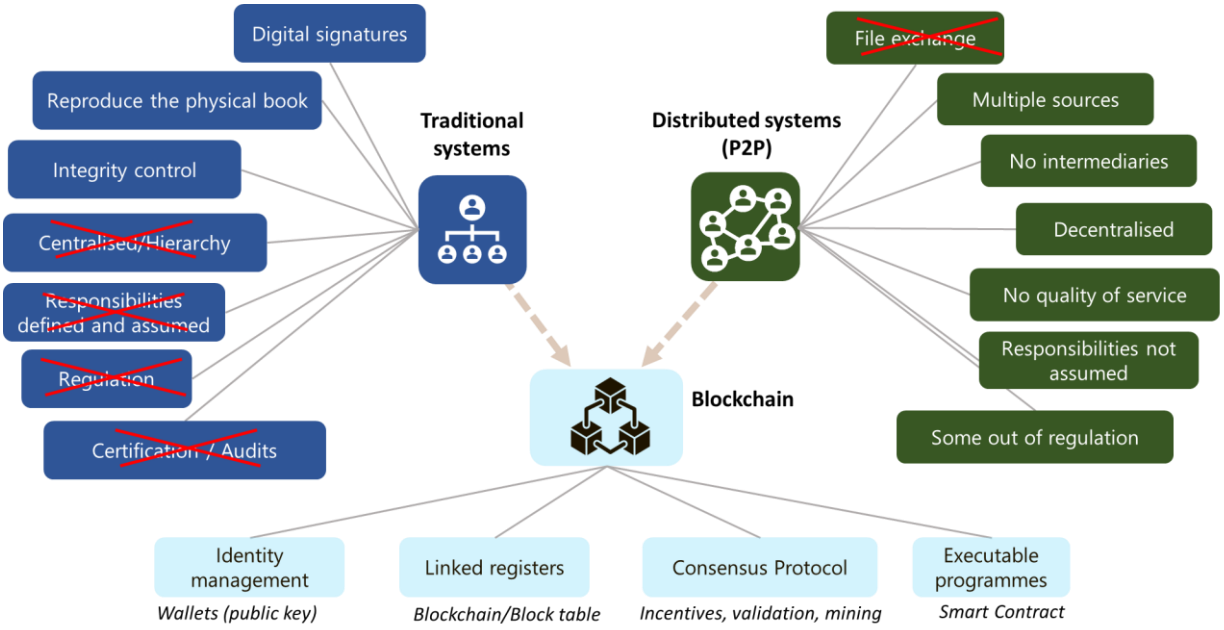


Figure 6. Blockchain inherits some features from traditional systems and many from distributed P2P systems, to which it adds new elements such as identity management, linked records with a hash, consensus protocols and executable programmes (in the figure, features not inherited by generic Blockchain technology are crossed out).

The general principles of Blockchain techniques, which have materialised in the various infrastructures to a greater or lesser extent, are:

- **Decentralisation**: Blockchain participants (nodes) form a decentralised P2P network, where each of them has a replicated copy of the chained table of records. There is no single entity that controls and manages the processing of the data. The nodes are independent, there is no hierarchical relationship between them, and they act on their own behalf. They do not assume any obligations or follow instructions and can stop their operation at any time. In many practical blockchain infrastructures, decentralisation does not exist in some respects or is limited.
- **Integrity control**: Once the data is recorded, cryptographic mechanisms are implemented to detect any modification, deletion or rearrangement of the records.
- **Auditability and transparency**: Transactions are visible to everyone who can access the network. Some infrastructures set limits to this property, e.g. only to participants/nodes or selected third parties.
- **Consensus mechanism:** The consensus mechanism establishes trust between participants/nodes (all or some assigned as mining nodes or validators) to cooperate in maintaining the consistency and integrity of the table and to incorporate new transaction blocks as a record in the table.

Integrity checking, coupled with massive replication of blocks across different nodes, is intended to provide some measure of "immutability" of information. However, the reality is that there are several factors that mean that this immutability can be compromised, as explained in misunderstanding 1.

From the perspective of the access, participation and control policies under which a Blockchain is designed and implemented, Blockchains infrastructures are often classified as public or private, and as permissioned or permissionless.

- A Blockchain infrastructure is public when any participant can freely decide to join it to become part of the infrastructure as a node, while a private one incorporates a governance process that makes it accessible only to a restricted number of participants, usually controlled by a private entity or consortium.
- A permissionless Blockchain infrastructure has no conditions for joining the infrastructure, without restrictions. A permissioned network incorporates governance processes that make it accessible to any participant as long as it passes an authorisation process.

The best-known cryptocurrencies (*Bitcoin, Ethereum, Solana, Tether*, etc.) correspond to the case of public, permissionless blockchain infrastructures. These characteristics make possible one of their main attractions, which is the elimination of intermediaries in the processing of transactions between parties, but it is also one of their main weaknesses when it comes to establishing legal guarantees for their operation.

These weaknesses have become manifest in cases of community and participant disagreements. For example, the Ethereum *DAO Fork*[29], in response to an attack that caused millions in losses, or the *Bitcoin Cash Fork*[30] , which was a proactive decision to address the original limitations of Bitcoin. In both cases, participants had to improvise a governance mechanism which, moreover, resulted in a bifurcation and separation into two distinct Blockchain networks and cryptocurrencies due to differences in decision-making: Ethereum (ETH) and Ethereum Classic (ETC), and Bitcoin (BTC) and Bitcoin Cash (BCH) respectively.

[29] https://ethereum.org/en/history/#dao-fork
[30] https://en.wikipedia.org/wiki/Bitcoin_Cash

The evolution of technology has meant that additional infrastructures have been built on top of the Blockchain infrastructure (in the style of *Over The Top* services in telecommunications infrastructures), which implement additional functionalities or restrictions, known as Layer 2 Blockchain[31] (state or payment channels, *sidechains*, *rollups*, etc).

B. MISUNDERSTANDINGS ABOUT BLOCKCHAIN TECHNOLOGIES AND INFRASTRUCTURES

As noted in the Basic Concepts chapter, when analysing from a data protection point of view the impact that the choice of certain technological options may have on a processing operation in relation to compliance with the GDPR, it is essential to rigorously define the terminology used and to ensure that the implications are accurately understood. The terminology used in the field of Blockchain technology can be misleading for those who do not know the details of the implementation of the technology[32]. In order to reach conclusions as to the precise legal implications, it is necessary to avoid generalities and more or less commercial labels and to establish definitions with rigour, so that such conclusions are based on objectively defined facts[33].

There are many terms used in the Blockchain environment that can lead to misunderstandings about the implications, possibilities and limitations of this technology. Some of the most important ones are explained below.

1. Misunderstanding: Immutability

Immutability is defined in ISO 22739:2024 as the property that data cannot be modified or deleted once it has been added to a distributed ledger. However, rather than a physical or technological property, it is a requirement or objective sought in some Blockchain infrastructures. As the nature of all digital data is volatility, some degree of immutability is sought with technical integrity management measures, which allow for the detection of altered or deleted data, with an assumed agreement among all participants/nodes not to alter the integrity of the already consolidated set of blocks/records, and with the provision for uncontrolled replication of that set of blocks/records in the hope that there will be someone interested in storing the information.

All these circumstances have at some point been altered. For example, the chain may be altered by an agreement between users to delete a block or transaction, which, while making an inconsistency manifest, does not allow the deleted information to be recovered. This is more feasible in private or permissioned Blockchain infrastructures and implies a change in the agreement between the participants/nodes, i.e. a governance measure.

The recalculation of cryptographic values by a majority of users would also allow the alteration of values on the Blockchain. It should be noted that the Blockchain infrastructure is not free from attacks and unwanted uses: malicious behaviour of nodes (if they exceed the majority, known as the 51% attack), vulnerabilities in the code or errors in the implementation can compromise the integrity of the chain.

Simpler is the fact, which has happened many times, of abandon of blockchain project participants/nodes[34], those temporarily set up for education, or the pruning[35] of old blocks. In

[31] https://ethereum.org/en/layer-2/

[32] This is common in any technology, where it is common to use commercial terms, some for marketing purposes, or with a very specific meaning in that context and different from other contexts in which it is used.

[33] Otherwise, it will be difficult to apply legal rigour to an ambiguously defined area, notwithstanding the fact that summaries may be generated for a non-specialist audience.

[34] Disappeared Blockchain Infrastructures are: *OneCoin and Finiko, BitConnect, GetGems, SpaceBIT, PayCoin, etc.*

[35] Removal of some of the Blockchain information stored on a node to save space.

the latter two cases, if governance mechanisms are not in place to provide quality of service guarantees on the availability of the data, the data disappears or becomes inaccessible.

Finally, many blockchains incorporate mechanisms to implement planned updates or improvements to protocols, such as *Bitcoin Improvement Proposals* (BIP)[36] or *Ethereum Improvement Proposals* (EIP)[37], which introduce network changes and modifications. On the other hand, decision-making in the face of unforeseen events in some Blockchain infrastructures, such as the well-known Ethereum *DAO Fork*, has led to changes in the consensus to reverse the effects produced, modifying the state of the information stored.

Therefore, there is no property of immutability in a Blockchain. Data has disappeared on previous occasions.

2. Misunderstanding: Blockchain infrastructure management is completely decentralised

Decentralisation is one of the fundamental principles of the Blockchain technology concept and is promoted as one of its most important features. However, in practice, even public and permissionless Blockchain infrastructures are partially centralised in certain aspects of their management, revealing a concentration of power in very few players.

One of the clearest examples of this concentration is the generation of new blocks/records. In most Blockchain infrastructures, the process of adding blocks to the chain is dominated by a small number of large groups of mining nodes or validators, known as *Pools*. These groups, given their processing capacity or participation in the network (depending on the consensus mechanism used) control a significant part of the capacity for block creation, which creates a real risk of operational centralisation in a system that was conceived to be fully decentralised.

Another case is manifested in the governance decision-making process of Blockchain infrastructures. This is often concentrated in a small group of developers or entities that have very significant influence over key decision-making, such as updates and changes to protocols and software.

3. Misunderstanding: There is no governance framework for blockchain infrastructures, and if there is, it is completely automated.

There is 'de facto' governance in any Blockchain infrastructure, but it is usually incomplete, especially from a data protection perspective. This is because GDPR compliance has not been included among the objectives of the governance framework defined from design.

It is often argued that governance in Blockchain infrastructures is automatic, democratic, fair and equitable, but the reality is quite different. For example, in most Blockchain infrastructures there is a concentration of decision-making in certain groups, such as founders, developers or the most influential community. Also, some consensus mechanisms favour those with the largest stake in each particular Blockchain infrastructure.

In particular, when faced with unforeseen events, some blockchain infrastructures have adopted specific management mechanisms on an ad hoc basis. In other words, they have had to improvise because they have not developed governance based on risk management. This has led to very serious crises in some infrastructures, causing disagreements between

[36] https://github.com/bitcoin/bips
[37] https://eips.ethereum.org/

the participants/nodes regarding the decisions taken in the face of these events, which have led to the splitting of the infrastructure[38].

4. Misunderstanding: In a Blockchain infrastructure, the nodes act in an automated way.

Mining nodes and validators are not, as has sometimes been conveyed, machines that cannot be held accountable in any way.

On the contrary, nodes are made up of a set of resources, selected and configured according to the specific interests of the people who manage them, and made available to a specific Blockchain infrastructure, and in a specific way, by the decision of these managers. They make autonomous decisions on which Blockchain infrastructure to participate in, which means to use[39], which transactions to include in a block, as participants/nodes can order or exclude transactions in a block that they themselves are producing, with the aim of obtaining an additional profit (*Maximum Extractable Value or MEV*)[40], and how to follow updates, among others.

Unless they are acting in the name and on behalf of another entity, they are not bound by specific instructions and, inter alia, may cease to operate at any time.

5. Misunderstanding: All Blockchain infrastructures have the same properties as the ideal model suggested in the original definition of Blockchain technology.

The ideal model originally suggested in Blockchain technology represents what would be an approximation to a concrete implementation of a Blockchain infrastructure. Even the original Bitcoin infrastructure deviates slightly from the ideal principles, e.g. in terms of decentralisation of governance management.

On the other hand, different general-purpose Blockchain infrastructures adapt these principles to their own requirements[41], which makes them incompatible with each other. For example, private-permissioned Blockchain infrastructures can deviate significantly from these principles.

6. Misunderstanding: Code is law

Code is law[42] is a statement that oversimplifies the reality of the practical application of Blockchain technologies.

This assertion is intended to establish that decisions are in the hands of a computer program alone, in particular, so-called *Smart Contracts*. In this way, it defines an environment in which, theoretically, human laws and principles are ineffective and a diversion of responsibility for decision-making and its consequences is realised.

However, *Smart Contracts* are programs made by people, to fulfil objectives defined by those people, and, in addition, the events that initiate the execution of the programs are initiated by people. Moreover, *Smart Contracts*, like all software, have vulnerabilities that are

[38] In addition to others already referenced in this paper, Bitcoin Satoshi Vision (BSV) arose as a result of the clash between two contentious proposals to upgrade Bitcoin Cash, leading to a *hard fork* of Bitcoin Cash.

[39] Connectivity, computing, management and human resources. This goes beyond using software compatible with a particular Blockchain, but also making decisions on which systems he runs the processes (own, in the cloud, outsourced, mixed), how he stores the information (some kind of database system, another kind of dataset, etc.), where he stores the information (in one of the participant's systems, in the cloud, etc.), how many blocks/records will be stored, with what criteria or when access to such information is given, etc., but acts on explicit instructions from a third party.

[40] https://ethereum.org/en/developers/docs/mev/

[41] Different variants may feature different degrees of decentralisation (such as Ripple, with limited decentralisation), different consensus mechanisms, or architectures that optimise performance and scalability (e.g. Solana or Polkadot).

[42] https://ethereumclassic.org/why-classic/code-is-law

exploited by attackers and there are variants of *Smart Contracts* that allow their behaviour to be modified (*Proxy*[43], Oracles[44], etc.).

7. Misunderstanding: Current Blockchain infrastructures guarantee user control over their own data.

The reality is that most Blockchain infrastructures, as they have been designed so far, do not allow the user (natural or legal person) to control who accesses their data, how long their data is stored, how to exercise their rights and for what purposes it will be processed.

This is not a problem with the Blockchain technology itself, but with the way current infrastructures have been designed, as they do not provide for data protection by design or by default.

8. Misunderstanding: Blockchain technology is incompatible with the GDPR

Most Blockchain infrastructures have been built with the purpose of being unregulated. Of course, if a processing operation or set of operations are designed to be non-compliant, it will be very difficult to adapt them to be so. In the case of Blockchain infrastructures, this is generally the case both with data protection regulations and with other regulations such as tax or commercial regulations.

The fact that compliance with the GDPR has not been contemplated as one of the objectives of a Blockchain infrastructure is not a problem of the technology itself, it is a problem of the objectives and decisions of the designers who have built it. If they had contemplated regulatory compliance objectives from the design stage, the appropriate management mechanisms would have been implemented. And this is independent of whether we are dealing with any type of Blockchain infrastructure.

9. Misunderstanding: Smart Contracts are autonomous and intelligent.

Smart Contracts are programs stored in the Blockchain infrastructure in which the result of any execution of the program is recorded on the infrastructure itself[45]. They are automated programs that execute predefined actions when certain programmed conditions are met, executed when invoked through transactions. Their operation is limited by their code and the data they receive, and there are also tools that allow the *Smart Contract* to update the internal states of the Blockchain based on external information (such as the so-called oracles)[46].

That is, they are not "intelligent" in the sense of possessing autonomy or decision-making capacity. Nor can it be considered as a contract in the strict legal sense.

The implementation of *Smart Contracts* in scenarios where personal data are handled falls within the scope defined by Article 22[47] of the GDPR and should therefore be carefully assessed from a compliance point of view, particularly with regard to the protection of individuals' rights against automated processing of their information.

---

[43] A *Smart Contract* of type *Proxy* delegates calls to other target *Smart Contracts*, allowing the update of these target *Smart Contracts*. The *Proxy* has as an updateable variable the address of the target *Smart Contract*.

[44] Services providing real-world data to *Smart Contracts* https://cointelegraph.com/learn/what-is-a-blockchain-oracle-and-how-does-it-work)

[45] ISO 22739

[46] ISO 22739: Service that updates the status of the Blockchain infrastructure using data external to the Blockchain/table.

[47] Article 22 of the GDPR provides that individuals have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning them or similarly significantly affects them, with some exceptional situations.

10. Misunderstanding: In a Blockchain infrastructure the data is only present in transactions and blocks.

The idea that in a Blockchain infrastructure data is stored only in the blocks/records and the transactions within them is incomplete. This is only true in the theoretical description of Blockchain technology, but the actual implementation of a concrete infrastructure is more complex.

For example, nodes, in addition to Blockchain/table, need to store a variety of additional data, including the *Smart Contracts*' own storage and transaction receipts involving calls to their functions and procedures. These transaction receipts, known as *receipts* or *logs*, are records that are stored or linked in the Blockchain infrastructure and contain information about the outcome of the transaction and the events issued by the *Smart Contract*. These events (programmed in the *Smart Contract*) facilitate communication between the *Smart Contracts* and the user interfaces of the applications deployed on the Blockchain infrastructure, and their purpose is to return values to the user interface to trigger actions on it, or as a log or cheap form of storage (taking into account the costs associated with the transactions). In this way, the Blockchain infrastructure is not only limited to managing financial transactions, but also becomes a platform for storing additional information related to the execution of *Smart Contracts*.

Another example is *off-chain* storage. *Off-chain* storage involves storing some of the information that would ideally be in the blocks/records, in a different data structure than the Blockchain/table. The linkage between the two sets of data would be stored in the block transactions. The data structure that makes up the *off-chain* could be database technology, or another type of blockchain/records. In addition, it could also be replicated across all nodes/participants, in only some of them, or in a centralised storage.

In any case, *off-chain* storage is part of the Blockchain infrastructure. The processing for the management of such storage, and those that are implemented on such storage, also using storage, must ensure and be able to demonstrate compliance with the GDPR. In other words, the GDPR compliance requirements do not disappear by moving personal data from one data structure (Blockchain/table) to another. Moreover, it must be demonstrated that there is no personal data in the Blockchain/table or in other data structures as outlined in this section.

Finally, another example is all the information that in temporary and working copies is stored in participants/nodes, in their outsourced services, and in any other player in the ecosystem that allows a Blockchain infrastructure to function.

C. PERSONAL DATA

The use of a Blockchain infrastructure for data processing generally requires the user (whether an individual or a legal entity) to have an account. The minimum requirement for having an account is to be in possession of a public/private key pair that allows him/her to sign transactions and identify him/herself on the Blockchain.

Pieces of software, outside the infrastructure called *wallets*[48], can be used to store accounts and perform transactions for a user. These wallets would be part of the ecosystem application set of a particular Blockchain infrastructure.

[48] ISO 22739: Application or mechanism used to generate, manage, store or use private keys (3.76) and public keys (3.79) or other digital assets. A wallet can be implemented in software, implemented as a hardware module, or written onto non-digital media such as paper or metal. Digital assets stored in wallets may include, for example, non-fungible tokens (NFTs).

Where a user is a natural person, the public key or address of a wallet is personal data, as it is any unique identifier that links all activity of that person on the Blockchain infrastructure[49].

In a Blockchain infrastructure personal data can be found at[50]:

- Transactions (sender account, destination account and transaction data).
- Account balances.
- *Smart Contracts* storage (programs stored on the Blockchain that execute actions on their own storage and other *Smart Contracts* when invoked by a transaction).
- Transaction receipts (*Smart Contracts* events/logs).
- *Off-chain* storage.
- Any other type of storage owned by the participant/node beyond the actual copy of the Blockchain/table, and temporary storages.

D. PROCESSING AND RESPONSIBILITIES

A Blockchain infrastructure allows certain data processing to be implemented on it, such as, for example, the consolidation of balances between a company branches, certificate management, or the accreditation of academic certifications. These processing use the Blockchain infrastructure as part of the operations that comprise the processing. Additional or intervening systems include wallets[51], *exchanges*, decentralised applications[52] (*dApps*), interface and APIs providers to interact with the Blockchain[53], oracles[54] (services that provide real-world data to *Smart Contracts*), mining/validation node pools[55], internet networks, etc.

In turn, the infrastructure itself, regardless of the number of processing that are implemented on it, will require the implementation of a series of processing for its own management[56] such as:

- Providing the service to the end user: the set of activities and resources to meet the user's needs and provide the appropriate quality of service.
- Conducting transactions on the Blockchain: transformation of data into transactions valid for the Blockchain.
- Validate or mine transaction blocks: verify and add transactions to the Blockchain according to the consensus mechanism.
- Manage the Blockchain according to the consensus mechanism: verify and store blocks, apply updates, access, etc.
- Implement any additional processing that a node performs on transactions, such as reordering transactions, managing pending transactions (*mempool*), providing

[49] There are many examples. Attached is a statement from the Royal Canadian Mounted Police which, in a Bitcoin scam money recovery action, states: "This investigation demystified the fact that cryptocurrency transactions are completely anonymous and that there's no chance of recovery of the victim's funds: https://rcmp.ca/en/gazette/police-help-victim-crypto-fraud-get-money-back

[50] Even when the user is not a natural person, but the actions recorded on the chain can be linked to a natural person by data that is part of the processing that is implemented on the Blockchain.

[51] In addition to storing cryptographic keys, the *wallets* also allow transactions to be carried out on the Blockchain https://en.wikipedia.org/wiki/Cryptocurrency_wallet.

[52] https://en.wikipedia.org/wiki/Decentralized_application

[53] For example, nodes as a service: https://ethereum.org/en/developers/docs/nodes-and-clients/nodes-as-a-service/

[54] https://es.cointelegraph.com/learn/what-is-a-blockchain-oracle-and-how-does-it-work, https://ethereum.org/en/developers/docs/oracles/

[55] Ethereum *Staking* Pools: https://ethereum.org/en/staking/pools/, Bitcoin Mining Pools: https://miningpools.com/bitcoin

[56] In the same way as other digital infrastructures, for example, the case of telecommunications networks, but also in other non-digital infrastructures, such as in the case of logistics companies. These are companies that handle many contracts or are organised in the form of cooperatives, that implement processing for their own management (such as HR), and that are used by other companies to implement other processing, such as Internet sales, distribution of medicines, etc.

data and services to applications (decentralised applications or *dApps*), storing historical data, processing and handling queries, processing *Smart Contracts* events, detecting patterns and behaviours, etc.

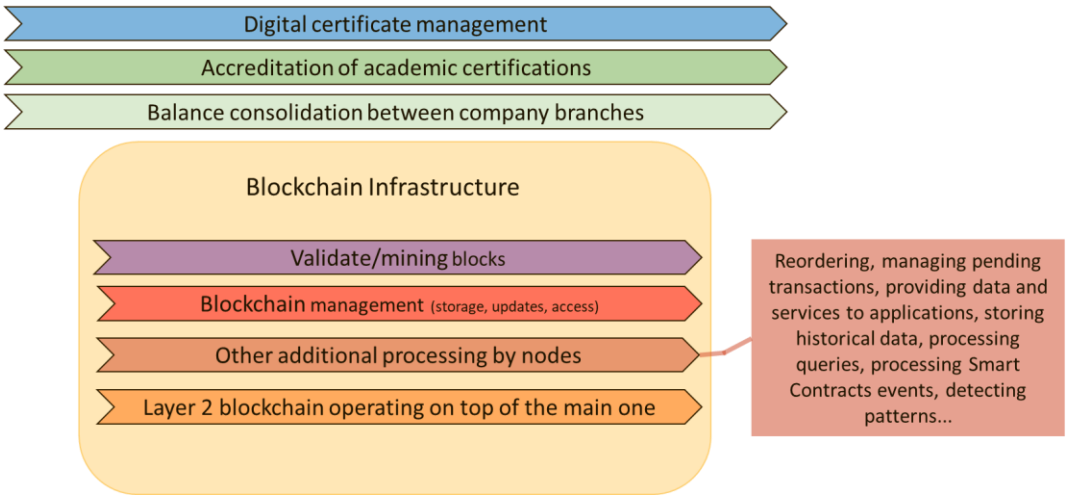


Figure 7. Processing in and on a Blockchain infrastructure.

In any Blockchain infrastructure it is crucial to clearly establish the responsibilities and obligations of each participant in the network, taking into account that any entity that processes personal data and does not do so on behalf of a controller cannot be a data processor[57].

E. GDPR COMPLIANCE

Any processing of personal data must comply with the GDPR. The design and characteristics of most of the currently implemented Blockchain infrastructures present significant challenges to comply with the GDPR principles for processing executed in or on the Blockchain infrastructure. This is due to the reasons explained below.

Firstly, the designs of the infrastructure have not taken into account the application of data protection by design. While Article 25 of the GDPR applies to controllers, it should be noted that Recital 78 of the GDPR states that "*When developing, designing, selecting and using applications, services and products that are based on the processing of personal data or process personal data to fulfil their task, producers of the products, services and applications should be encouraged to take into account the right to data protection when developing and designing such products, services and applications and, with due regard to the state of the art, to make sure that controllers and processors are able to fulfil their data protection obligations*." This does not shift obligations from controllers to services and solution providers. Controllers have the obligation to select services and solutions that enable them to '*ensure and be able to demonstrate that processing is performed in accordance with this Regulation*' (Article 24(1) of the GDPR).

On the other hand, a large part of Blockchain infrastructures are generally built with open source modules or components[58]. The fact of using open source should not imply, as is often the case, that these components are not properly documented in their design and in relation to the tests that allow them to demonstrate what is established in the previous paragraph. In fact, many of these components are open source but very opaque in their functionalities, include a high level of unidentified dependencies, and it is very difficult to update them or

[57] Paragraph 76 of the Guidelines 07/2020 on the concepts of "controller" and "processor" in the GDPR.
[58] (there may be some components or enterprise versions that do not).

even guarantee their functioning. The latter is especially critical when it comes to ensuring and being able to demonstrate compliance with the GDPR (Article 24(1) of the GDPR), implementing data protection by design (Article 25 of the GDPR) and ensuring the resilience of the processing built upon them (Article 32(1)(b) of the GDPR).

Complying with the GDPR when processing personal data using one or more Blockchain infrastructures requires two steps: 1) ensuring compliance and, once compliance has been achieved, 2) assessing and evaluating risks, some of which can be mitigated by legal, organisational and technical measures. In case of high risk, the assessment of the appropriateness, necessity and proportionality of the processing has to be passed.

In the case of Blockchain infrastructures, the aspects of particular relevance in relation to GDPR compliance are those relating to accuracy and storage limitation. This is due to the lack of governance measures that engage participants beyond those that are automated in the code. In addition, these measures do not include inconsistency management, which makes it difficult to modify data once it is included in a block/record of the Blockchain/table. These circumstances make it very difficult to exercise the rights of rectification and erasure.

To solve these problems, a governance framework has to be defined, management processes developed, and the infrastructure components adapted to facilitate the exercise of stakeholders' rights.

It should not be overlooked that Blockchain infrastructures already have mechanisms in place that may involve changes to the protocol, consensus rules and other aspects (including modifications as occurred in the Ethereum *DAO Fork* mentioned above) that take the form of software updates, which should be adopted by participants/nodes. These changes, also called improvement proposals, are, for example, the *Bitcoin Improvement Proposal* (BIP)[59] and the *Ethereum Improvement Proposal* (EIP)[60] . The process of managing these changes is not carried out through formal procedures, but they are discussed and assessed in public forums, mailing lists, social networks, by the community, core developers and other important stakeholders such as participants/nodes (in some cases changes are supported and registered on the Blockchain by a subset of participants, those who manage mining nodes).

Another factor to take into account is that the offshoring of nodes may involve international transfers of personal data, which have to comply with the provisions of the GDPR in this respect.

Finally, and as already mentioned, *Smart Contracts* are programmes that are stored in the Blockchain and that execute automated decisions (those for which they have been programmed). Insofar as these decisions may significantly affect natural persons, it is essential that the requirements established in Article 22 of the GDPR are complied with from the design stage, and that the necessary guarantees and measures are incorporated to protect the rights of data subjects.

[59] https://github.com/bitcoin/bips
[60] https://eips.ethereum.org/

IV. AEPD PROOF OF CONCEPT

The developed Proof of Concept shows, in full detail, the use case of deleting the activity of a user (natural person) in a Blockchain infrastructure. This will be done by deleting the address of an account, which is an identifier and therefore a personal data. The chosen approach involves either overwriting the account address, or the signature of the transaction from which the address can be derived.

Particularly, to ensure the rights to rectification and erasure, which involves the deletion and updating of transaction data in the Blockchain/table, governance mechanisms and technical measures will be needed to materialise them.

A. BACKGROUND

A number of research and projects have already explored strategies for developing "*redactable*" Blockchain infrastructures, i.e. allowing data to be edited or deleted under certain circumstances, for the purpose of regulatory compliance or error correction.

These projects have had certain limitations and have been based on designing protocols that support editing and modification of transactions through operations carried out by authorised participants and seek to maintain consistency at transaction or block level. Chameleon hashes and cryptographic variants, modification of the block data structure, mutable transactions (multiple versions), *off-chain* storage, local erasure, ZKP techniques, pruning techniques or techniques based on consensus mechanisms are some of these strategies[61]. However, they have not been implemented in the most popular blockchain infrastructures.

B. EXPLOITING EXISTING STRATEGIES

The AEPD's Proof of Concept (PoC) does not propose a disruptive change to ensure compliance with the GDPR but is based on using existing and well-known strategies of Blockchain infrastructures and distributed data storage models.

1. Governance

Governance involves exercising authority and control to decide the objectives of an organisation, making decisions, based on assets, resources, context and risk management, to achieve those objectives in a prioritised and balanced way, and continuously monitoring

[61] Solanki, A. R. (2024). Redactable Blockchain Solutions for IoT: A Review of Mechanisms and Applications https://engrxiv.org/preprint/view/3792/6703

Weiqi Dai et al, (2024) PRBFPT: A Practical Redactable Blockchain Framework With a Public Trapdoor https://ieeexplore.ieee.org/document/10380599,

Shams Mhmood A.A et al (2023). Redactable Blockchain: Comprehensive Review, Mechanisms, Challenges, Open Issues and Future Research Directions, https://www.mdpi.com/1999-5903/15/1/35,

Xiayu Wang et al (2024). A Redactable Blockchain Scheme Supporting Quantum-Resistance and Trapdoor Updates https://www.mdpi.com/2076-3417/14/2/832,

Xin-Yu Li et al (2021). Escaping from Consensus: Instantly Redactable Blockchain Protocols in Permissionless Setting https://eprint.iacr.org/2021/223.pdf,

Damiano Sartori. University of Twente (2020). Redactable Blockchain. How to change the immutable and the consequences of doing so. http://essay.utwente.nl/82755/1/Sartori_MA_EEMCS.pdf,

Yueyan Dong et al (2023). Redactable consortium blockchain with access control: Leveraging chameleon hash and multi-authority attribute-based encryption https://www.sciencedirect.com/science/article/pii/S2667295223000661

Dominic Deuber et al. University of Manchester (2019). Redactable Blockchain in the Permissionless Setting. https://pure.manchester.ac.uk/ws/portalfiles/portal/211464559/Redactable_Blockchain_in_the_Permissionless_Setting.pdf

that the progress of each of the actions taken is on track[62]. Governance has to be implemented by defining roles, policies, procedures, plans, organisational, legal and technical measures to manage the organisation.

The success of a governance framework is measured by the long-term success of the organisation. A factor in this success will be the clear and *accountable* definition of the elements outlined in the previous paragraph. These include the identification of objectives, definition of decision-making and management roles, well-defined decision-making and implementation procedures and their traceability, documentation, etc. Depending on the objectives, we can speak of corporate governance, strategic governance, data governance, sustainability governance, AI governance, etc. However, in an organisation there is only one governance, which will encompass several aspects, and which will be translated into a single management, which will be oriented towards achieving each of the objectives.

Governance is a basic element in any implementation of a Blockchain infrastructure and also a differentiating factor in each of them. In fact, the main classification of Blockchain infrastructures is based on one aspect of their governance: public/private and permissioned/permissionless.

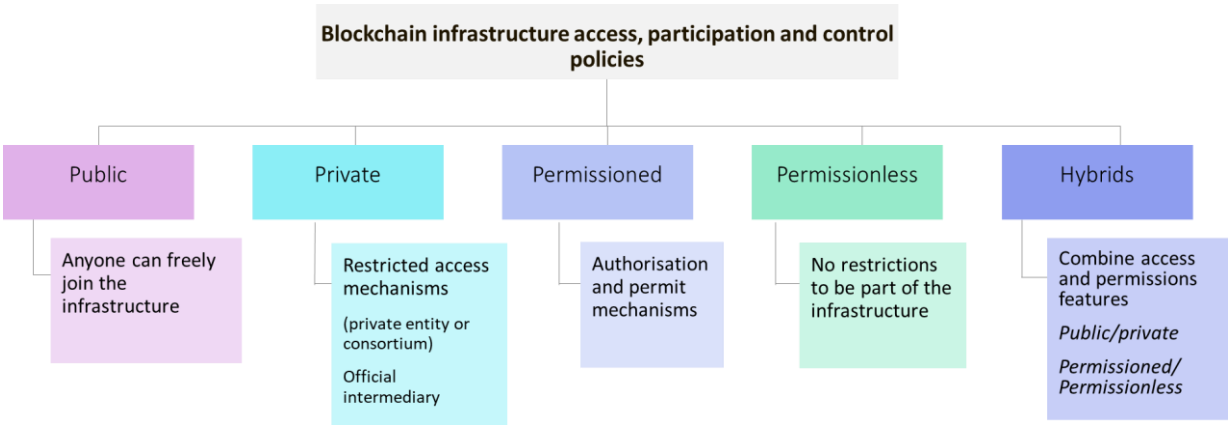


Figure 8. Classification of Blockchain networks according to access, participation and infrastructure control policies. It is common to find public-permissionless (Bitcoin, Ethereum, and most of the known cryptocurrencies) and private-permissioned (by private entities and consortiums) networks.

Even permissionless public Blockchain infrastructures include elements of governance without which they would not be able to operate. First, a more or less distributed allocation of roles over who can make decisions to alter the Blockchain. Then, more or less assembly-based decision-making procedures. A definition of framework objectives for the infrastructure[63] and other lower-level management procedures (consensus mechanisms, permitted code versions, register of updates, etc.).

[62] ITIL4 "the framework of authority, accountability, and decision-making required to achieve an organisation's objectives and manage risks appropriately", COBIT 2019 "ensures that stakeholder needs, conditions and options are evaluated to determine balanced, agreed-on enterprise objectives to be achieved; that direction is set through prioritisation and decision making; and that performance and compliance are monitored against agreed-on direction and objectives".

[63] In the case of Bitcoin, Nakamoto's 2008 email set the main goal of having a currency and economic exchanges outside the control of a central authority. In his original publication https://bitcoin.org/bitcoin.pdf, he devotes chapter 10 to privacy, considering that all privacy issues were solved by using public keys, and inferring from this that transactions were anonymous. This has turned out to be a mistake. The reality is that encryption is not anonymising, not least because public keys are unique identifiers. The facts have shown that the authorities re-identify users, and there are even companies that provide such services.

All these elements will be more or less documented, more or less explicit[64], supported by legal, organisational or technical measures, more or less automated, and will have been established by a risk management process or created on the basis of events[65].

The fact that a Blockchain infrastructure does not include compliance management as one of its objectives is not a problem with the technology itself. But if this objective is not contemplated, the appropriate management mechanisms are not implemented from the design stage. And this is irrespective of whether we are dealing with any type of Blockchain, whether private, public, permissioned or not. In this regard, Recital 78 of the GDPR explains that "*In order to be able to demonstrate compliance with this Regulation, the controller should adopt internal policies and implement measures which meet in particular the principles of data protection by design and data protection by default"*. Therefore, at least from the point of view of the GDPR, the governance of a Blockchain infrastructure must include data protection compliance objectives and therefore define policies that develop the management processes that implement them.

2. Hard Fork/ Soft Fork and software updates

In the context of a blockchain infrastructure, a *fork*, i.e. a split of the Blockchain/table, is the main mechanism by which software updates are implemented. It can occur for various reasons, such as modifications to the protocol, modifications to the code to fix bugs or add functionality, decision making in the face of specific unforeseen events, or other changes.

In some cases, new software updates do not directly affect the operation and previous versions are compatible, i.e. those participants/nodes that do not update their software can continue to operate, this is known as a *Soft Fork*. For a *Soft Fork* to be successful, it is necessary for most of the mining/validator nodes to be upgraded.

On the other hand, *a hard fork* occurs when the software update is not compatible with previous versions and, therefore, all nodes must update their software in order to continue operating. In case not all nodes update their software, separate and incompatible Blockchain/tables will be created, giving rise to new infrastructures, as has already happened on several occasions, mainly due to disagreements and conflicting interests.

The mechanisms for introducing changes through *Bitcoin Improvement Proposals* (BIPs) and *Ethereum Improvement Proposals* (EIPs), which are included in the corresponding software updates, sometimes implementing more than one BIP or EIP in each update, have been mentioned above.

The official Ethereum client currently has more than 100 EIPs built in, around 60 of which are consensus-related, and has had around twenty *Hard Fork*[66] versions over its history (ten years), i.e. between one and three times a year. On the other hand, the official Bitcoin client has incorporated more than 60 BIPs, 17 of them related to consensus in its fourteen-year history.

Hard forks are part of the process of developing and upgrading Ethereum. They generally do not result in separate Blockchain/tables, as the community usually adopts the updates. In Bitcoin all major updates so far have been deployed as *Soft Forks*, several of them with major changes to the consensus. In any case, a mining node or validator using older versions of a

[64] For example, any Blockchain infrastructure, as in any business sector, requires those who want to participate in it (finance, telecommunications, pharmaceuticals, transport, etc.) to comply with regulations, standards and de facto agreements in order to operate.

[65] Risk management involves identifying what could go wrong in the future and putting in place the means to manage setbacks. In the case of the Ethereum DAO Fork (https://ethereum.org/en/history/#dao-fork), it was not foreseen that a Smart Contract could be manipulated, and when it was, a procedure had to be improvised to manage such an event with the negative consequences it entailed.

[66] https://ethereum.org/en/history/

Soft Fork may have limited functionality and risk of rejection of the blocks it generates by those who have upgraded, a risk that will be greater if the *Soft Fork* is relative to the consensus.

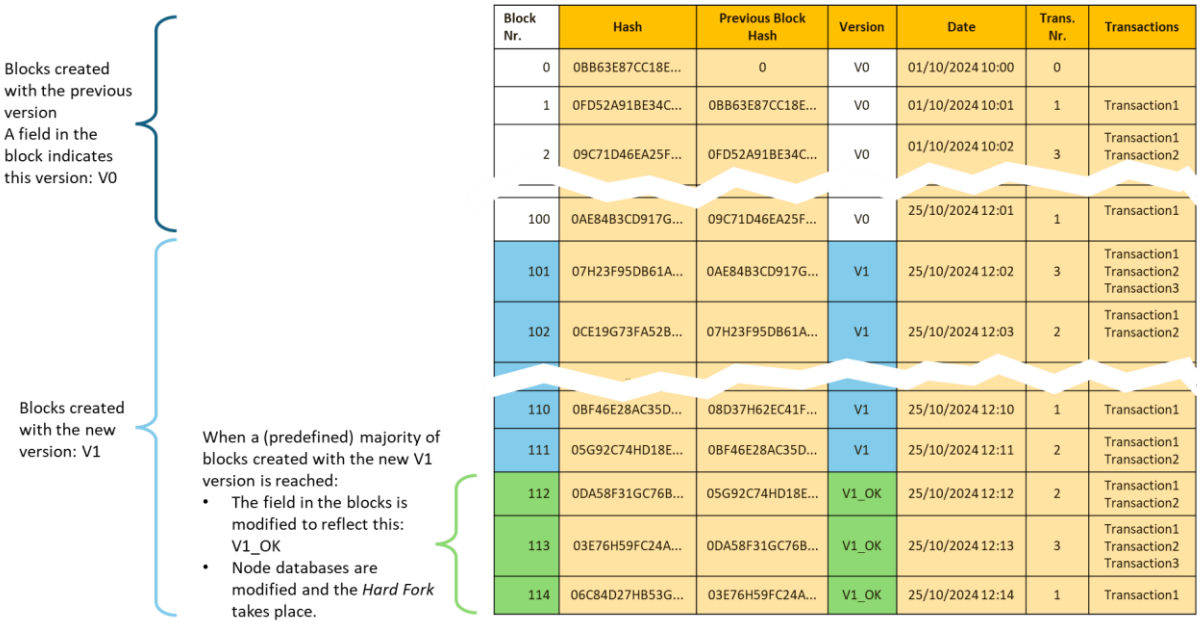
The PoC presented in this document makes use of the *Hard Fork* mechanism to implement the changes to the node databases necessary for the execution of the right to erasure, so that the new version of the Blockchain infrastructure does not contain the personal data subject to erasure.

3. Validation procedures for new versions

The *Bitcoin Improvement Proposal* BIP-0009[67] is one of the mechanisms implemented in this infrastructure to validate a new version of the software (*Soft Fork*).

The PoC presented here incorporates the BIP-0009 mechanism in a simplified way. When a validator node updates to the new version, the blocks it generates incorporate an indicator in one of their fields. When, in a predefined interval of the last blocks added to the chain, a certain number of them (also a predefined majority) contain that indicator, then it is concluded that the majority of the validator nodes have updated to the new version (they agree with the changes) and the change is accepted.

Once this majority is reached, the new software version causes each validator node to proceed to modify its local database, making the right to erasure effective, thus completing the *Hard Fork*. Thereafter, the block field that has been used to indicate the new version is updated, and in the new blocks generated this field also contains the information that the update has been accepted and carried out. A node that synchronises with the network subsequently verifies the information in that field and modifies its local database as well.



Block Nr.	Hash	Previous Block Hash	Version	Date	Trans. Nr.	Transactions
0	0BB63E87CC18E...	0	V0	01/10/2024 10:00	0	
1	0FD52A91BE34C...	0BB63E87CC18E...	V0	01/10/2024 10:01	1	Transaction1
2	09C71D46EA25F...	0FD52A91BE34C...	V0	01/10/2024 10:02	3	Transaction1 Transaction2
100	0AE84B3CD917G...	09C71D46EA25F...	V0	25/10/2024 12:01	1	Transaction1
101	07H23F95DB61A...	0AE84B3CD917G...	V1	25/10/2024 12:02	3	Transaction1 Transaction2 Transaction3
102	0CE19G73FA52B...	07H23F95DB61A...	V1	25/10/2024 12:03	2	Transaction1 Transaction2
110	0BF46E28AC35D...	08D37H62EC41F...	V1	25/10/2024 12:10	1	Transaction1
111	05G92C74HD18E...	0BF46E28AC35D...	V1	25/10/2024 12:11	2	Transaction1 Transaction2
112	0DA58F31GC76B...	05G92C74HD18E...	V1_OK	25/10/2024 12:12	2	Transaction1 Transaction2
113	03E76H59FC24A...	0DA58F31GC76B...	V1_OK	25/10/2024 12:13	3	Transaction1 Transaction2 Transaction3
114	06C84D27HB53G...	03E76H59FC24A...	V1_OK	25/10/2024 12:14	1	Transaction1

Figure 9. Schematic of the simplified BIP-0009 mechanism used in the PoC. It shows the Blockchain/table that any node stores (the same thing happens in all of them). In each block a field indicates the status of the version used. When the predefined number of last blocks generated with the new version is reached, the change is agreed, the node modifies its database and the new blocks it generates reflect the new agreed state of this new version. The Hard Fork has been triggered.

[67] https://github.com/bitcoin/bips/blob/master/bip-0009.mediawiki

4. Traceability

Data traceability in any information sharing environment is the ability to know the entire data lifecycle. Traceability is the management process that answers the questions of who, when, how, where and why data are processed. Traceability is essential to enforce a right to erasure in complex organisations and processing operations with multiple actors (different controllers, multiple processors and sub-processors, etc.). In such cases, it is essential to know in which different repositories the data is located, in which temporary copies, and in which different sites, processors or sub-processors (Article 28(3)g of the GDPR). Traceability management, among other processes, underpins the legal[68] and organisational measures that oblige each of the actors (and their departments or branches) to effectively execute the right to erasure.

Traceability of individual data, or of the data set, is one of the essential functionalities in a Blockchain infrastructure. Even in permissionless public infrastructures, knowing who is processing the data is necessary for the implementation of certain consensus mechanisms, to determine whether there is control of the infrastructure by a majority group, or to provide measures to guarantee accessibility to the data (even with *Best Effort* procedures[69]), among others. In other types of Blockchain infrastructures, they will fulfil other functions such as permission management, service billing or membership control.

Therefore, traceability mechanisms are another necessary management tool to implement the governance framework in a Blockchain infrastructure.

In the online environment, it is also mandatory to be able to determine which data controllers are processing data made public (Recital 66[70] of the GDPR). This obligation comes in response to cases that arise in P2P networks, such as Blockchain infrastructures. A typical case occurred in Spain when a medical centre disseminated a file containing highly sensitive patient information[71] through a popular P2P network[72]. This meant that the information was distributed among people[73] who had joined the permissionless public network and decided to download the file. The data controller (the medical centre) had not implemented a mechanism to control the distribution of such data to different controllers (who had no legal basis to process such content, CJEU 62006CJ0275[74]), and of course there was no controller-processor relationship between the participants in the P2P network. Each participant in the P2P network had freely chosen to participate in the sharing of content

[68] Regardless of whether Blockchain and other data sharing technologies are used, it is the legal measures that impose obligations on data controllers to comply with the fundamental data protection principles set out in the GDPR. Among these measures, contracts between the controller and processors must ensure the deletion of data when required by the controller.

[69] https://en.wikipedia.org/wiki/Best-effort_delivery

[70] Recital 66 of the GDPR: *"in the online environment, the right to erasure should also be extended in such a way that a controller who has made the personal data public should be obliged to inform the controllers which are processing such personal data to erase any links to, or copies or replications of those personal data. [2]In doing so, that controller should take reasonable steps, taking into account available technology and the means available to the controller, including technical measures, to inform the controllers which are processing the personal data of the data subject's request."*

[71] Procedure PS/00059/2008

[72] eMule, which differs from a Blockchain infrastructure in its consensus mechanisms. All content was identified with a hash, which gave it integrity control at the level of individual content, but not at the level of the content as a whole or its chronological order.

[73] Natural or legal persons who had decided to put in place proprietary systems, connections and programmes to share content, some of which were subject to restrictions on use and for which they could be considered liable under different regulations, for example, industrial or intellectual property.

[74] https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:62006CJ0275 rejection of the obligation to provide personal data relating to internet use in order to protect copyright. The judgment considers that Promusicae was attempting to exercise a legitimate right against persons who had infringed copyright law in the context of civil proceedings. The IP addresses of the machines made it possible to identify the persons responsible for these machines and processing. These persons were acting on their own behalf in infringing copyright law and not on behalf of third parties. Paragraph 43 "*It should be observed to begin with that the intention of the provisions of Community law thus referred to in the question is that the Member States should ensure, especially in the information society, effective protection of industrial property, in particular copyright, which Promusicae claims in the main proceedings*".

(purposes and means) and was acting on his own behalf, not on the instructions of the medical centre[75].

5. The application of rights in the GDPR

Article 12.3 of the GDPR establishes a maximum period of one month for the controller to take action on an exercise of rights request, which may be extended to a total of three months if necessary, taking into account the complexity of the process of exercising these rights:

> *3. The controller shall provide information on action taken on a request under Articles 15 to 22 to the data subject without undue delay and in any event within one month of receipt of the request. That period may be extended by two further months where necessary, taking into account the complexity and number of the requests.*

This is a factor that must be taken into account when considering strategies for the exercise of rights in the processing of or based on Blockchain infrastructures, both in the maximum time limits established and in the consideration of certain approaches to the effective application of these rights. In this case, it would mark the maximum time limit for the execution of a *Hard Fork*.

C. PROOF OF CONCEPT DESIGN

The PoC implementation approach is based on the use of the *Hard Fork* mechanism to implement inconsistency management, allowing the system to continue to operate with guarantees.

For the development of the PoC, the official Ethereum implementation[76] has been used, in its version 1.13.15 of April 2024, which incorporates a consensus protocol based on Proof of Authority (called '*clique*'[77] , supported until the indicated version). In this configuration, the validator nodes are previously known and authorised, from the beginning (although it implements an approval mechanism to add validator nodes once the Blockchain infrastructure is up and running).

To do this, firstly, an analysis had to be made of the existing documentation and source code, as well as the implicit and explicit management procedures that implement the governance defined in Ethereum.

Based on this, a goal has been transferred to Ethereum's governance framework: compliance with the GDPR, in particular that of exercising the right to erasure. For this purpose, the management, organisational and technical tools necessary for its implementation have been developed. These tools are:

1. Procedure for detecting the affected records in the different nodes. This includes not only the transactions stored in blocks, but also the storage of *Smart Contracts* and the receipts of the transactions involving them, *receipts/logs*, as explained in chapter III section B.9.
2. Procedure for generating a new software version of the Blockchain infrastructure, implementing a *Hard Fork*, as explained in section B.2 of this chapter. The

[75] The AEPD set up a team to locate the hashes of the files on the network and to order those responsible under its competence to remove them from their systems, and to request those outside its competence, for example, because they were in another country, to remove them.

[76] https://geth.ethereum.org/

[77] Clique is the standard implementation of the PoA consensus in Ethereum, according to EIP-2-5. https://eips.ethereum.org/EIPS/eip-225

technical strategy employed consists of implementing in the source code a consensus mechanism of the validator nodes in the new version of the Blockchain infrastructure, based on the BIP-0009, together with the modification of the node's local database when agreement is reached on the new version, where the user's right has been fulfilled. In addition, the subsequent synchronisation of new nodes with the new version is envisaged.

3. Procedure for distributing the new software version and running it on the nodes.
4. Technical strategy to implement a consensus mechanism in the new version of the Blockchain infrastructure.
5. Organisational measures to be incorporated in the management of the Blockchain infrastructure to implement the governance objective of compliance with the GDPR.

1. Procedure for detecting affected records

The PoC implements a mechanism for detecting the affected records in the database that make up the Blockchain/table in the different nodes. The databases used by the nodes store information in key-value pairs, unlike traditional relational databases. The process focuses on identifying what information needs to be deleted and replacing it.

This procedure consists of the following steps:

- It searches the blocks for transactions involving the account of the user who wants to exercise his right to erasure.
- Checks whether any of these transactions are sent to *Smart Contracts*, i.e. involve calls to some of their functions, or whether it is the *Smart Contract* creation transaction.
- Obtains *Merkle* trees[78] from the state and balance of the accounts.
- Overwrites the address of the account to be deleted in the affected transactions.
- Overwrites the address of the account to be deleted in the storage of the affected *Smart Contracts*.
- Overwrites the address of the account to be deleted in the *Smart Contract* transaction *receipts/logs*.
- Overwrites the address of the account to be deleted in the *Merkle* trees of the state and balance of accounts.
- Finally, it saves the modified database in a JSON data structure, which contains the modified key-value pairs, where the address of the account to be deleted no longer appears. This data structure will be incorporated in the *Hard Fork* software update generation procedure.

For simplicity, this PoC performs this procedure by running on an auxiliary node that contains a copy of the original database. The execution of the auxiliary node facilitates the obtaining of blocks, transactions and their data through queries to the API that a node exposes. Additionally, by executing a program developed in NodeJS, read and write libraries are used in the node's databases (*LevelDB* database) to access and modify the rest of the necessary data. This approach allows the necessary modifications to be made without directly affecting the main network until the *Hard Fork* update is executed.

[78] ISO 22739: Tree data structure in which every leaf node is labelled with the hash value of a data element and every non-leaf node is labelled with the hash value of the labels of its child nodes.

2. Procedure for the generation of a new Blockchain infrastructure software version

Once the affected records have been detected, the official Ethereum source code (*geth*[79]) must be modified to generate the software version update. To do so, the following steps are followed:

- Add logic and functions to implement the BIP-0009 inspired agreement and modifications to the node databases.
- Add the key-value pairs to be replaced in the node's database, obtained in the previous procedure.
- Compile the new software version and distribute it.

The first stage in this procedure includes adding new functionalities to the official Ethereum source code, as well as the relevant updates to manage the modification of the node's local database and to manage the subsequent synchronisation of a node. These changes can remain in successive updates of the software, where only the version number would need to be taken into account.

The second stage involves including, as part of the source code, the modified key-value pairs that will replace the corresponding original key-value pairs in the node's local database. This operation will be performed when the BIP-0009 inspired agreement is verified, or when a node synchronizes with the chain where the agreement has been previously reached.

3. Procedure for distributing the new software version and running it on the nodes.

Once the above procedure is completed, the new software version will be available. This new version is made available to all nodes to run. The nodes must be aware of the software updates in order to execute them, as they will involve a *Hard Fork*. Traceability, as indicated in section B.4 of this chapter, underpins the legal and organisational measures that oblige each node to effectively execute the right to erasure.

In the PoC context the new version is distributed manually to each node. Since the nodes use the same operating system and are configured in the same way, the update process is simplified, as the modified version can be copied directly to each node.

4. Technical strategy for implementing a consensus mechanism in the new version of the Blockchain

A second fundamental pillar on which the PoC is based, once the records to be modified have been detected, is a mechanism for agreement among the validator nodes to adopt this software update by consensus and majority, registering it on the Blockchain itself in the updated version of the software that these nodes run.

As described above (section B.3 of this chapter), the PoC incorporates the BIP-0009 mechanism in a simplified way, modifying and adapting the official Ethereum code to implement it. Particularly, it takes advantage of a block field not used by the *clique* consensus mechanism in Ethereum.

5. Organisational measures for governance management

The procedures described in this PoC have to be articulated as management processes that implement the objectives defined in the governance framework. To this end, the following would have to be done:

[79] https://geth.ethereum.org/

a) A definition of roles involved in the processes of managing the execution of the right to erasure.
b) A definition of the entities in the Blockchain infrastructure that execute each of the roles.
c) Documentation of both the policies for the execution of the erasure procedures and the actual execution of the erasure procedures.

There are many solutions for the concrete implementation of management measures. One solution could be to delegate to one entity all the roles in the process of executing the right to erasure. At the other extreme, all those roles could be executed in a distributed manner.

A mixed option has been followed in this PoC, as the main objective has been to demonstrate technical feasibility. The solution adopted is described below, as well as other possible options.

a) Defining roles in the processes of managing the enforcement of the right to erasure

Each of the following procedures shall be carried out by one or more entities that have been assigned the obligation to carry them out:

1. Documentation of policies for the execution of erasure procedures. Assigned entities should develop and maintain policies describing how erasure requests will be handled, defining the deadlines and obligations of each party involved in the process.
2. Collection of requests for the exercise of rights. A secure and accessible channel must be ensured for users to send their requests to the entity/entities in charge of collecting requests.
3. Assessment of requests to ensure their validity, authenticity and compliance. This assessment would include measures to prevent fraud or malicious use, or to determine whether or not to attend the request, among other aspects.
4. Decision to start the process of implementing the execution of exercise of rights (which could take into account other purposes than GDPR, e.g. changes to correct errors). Before initiating the erasure process, an assessment must be made to consider the effects of the erasure, whether for legal obligations or for the protection of third party rights.
5. Execution of the affected records detection procedure. All records containing the personal data that are the subject of the request, including *Smart Contracts* and *receipts/logs*, must be identified.
6. Execution of the procedure for the generation of the new software version. The new version must incorporate all the records to be modified, as well as the mechanism of agreement between the nodes, inspired by the BIP-0009 in this PoC. A record and version control of the changes made must be kept.
7. Execution of the software distribution procedure. The distribution of the update has to be coordinated in a controlled and complete manner, so that it can be ensured that the new version that a node is going to use is the correct one.
8. Validation of the new software version. A node could run the affected records detection procedure and check the changes in the databases.
9. Execution of the new software version. Nodes shall stop their operation to restart with the new version.
10. Execution of a purge process of databases or temporary storage, and any other storage containing the data to be deleted, e.g., node *keystore*, etc. It is essential that no residual or *backup* copies of the data remain with any participant in the network.

11. Maintenance of records of decision-making in relation to the exercise of right to erasure.
12. Monitoring of the management of the execution of the right to erasure and information to the data subject. Finally, organisations must implement a system of continuous monitoring to verify that the right to erasure is properly executed, including audits and review of procedures and policies.

b) Definition of the entities of the Blockchain infrastructure that execute each of the roles

In the case of this PoC, we have considered the existence of an entity[80] that assumes all the above management roles, except those numbered 8, 9 and 10, which are carried out by the participants/nodes of the infrastructure. This entity does not necessarily have to be an entity hierarchically superior to the participants/nodes, but only one that has been assigned these management roles.

On the other hand, each node would be responsible for executing the software validation procedures, managing its temporary copies and any other storage that might be affected, and executing the new software version.

This approach could easily occur in private and/or permissioned networks where such entities already exist. In another type of Blockchain infrastructure, it could be the case that different entities take on different roles. For example, nodes that are data controllers could have the obligation to collect requests for the exercise of rights without delegating it to a single node or a central entity. Alternatively, one (single) entity could manage the rights requests and another, separate entity could manage the monitoring of the procedure[81].

However, in the case of permissionless public networks, other approaches are possible. For example, distributed decision-making mechanisms could be implemented. In this context, governance is distributed among the network participants (or a selected subset of them), who through mechanisms such as consensus or voting decide the rules and operations on the Blockchain. In the case of Bitcoin or Ethereum, it is a small group that makes the decisions. In Ethereum, it was a decision by the developers to split an update called *Pectra* into two phases[82].

c) Documentation of the policies for the execution of erasure procedures, as well as the actual execution of the erasure procedures.

In the case of this Proof of Concept, a given entity is assigned the obligation to document management policies and procedures. The current chapter could be part of such documentation.

It would also be part of the documentation to establish the obligation for all nodes to execute the procedures for the detection of records affected by the right to erasure, the generation of a new software version, its distribution and execution.

In addition, there would be an obligation for all parties involved in the infrastructure to:

- Establish communication channels to inform all participants about updates. These channels should be accessible to nodes, operators and other key actors, ensuring timely notification of any updates. They should be auditable and verifiable channels.

[80] In the PoC, one and the same person takes on these roles.

[81] It may also be the case that a controller node contracts nodes for the operation on the Blockchain, by means of a controller-processor party relationship, which will have to comply with Art. 28 of the GDPR.

[82] https://www.coindesk.com/tech/2024/09/19/ethereum-developers-confirm-plan-to-split-pectra-upgrade-in-two/

- Implement an update query procedure for nodes to verify and apply updates. This ensures that updates are legitimate and prevents malicious or incorrect updates from being implemented.
- Establish a monitoring procedure to ensure that erasure is carried out correctly throughout the network and in accordance with the governance policies.
- Establish a dispute resolution procedure. It should address conflicts between participants, such as disagreements, enforcement issues or disputes over contractual obligations between controllers and processors.

Policies should include setting out the responsibilities of the actors, both in their actions as controllers and specific clauses in controller-processor contracts to ensure effective compliance with Article 28(3)(g) of the GDPR.

V. PROOF OF CONCEPT EXECUTION AND RESULTS

For the execution of the PoC use cases, several user accounts and two validator nodes have been created, two *Smart Contracts* have been deployed (simple ones, creating respective tokens and different functions), transactions have been carried out that invoke them, in particular, the creation of tokens, the purchase of tokens and the transfer of tokens. In addition, several transactions have been carried out to transfer the Blockchain's own cryptocurrency, *Ether* (which has no economic value in this private Blockchain), between users.

One of the accounts, belonging to a user, will be the object of the right to erasure request in accordance with the GDPR, and therefore any trace of the user's interactions in the Blockchain will have to be deleted. In the implementation of the PoC, an attempt has been made to cover a wide range of transactions of this account: as both the sender and recipient of *Ether* cryptocurrency transfer transactions, recipient of a token transfer in one *Smart Contract*, creator of the other *Smart Contract* and as the issuer of a token creation transaction in the latter.

A. TECHNICAL STRATEGY FOR POC IMPLEMENTATION

The governance policies clearly define the obligations and procedures for data processing, which are embodied in a technical strategy for responding to requests for rights, the steps of which are as follows:

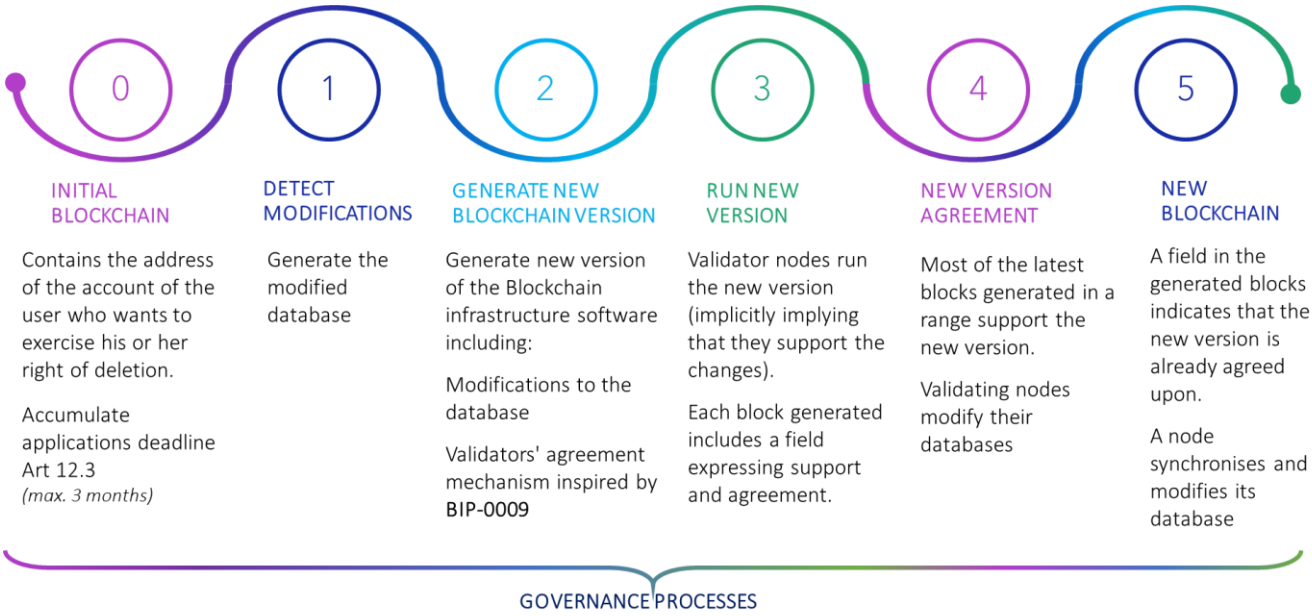


Figure 10. Phases of the PoC, each phase lists the procedures to be carried out.

0. Accumulate exercise of rights requests on a regular basis (every month or every three months if duly justified) in order to execute the requests in that period on the infrastructure.

   As indicated above, the PoC provides the right to erasure request for one account for simplicity, providing for more accounts would simply require repeating the next step 1 for each of them.

1. Detect the necessary modifications to be made in the databases of the nodes where the table of records and the blockchain materialise and generate the modified database (up to the last necessary block).

- Relevant data that are part of the blocks, *Smart Contracts* storage, transaction receipts, balance and states, etc. are overwritten. For each block, the different transactions that are recorded in it are observed, and for each transaction it is checked whether the account in question is the sender, the recipient of the transaction, or whether it is included in additional transaction data. When one of these transactions is sent to a *Smart Contract*, its storage and any transaction logs or receipts that may have been generated and stored are also checked.

  In the PoC, the value of the account in question (`0x17c3b445750221cfc48b1ea6a8d13b1eef1da197`) is overwritten by a constant value (`0xaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa`). In the case that this account is the sender of a transaction, what is overwritten with the constant value is one of the parameters of the transaction signature, since the data stored in the Blockchain does not explicitly include the field with the sender's account (*from*), what it includes is the signature of the transaction, which incorporates the public key, from which the address of the account can be derived. In Ethereum three fields are used for this purpose '*r*','*s*','*v*', the Proof of Concept overwrites 'r'[83] .

- To perform these operations, the PoC, for simplicity, runs an auxiliary node that contains a copy of the original database. This makes it easier to obtain blocks, transactions and their data through queries to the API exposed by the node. Additionally, by executing a program developed in NodeJS, read and write libraries are used in the node's databases (*LevelDB* database) to access and modify the rest of the necessary data. This approach allows the necessary modifications to be made without directly affecting the main network until the *Hard Fork* update is executed.

2. Generate the new version of the Blockchain software, which incorporates the key-value pairs from the modified database that will replace the corresponding original key-value pairs in the node's local database, and where the address of the account to be deleted will no longer appear. In addition, the new software version includes the validator node agreement mechanism, inspired by Bitcoin's BIP-0009, as described above.
3. Validator nodes update their software version. The blocks they generate include an indicator reflecting their agreement with the change.
4. When a majority of blocks in a predefined range include the change indicator in the block, then the change is agreed. At that moment, the validator nodes with the new software version modify their databases, by replacing the key-value pairs with the corresponding ones contained in this version. At this moment, the *Hard Fork* takes place, from that point on, the validator nodes modify the indicator in the blocks they generate, now signalling that this version is accepted and the database changes are implemented.
5. From this moment on, any node that wants to synchronise with the network will have to update the software to the new version, otherwise it will not be able to synchronise as it is a *Hard Fork*. The node, in the process of synchronising and downloading blocks, will check that the last existing block in the Blockchain/table contains the indicator that the version is accepted and will consequently update its database. In case the synchronisation occurs before the agreement has been reached, the node will modify its local database when it is reached.

[83] The values r and s are the components of the ECDSA elliptic curve that generates the public key, v is an identifier that facilitates the extraction of the public key. If the sender account is derived having overwritten r, a different account will be obtained (the one corresponding to the new values r,s,v, with very low probability that it already exists, in any case, the transaction has no effect on it.

B. PoC RESULTS

The PoC has been developed following the strategy detailed in the previous section, deploying two virtual machines on a Windows computer, each of which will act as a validator node, connected to a local network. Transactions are carried out from the Windows computer, connecting alternatively to each of the nodes, reproducing the usual and normal operation of a Blockchain infrastructure, thus simulating the activity of users in decentralised applications (*dApps*) or other applications (*Wallets*, *Exchanges*, trading platforms, etc.).

The following two figures show the execution of transactions from the Windows computer and the operation of the Blockchain infrastructure receiving and validating them:

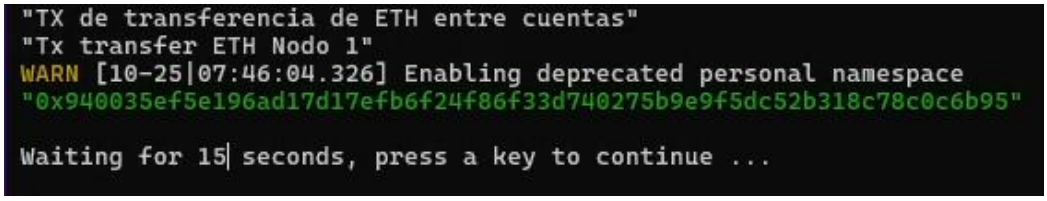


Figure 11. Execution of transactions from the Windows computer.

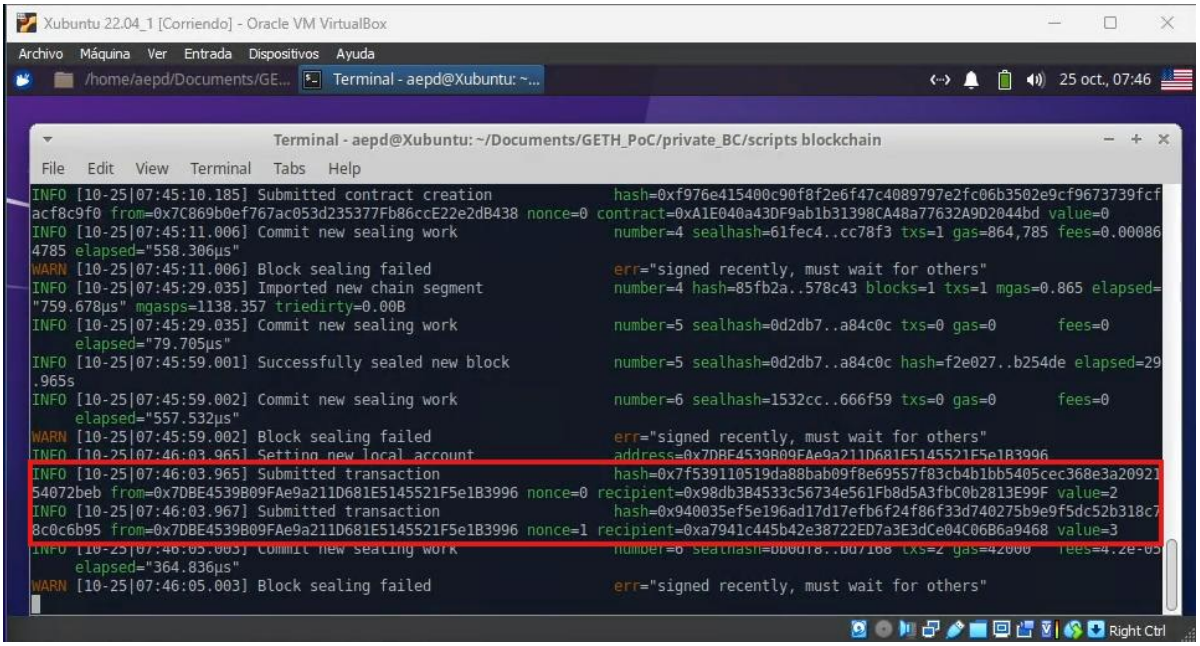


Figure 12. Operation of the Blockchain infrastructure with the 2 validator nodes and different transactions on them, marked two Ether transfers.

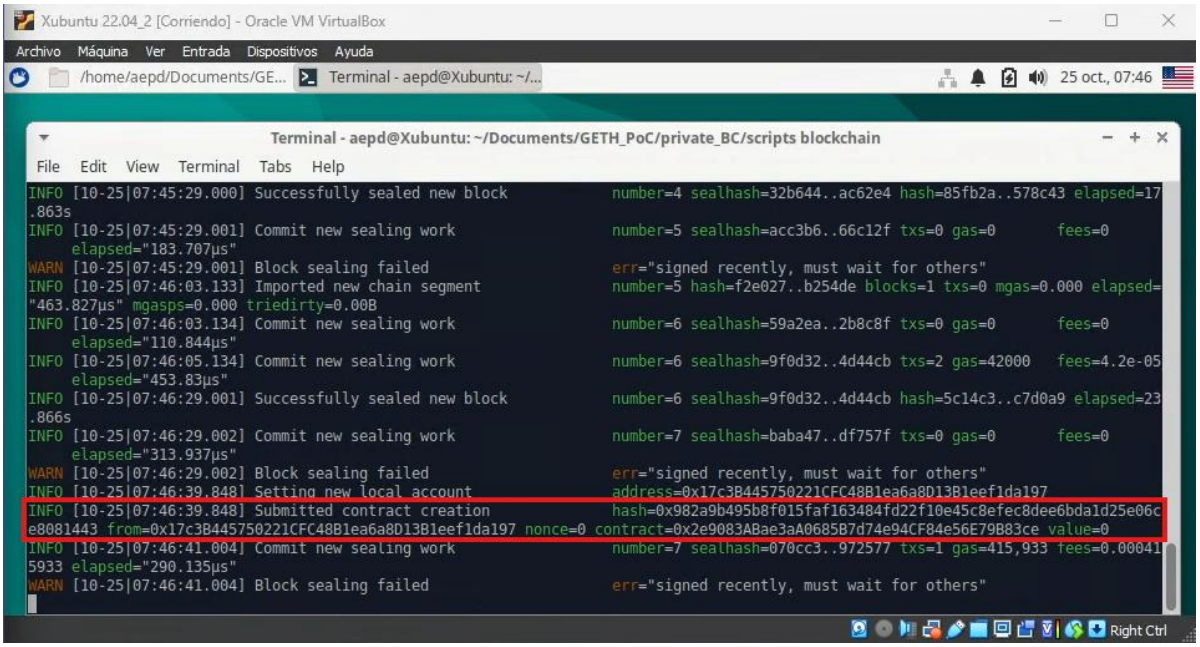


Figure 13. Operation of the Blockchain infrastructure with the 2 validator nodes. The transaction for the creation of one of the two Smart Contracts is marked.

All use cases related to the account requesting deletion have been considered. These cases are grouped into five types of transactions, which cover all possible situations in which this account appears, either as sender, receiver or argument of a call to a *Smart Contract* function.

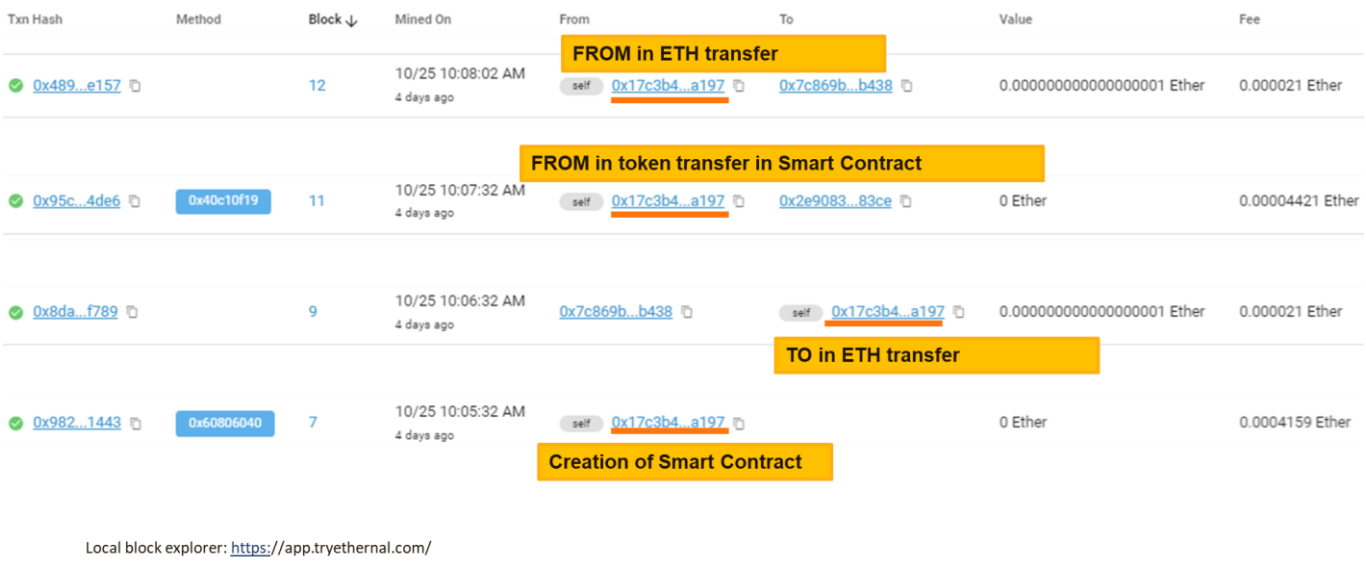


Figure 14. Transactions where the account to be deleted is listed as sender (from) or receiver (to).

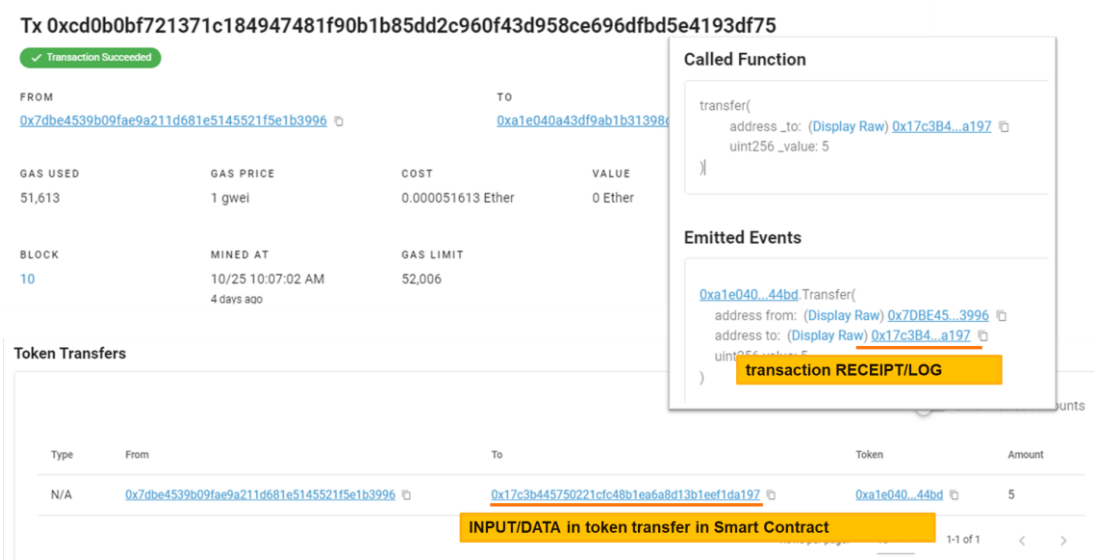


Figure 15. Transaction in which the account does not appear as sender or receiver, but in the transaction data, in this case in the argument to a call to the first Smart Contract, as the destination of a token. This transaction generates an event or log, which is stored in the transaction receipt (receipt), in the figure shown in the box overlay.

Subsequently, the necessary data to be modified is identified, and the new software version is prepared. The following figure shows how the account address has been overwritten by the value "*aaaa...a*".

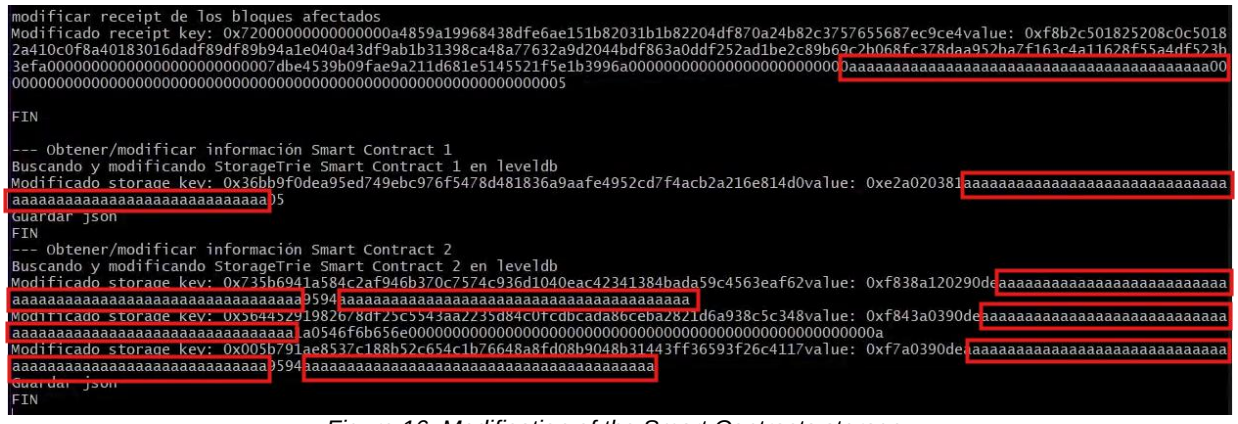


Figure 16. Modification of the Smart Contracts storage.

The following figure shows the agreement process of the two validator nodes running the updated version of the software. In each block they generate, they check the value of the indicator field in a predefined range of blocks (5 blocks in this case).

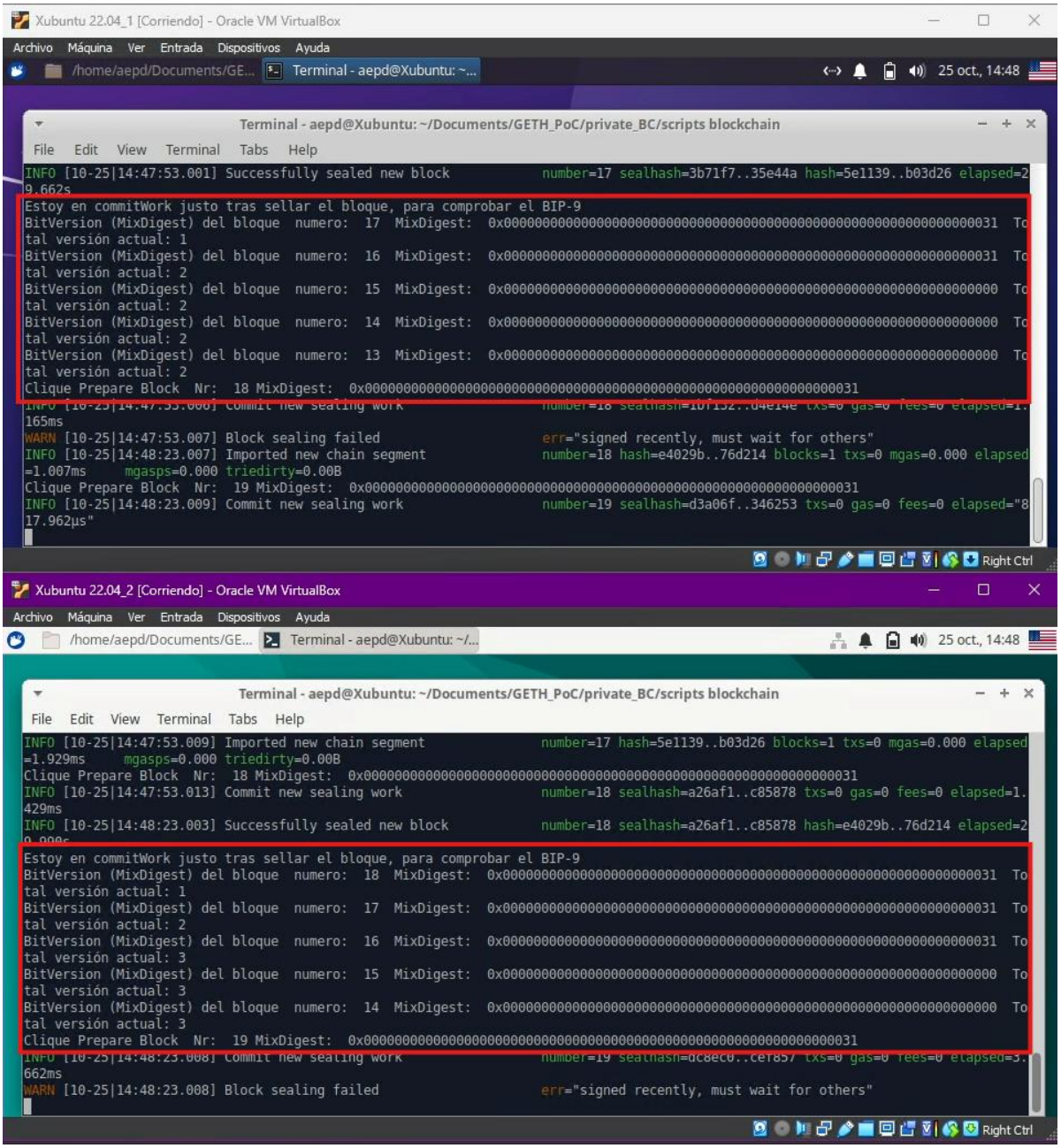


Figure 17. Agreement process between validator nodes (inspired by BIP-0009). The blocks generated with this new version show a value '0...0031' (text string '1' encoded to hexadecimal format), while the previous ones show '0...0000'.

When the majority of the blocks generated with the indicator of this new version are reached, the indicator is modified to indicate that the modification is accepted while the validator nodes simultaneously modify their databases, making the right to erasure effective.

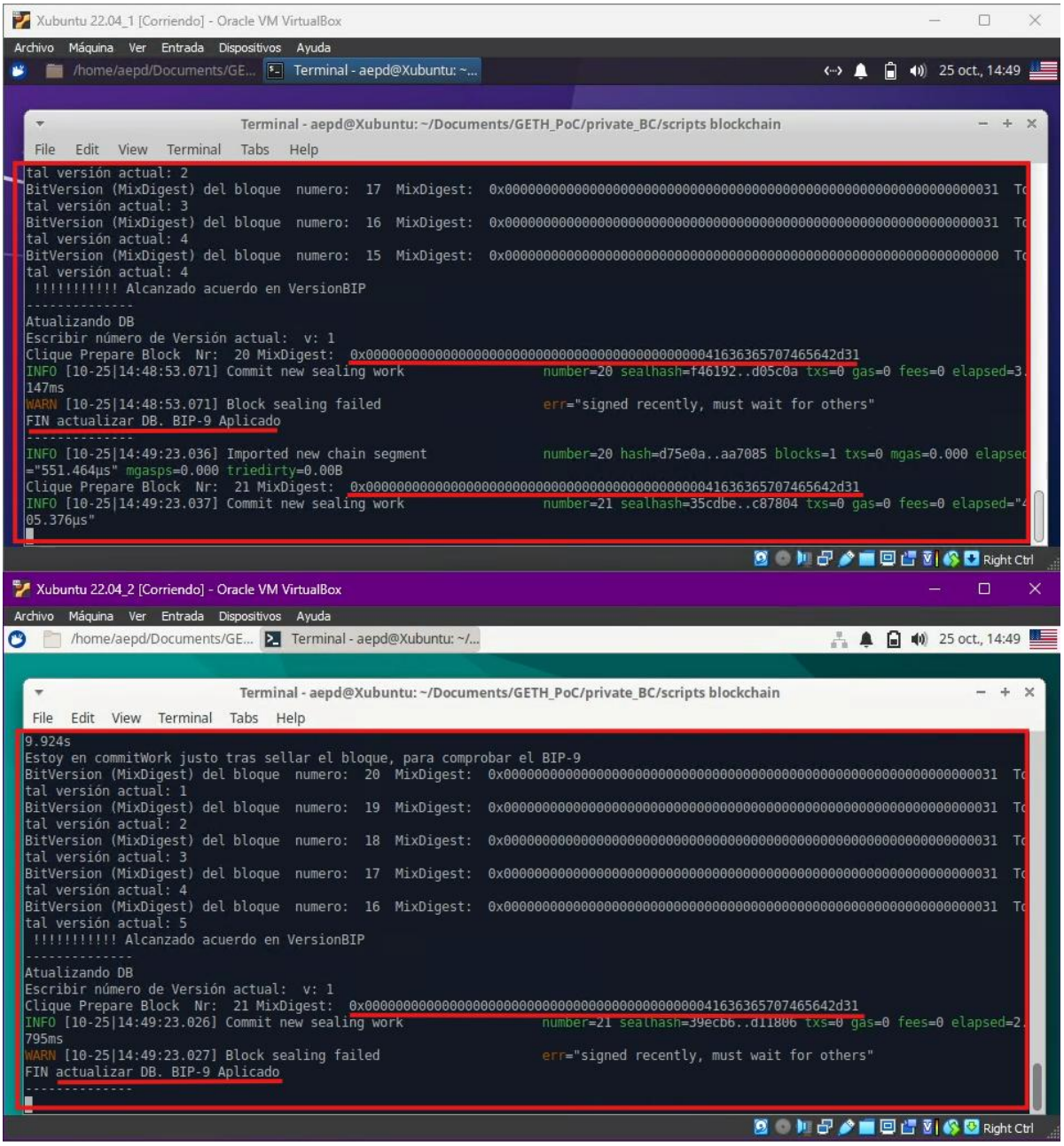


Figure 18. Agreement reached between validator nodes (inspired by BIP-00009). The nodes modify their databases, and the indicator of the new block is updated to indicate that the modification has been accepted (the blocks generated from now on show the value 'Accepted-1', encoded in hexadecimal format).

1. Transactions

Once the agreement has been reached and the *Hard Fork* has taken effect, it can be verified that the address of the account in question no longer appears in the interactions it had with the Blockchain infrastructure. To show the results in a visual and simple way, the local block explorer *Ethernal*[84] has been used. The block explorer does not show transaction data since controlled inconsistencies have been caused, and therefore cannot decode the corresponding data correctly (however, the node databases contain the raw data of the blocks and transactions with the modifications made).

[84] https://app.tryethernal.com/

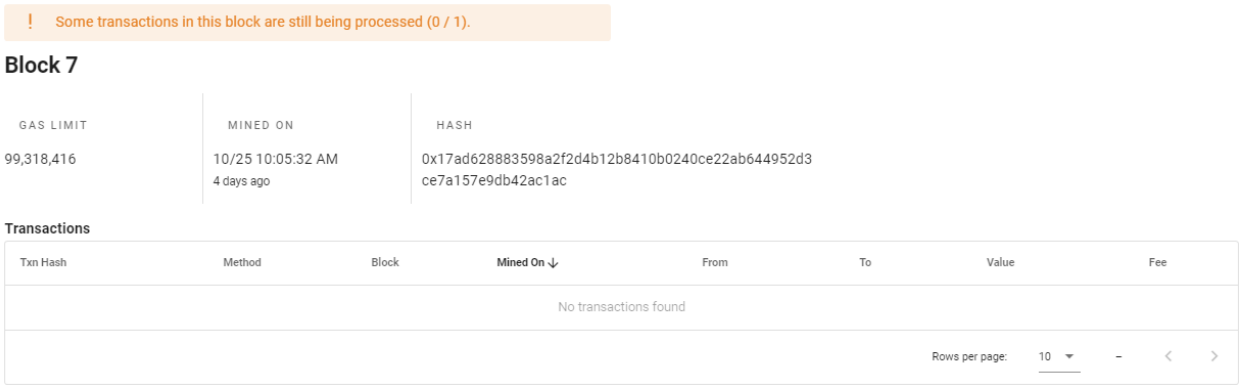


Figure 19. The block explorer cannot decode the transaction and warns with an error. The account in question does not appear and the transaction data is not displayed.

Txn Hash	Method	Block	Mined On	From	To	Value	Fee
0x492...d227		13	10/25 10:08:32 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0xa3d...ea18		13	10/25 10:08:32 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0xa2b...7a70	0xd0679d34	12	10/25 10:08:02 AM 4 days ago	0x98db3b...e99f	0x2e9083...83ce	0 Ether	0.00005239 Ether
0x489...e157		12	10/25 10:08:02 AM 4 days ago	0x17c3b4...a197	0x7c869b...b438	0.000000000000000001 Ether	0.000021 Ether
0x95c...4de6	0x40c10f19	11	10/25 10:07:32 AM 4 days ago	0x17c3b4...a197	0x2e9083...83ce	0 Ether	0.00004421 Ether
0xdb7...a5cd	transfer	11	10/25 10:07:32 AM 4 days ago	0x7dbe45...3996	0xa1e040...44bd	0 Ether	0.00005161 Ether
0x20f...2cbe		11	10/25 10:07:32 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0xf88...d2e5		11	10/25 10:07:32 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0xd69...a7a2		10	10/25 10:07:02 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0xdd9...2080		10	10/25 10:07:02 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0xcd0...df75	transfer	10	10/25 10:07:02 AM 4 days ago	0x7dbe45...3996	0xa1e040...44bd	0 Ether	0.00005161 Ether
0x8da...f789		9	10/25 10:06:32 AM 4 days ago	0x7c869b...b438	0x17c3b4...a197	0.000000000000000001 Ether	0.000021 Ether
0x509...e6a4		9	10/25 10:06:32 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0x621...12fa		9	10/25 10:06:32 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0xc47...98fe	0xf088d547	9	10/25 10:06:32 AM 4 days ago	0x7c869b...b438	0xa1e040...44bd	0.007 Ether	0.00006408 Ether
0x982...1443	0x60806040	7	10/25 10:05:32 AM 4 days ago	0x17c3b4...a197		0 Ether	0.0004159 Ether
0x940...6b95		6	10/25 10:05:02 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0x7f5...2beb		6	10/25 10:05:02 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0xf97...c9f0	0x60806040	4	10/25 10:04:02 AM 4 days ago	0x7c869b...b438		0 Ether	0.0008648 Ether

Rows per page: 100

Figure 20. The block explorer shows the 19 transactions in the chain.

Txn Hash	Method	Block	Mined On	From	To	Value	Fee
0xa3d...ea18		13	10/25 10:08:32 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0x492...d227		13	10/25 10:08:32 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0xa2b...7a70	0xd0679d34	12	10/25 10:08:02 AM 4 days ago	0x98db3b...e99f	0x2e9083...83ce	0 Ether	0.00005239 Ether
0x20f...2cbe		11	10/25 10:07:32 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0xf88...d2e5		11	10/25 10:07:32 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0xdb7...a5cd	transfer	11	10/25 10:07:32 AM 4 days ago	0x7dbe45...3996	0xa1e040...44bd	0 Ether	0.00005161 Ether
0xdd9...2080		10	10/25 10:07:02 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0xd69...a7a2		10	10/25 10:07:02 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0xc47...98fe	0xf088d547	9	10/25 10:06:32 AM 4 days ago	0x7c869b...b438	0xa1e040...44bd	0.007 Ether	0.00006408 Ether
0x621...12fa		9	10/25 10:06:32 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0x509...e6a4		9	10/25 10:06:32 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0x940...6b95		6	10/25 10:05:02 AM 4 days ago	0x7dbe45...3996	0xa7941c...9468	0.000000000000000003 Ether	0.000021 Ether
0x7f5...2beb		6	10/25 10:05:02 AM 4 days ago	0x7dbe45...3996	0x98db3b...e99f	0.000000000000000002 Ether	0.000021 Ether
0xf97...c9f0	0x60806040	4	10/25 10:04:02 AM 4 days ago	0x7c869b...b438		0 Ether	0.0008648 Ether

Rows per page: 100

Figure 21. The block explorer shows14 transactions in the modified chain (Hard Fork), the 5 where the account of the user exercising the right to erasure was revealed cannot be obtained.

2. Balance

Similarly, after the *Hard Fork,* the balance of the deleted account can no longer be displayed (it shows a value of 0, just like any other non-existent account in the Blockchain infrastructure) because the node is not found in the *State Trie*[85], the data structure that contains the balances and data of the accounts.

```
> eth.getBalance('0x17c3b445750221cfc48b1ea6a8d13b1eef1da197')
399518855000000000
>
```

Figure 22. In the initial Blockchain/table, before the deletion, a node console shows the value of the balance of the account that has been deleted in the Wei unit. In the genesis block this account was founded with 400000000000000000 Wei, which has been reduced by the amounts used in transfers and cost (gas) of the transactions made by this account.

```
> eth.getBalance('0x17c3b445750221cfc48b1ea6a8d13b1eef1da197')
0
>
```

Figure 23. After the Hard Fork, a node console displays a 0 value for the balance of the account that has been removed. NOTE: This command displays a 0 value for the balance of any account that does not exist.

[85] https://ethereum.org/en/developers/docs/data-structures-and-encoding/patricia-merkle-trie/#state-trie

3. *Smart Contracts* storage and transaction receipt *(receipts/logs)*

The *Smart Contracts* storage will also not show the information corresponding to the deleted account, as the nodes of the *Storage Trie*[86] have been overwritten.

For example, in the first *Smart Contract*, whose address is '`0xa1e040a43df9ab1b31398ca48a77632a9d2044bd`', the account being deleted got a transfer of 5 tokens. The following figure shows the storage access before and after the *Hard Fork*.

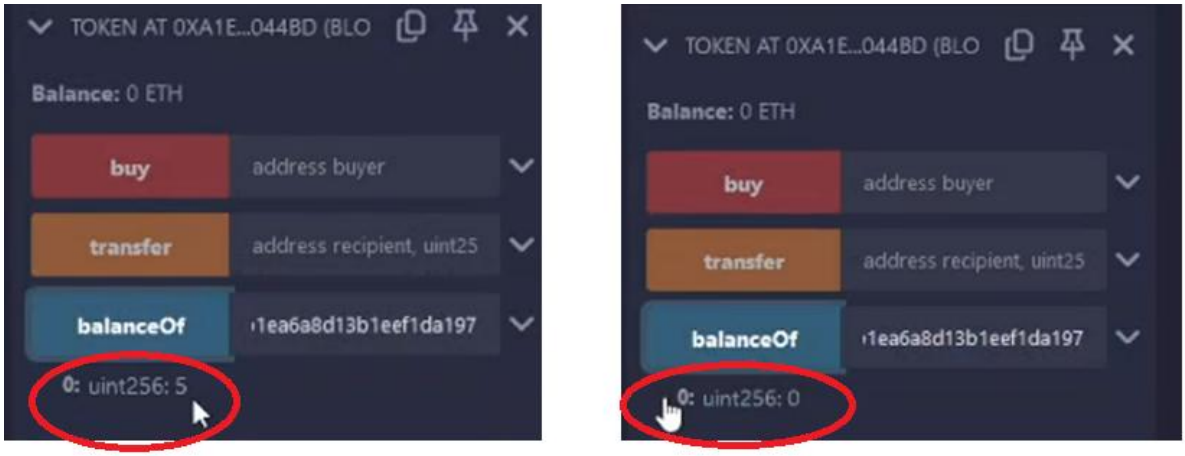


Figure 24. Balance of tokens for the account in question before (left) and after (right) the Hard Fork, as seen from an application to interact with the Smart Contract (REMIX IDE).

The storage of the token balances of the accounts in this *Smart Contract* is done through a mapping between the account addresses and their token balance and can be verified through a console connected to one of the nodes.

```
> eth.getStorageAt('0xa1e040a43df9ab1b31398ca48a77632a9d2044bd','0xc0f2fb06a666264727960e90672
27c73b3561f16ddd0c96159193bf520102c8f')
"0x0000000000000000000000000000000000000000000000000000000000000005"
>
```

Figure 25. Balance of the first Smart Contract storage for the account to be deleted. The function that obtains it has two parameters, the first the address of the Smart Contract, the second obtained with the index of the variable where it is stored in the source code of the Smart Contract, of type address mapping. It is calculated by hashing the address of the account concatenated with the index of the variable, 7 in this case. The value of the second parameter is Keccack256(0000000000000000000000000000000000000017c3b445750221cfc48b1ea6a8d13b1eef1da1970000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000000007).

```
> eth.getStorageAt('0xa1e040a43df9ab1b31398ca48a77632a9d2044bd','0xc0f2fb06a666264727960e90672
27c73b3561f16ddd0c96159193bf520102c8f')
"0x0000000000000000000000000000000000000000000000000000000000000000"
>
```

Figure 26. After the Hard Fork the token balance of the deleted account in the first Smart Contract shows a null value. NOTE: Any account that does not exist in the Smart Contract will show a null token balance.

[86] Structure where the data of a *Smart Contract* resides. https://ethereum.org/en/developers/docs/data-structures-and-encoding/patricia-merkle-trie/#storage-trie

The second *Smart Contract*, whose address is '0x2e9083abae3aa0685b7d74e94cf84e56e79b83ce', stores in a variable the account that created the contract, which in the PoC was the one to be deleted.

```
> eth.getStorageAt('0x2e9083abae3aa0685b7d74e94cf84e56e79b83ce',0)

"0x0000000000000000000000017c3b445750221cfc48b1ea6a8d13b1eef1da197"
>
```

Figure 27. Storage of the address of the account that created the second Smart Contract, showing the account to be deleted. The function that obtains it has two parameters, the first one the address of the Smart Contract, the second one obtained with the index of the variable where it is stored in the source code of the Smart Contract, of type address (account address), 0 in this case.

```
> eth.getStorageAt('0x2e9083abae3aa0685b7d74e94cf84e56e79b83ce',0)

"0x0000000000000000000000000000000000000000000000000000000000000000"
>
```

Figure 28. After the Hard Fork, the result of the account creating the second Smart Contract shows a null value.

Regarding the transaction receipt, the second transaction of block 10, which corresponds to the transfer of tokens from the first *Smart Contract* to the account being deleted, generates a receipt (*receipt/log*) that is stored on the node in another data structure than the Blockchain/table. The transaction receipt contains the address of the account and therefore also needs to be taken into account when exercising the right to erasure.

```
> eth.getBlockReceipts(eth.getBlockByNumber(10).hash)[2]
{
  blockHash: "0x1a1d4fd85c1c9af21253eb9daf3cbc8ad740d527eacd04bf725c326e1b311f85",
  blockNumber: "0xa",
  contractAddress: null,
  cumulativeGasUsed: "0x16dad",
  effectiveGasPrice: "0x3b9aca00",
  from: "0x7dbe4539b09fae9a211d681e5145521f5e1b3996",
  gasUsed: "0xc99d",
  logs: [{
      address: "0xa1e040a43df9ab1b31398ca48a77632a9d2044bd",
      blockHash: "0x1a1d4fd85c1c9af21253eb9daf3cbc8ad740d527eacd04bf725c326e1b311f85",
      blockNumber: "0xa",
      data: "0x0000000000000000000000000000000000000000000000000000000000000005",
      logIndex: "0x0",
      removed: false,
      topics: ["0xddf252ad1be2c89b69c2b068fc378daa952ba7f163c4a11628f55a4df523b3ef", "0x0000000000000000000000007
dbe4539b09fae9a211d681e5145521f5e1b3996", "0x0000000000000000000000017c3b445750221cfc48b1ea6a8d13b1eef1da197"],
      transactionHash: "0xcd0b0bf721371c184947481f90b1b85dd2c960f43d958ce696dfbd5e4193df75",
      transactionIndex: "0x2"
  }],
  logsBloom: "0x00000000000080000000000000000000000000000000000000000000000000000000000000000000800000000000000000000000000
00000200000000000000000000000000000000000000000000000000000000000000000800000000002000000000000000000000000000000000000000000000
00000000000000000000000000000000010000000000000000000000000000000000000000010000000000000000000000000000000000000000000000000
00000000020000020000000000000000000000000000000010000000000000002000000000000000000000000000000000000000000000000000000000000
00000000000000000000000000000000000000000000000000000000000000000400000000000",
  status: "0x1",
  to: "0xa1e040a43df9ab1b31398ca48a77632a9d2044bd",
  transactionHash: "0xcd0b0bf721371c184947481f90b1b85dd2c960f43d958ce696dfbd5e4193df75",
  transactionIndex: "0x2",
  type: "0x0"
}
>
```

Figure 29. The receipt of the token transfer transaction of the first Smart Contract shows the value of the address of the account to be deleted (red box in the figure).

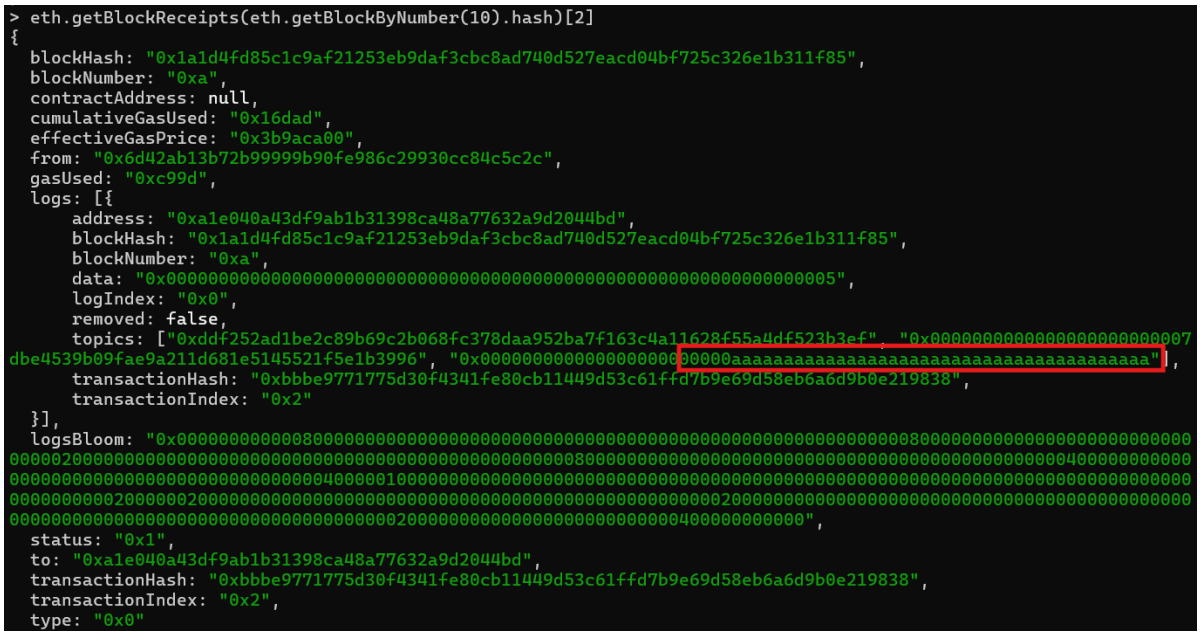


Figure 30. After the Hard Fork, the address of the deleted account is no longer in the receipt information of the token transfer transaction of the first Smart Contract, it has been properly modified.

VI. FUTURE DEVELOPMENTS

The PoC presented in this document is not intended to be a commercial solution for direct market application, nor a development ready for production deployment. Nor does it intend to validate that achieving GDPR compliance in a data processing activity can be done by making modifications after the processing is already designed or even in operation. On the contrary, the failure to consider data protection by design is, in itself, a regulatory non-compliance.

At the present time, the Proof of Concept essentially contemplates the application of the right to erasure (deletion of an account and all its interactions occurring on the Blockchain infrastructure). As mentioned above, it is based on the use of the Hard Fork mechanism to implement inconsistency management and allows the system to continue operating with guarantees. This solution does not recompose the integrity of blocks with deleted data and takes advantage of the fact that a node does not constantly validate the past but uses its local database for validation. However, it is possible to use other approaches, such as techniques based on restoring the mathematical integrity of the Blockchain blocks through, for example, Chameleon hashing, mutable transactions (multiple versions of a transaction), new block structures, off-chain strategies, ZKP zero-knowledge proofs, local erasure, *pruning*, as used in various research and projects that have explored strategies to develop "*redactable*" Blockchain.

Finally, the PoC has not taken into account the effects of the right to erasure on the user exercising the right to erasure, i.e. the recovery of his funds or balances in *Smart Contracts*, it is assumed that the user would have to manage these effects before making his right to erasure request, e.g. by transferring the funds and balances to other accounts. Future developments could provide for this.

VII. CONCLUSIONS

Using Blockchain technologies, infrastructures that enable the storage and exchange of information in a distributed and decentralised manner can be created. Such Blockchain infrastructures adapt the Blockchain technology to their specific objectives and are complemented by additional systems and applications. A data processing might decide to implement some of its operations, e.g. storage, using such an infrastructure. In turn, the Blockchain infrastructure could support the operations of different processing. Finally, the management of the infrastructure itself could involve specific processing of personal data.

Any project that involves developing new personal data processing using this technology must ensure that compliance with the principles, rights and obligations set out in the GDPR is not limited by the choice of a particular technology option. If the original design prevents compliance, that design should be modified to ensure and demonstrate compliance with the GDPR. If, after assessing its technical and regulatory feasibility, it is decided to implement any processing operation on a Blockchain-based solution, the processing, its context, scope and implications should be carefully analysed to ensure compliance. Beyond minimum compliance, if a technological option involves a high risk to the rights and freedoms of data subjects, this risk must be managed and subjected to an analysis of suitability, proportionality and necessity in the framework of passing a data protection impact assessment.

In many of today's Blockchain infrastructures, data protection principles, in particular accountability, are not applied by design. The fact that the design of a Blockchain infrastructure has not considered compliance management as an objective is not a problem with the technology itself. The problem is that the designers have not considered these objectives in the governance framework of the infrastructure and therefore the appropriate management mechanisms have not been included in the design.

In many cases, Blockchain infrastructures are being implemented, or processing on these infrastructures, with a great lack of knowledge of how the components, codes or applications used in these infrastructures work. In many cases, there is a lack of minimum documentation describing them, the necessary analyses have not been carried out to demonstrate that they work with the necessary quality, and there is insufficient information to adapt them to regulatory requirements. No processing should be built with systems, products, services or components if controllers and processors are not in a position to fulfil their obligations to ensure and demonstrate data protection compliance (Recital 78 and Articles 24(1), 25 and 32(1)(b)).

However, compliance with the GDPR is possible, and must be achieved, regardless of the technologies used for the implementation of processing operations, by implementing appropriate legal, organisational and technical measures (Recital 15 of the GDPR).

The European Data Protection Board has stated that technical impossibility cannot be invoked to justify non-compliance with the requirements of the GDPR[87]. Especially in view of the fact that Article 25(1) of the GDPR provides that data protection by design implies taking it into account at the time of the determination of the means of processing and at the time of the processing itself. In the case in question, as in others dealt with by the AEPD[88], the question arises as to when we are faced with a technical impossibility of protecting the rights and freedoms of individuals, and when we are faced with a resistance to applying solutions for the protection of such rights.

[87] EDPB, Report of the work undertaken by the ChatGPT Taskforce of May 2024, paragraph 7: "*In particular, technical impossibility cannot be invoked to justify non-compliance with these requirements, especially considering that the principle of data protection by design set out in Article 25(1) GDPR shall be taken into account at the time of the determination of the means for processing and at the time of the processing itself*".

[88] https://www.aepd.es/en/areas/innovation-and-technology#Minors

It is therefore imperative for controllers, processors and developers to know the real implications, possibilities and limitations of the components they select to build products, services and data processing. This knowledge must be based on objective evidence obtained through documentation, certification and/or auditing. Controllers, processors and developers, including control authorities, must be careful about the terminology used. The current terminology generates confusion in concepts and their implications. It is necessary to avoid drawing conclusions on imprecise terms, in order to avoid misunderstandings about the technological reality.

The Proof of Concept, presented here, develops an approach to the implementation of technical measures and a data protection policy in the governance of the Blockchain infrastructure, which demonstrates that compliance with the GDPR is possible. More specifically, it enables the exercise of the right to erasure through inconsistency management that allows the infrastructure to continue operating with the same guarantees.

To this end, procedures have been designed for the execution of the right to erasure. In turn, the code and applications have been adapted from a real, globally used infrastructure, which is the official Ethereum infrastructure, configured with the *clique* proof of authority consensus mechanism. In addition, all the personal data that can be processed in the infrastructure, such as the sender and recipient of a transaction, the additional data contained in a transaction, the storage of *Smart Contracts* and the receipts of transactions involving *Smart Contracts* (*receipts/logs*), have been taken into account. This has required intensive research work on the infrastructure code and documentation and the deployment of a use case. Finally, the necessary governance measures for its implementation have been defined.

In this specific case, the technical strategy employed on a Blockchain infrastructure, which originally did not take data protection into account by design, requires direct modification of the nodes' databases. Such measures, in terms of their impact on the management of the infrastructure, are periodically taken when it has been necessary to deal with critical incidents in its operation. Blockchain's theoretically rigid principles are adapted to specific infrastructures, as well as to the incidents that occur in them, such as those mentioned in Chapter III, Ethereum's *DAO Fork* and *Bitcoin's Cash Fork*, among many other cases.

Although it has been developed on a widespread Blockchain infrastructure, it is a laboratory test, as the development of commercial solutions is outside the competence of a supervisory authority. In no case is the AEPD imposing the use of this example, nor does it imply a commitment in relation to the exercise of the powers granted to it by the GDPR as a supervisory authority.

To conclude, this Proof of Concept does not validate that GDPR compliance can be achieved by modifying blockchain infrastructures that have failed to comply with data protection by design or the accountability principle. In no kind of processing is this the way to go. On the contrary, it is intended to promote among designers, authorities and organisations that have a role in the design and development of these technologies the adoption of data protection by design and by default strategies that comply with the requirements of the GDPR.