ANNEX 3.2

REFINING RESEARCH CHALLENGES AND DIRECTION

SECOND OPEN CALL FOR PROPOSALS

RESEARCH CHALLENGES

Blockchain technology and its products are in constant development and fruition, thus, new characteristics and attributes emerge along with new issues and challenges. The monitoring of this technological domain is very important for the accurate and efficient management of the technology by the stakeholders. The purpose of this action is to monitor such scientific and technological evolution and support the relevant research, development and financial efforts that are in progress. Through experienced methodologies and by approaching the state of the art solutions and related work, it is easier to manage and monitor such dynamic technologies, while the approach will work as a feedback loop for the next iterations of these efforts.

1 TRUST AND REPUTATION SYSTEMS IN BLOCKCHAIN TECHNOLOGY

Several distributed trust and reputation systems exist without blockchain as the underlying technology, though they are susceptible to attacks and malicious control[1]. PTM[2], a pervasive trust management model that is based on trust relations between entities, suggests that trust can be gained on both direct and indirect basis, and the overall trust score is calculated on a reference model as the average of all recommendations, weighted by the trust degree of the recommender. Another system called FIRE[3] proposes an integrated trust and reputation model for open multi-agent systems to compute participant reputation. Based on different sources of trust information, such as direct experience, witness information, role-based rules, and third-party references

[1] https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8970496

[2] http://www.it.uc3m.es/~florina/ptm/

[3] https://eprints.soton.ac.uk/262593/1/jaamas-dong.pdf

reputation is computed. Furthermore, there are trust and reputation systems where every peer holds some part of the information, thus, the reputation of a service provider can be computed[4] [5]. These systems also use witnesses for each transaction, which guarantees that the reputation submission will be correctly performed even if one of the two parties was to abort the protocol. Additionally, there exist other protocols that try to achieve ideal decentralization and the reputation feedback is retrieved from the participants each time a party wishes to know the reputation of another party[6]. In those cases, all network nodes should stay online to contribute to the reputation calculation. Such approaches have been used in P2P applications but seem to not be suitable for many other domains, such as in e-commerce or IoT[7]. Furthermore, these protocols are rather confidentiality-preserving than privacy-preserving, in that they do not hide the list of users who participated in the rating[8]. This way of partially hiding information leads to multiple issues linked to the mutability of the set of participating peers[9]. For instance, the contribution of a user to the aggregated reputation might be revealed if the user goes offline between two reputation-queries[10]. These distributed approaches, however, are not immune to attacks and potential manipulation.

Regarding blockchain-based trust and reputation systems, there exist numerous paradigms in diverse domains. For instance, a blockchain-based trust and reputation system is used to implement security in the DNS[11], overcoming the highly centralized field limits of the Domain Name System Security Extensions (DNSSEC) protocol. The solution is based on Namecoin[12] (namely Flatcoin) using the popular consensus mechanism Proof-of-Work (PoW)[13]. The solution aims to reduce the number of built-in CA certificates to the minimum required level defined by the end-user, with the ultimate goal each user to become a Trusted Third Party (TTP), and the customers to purchase the certificate from the TTP with the highest reputation score (Figure 1). The gross amount of transaction fees paid serves as a reputation score metric.

[4] https://ieeexplore.ieee.org/document/6654809

[5] https://hal.inria.fr/hal-01104837v1

[6] http://leibniz.cs.huji.ac.il/tr/693.pdf

[7] https://link.springer.com/article/10.1186/2196-064X-1-8

[8] https://www.sciencedirect.com/science/article/pii/S0167404811001465

[9] https://www.sciencedirect.com/science/article/abs/pii/S157087051300098X

[10] https://dl.acm.org/doi/abs/10.5555/1139711.1648680

[11] https://ieeexplore.ieee.org/document/6550483

[12] https://namecoin.org

[13] https://bitcoin.org/bitcoin.pdf

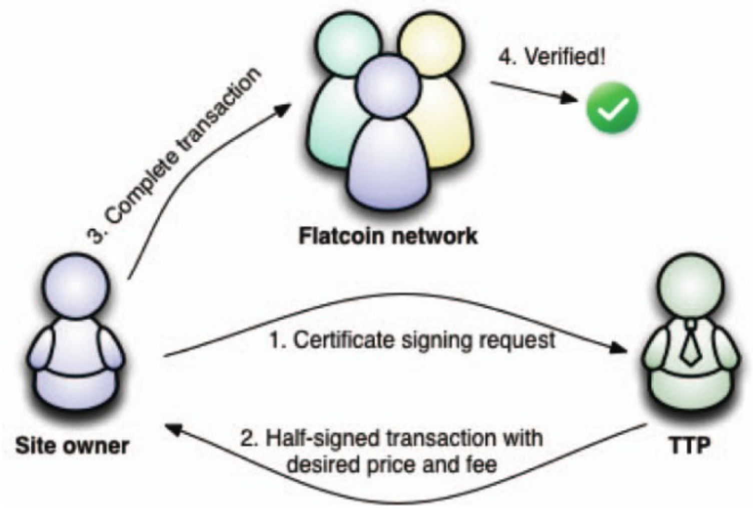

3. Complete transaction
Flatcoin network
4. Verified!
1. Certificate signing request
Site owner
2. Half-signed transaction with desired price and fee
TTP

FIGURE 1: FLATCOIN VALIDATION PROCESS.[14]

Furthermore, another solution of such a blockchain-based system is aiming to manage quantifying reputation by subtracting the personal opinion from the transaction in a P2P network. The solution suggests a Proof-of-Stake (PoS)[15] approach in order to reduce malicious transactions on the network. Reputation is saved on the blockchain, and the client calculates the reputation score based on its parameters and only over a short period (Error: Reference source not found). Another solution tries to improve resource sharing in P2P networks by suggesting a custom multi-level reputation scoring system based on rewards that aim in regulating a fair usage of resources among all nodes of an inter-organization cluster. Blockchain is used for logging node activities, allowing any single node to calculate the reputation of a given node, to identify and eliminate nodes that tend to overuse resources of the whole cluster and do not contribute by their computation resources or contribute by false results[16].

[14] https://ieeexplore.ieee.org/document/6550483
[15] https://eth.wiki/en/concepts/proof-of-stake-faqs
[16] https://ieeexplore.ieee.org/document/8026098

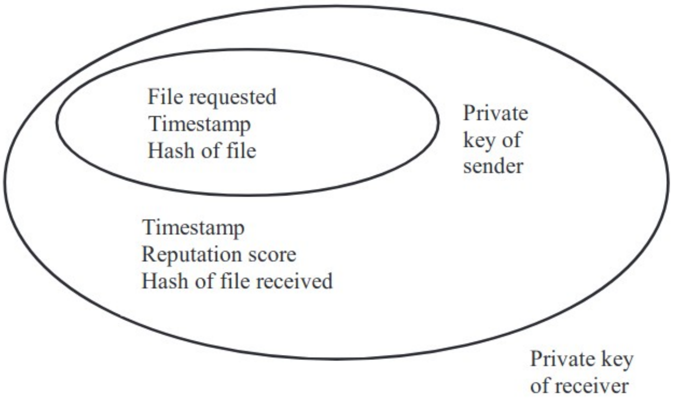


FIGURE 2: RECEIPT OF TRANSACTION SENT TO THE MINERS.[17]

In e-commerce there exist various solutions that utilize blockchain-based trust and reputation systems as well. For instance, in an attempt to lower the overhead for the processing of transactions, the following architecture is proposed[18]. The customer retrieves the Service Provider (SP) reputation to decide whether to engage in a transaction and once the transaction is completed, the customer receives a token from the SP based on the amount available on its account. Afterwards, the customer broadcasts a message containing the address of the SP, the token, the rating of the transaction as well as a pointer to the last review concerning the same SP. This pointer enables any participant to compute the reputation much faster since it does not need to retrieve the entire reputation history. In another e-commerce solution, the main goal is to eliminate the need for third parties by analyzing the underlying transaction network structure and building a history of transaction outcomes[19]. This reputation system is composed of a series of dedicated smart contracts executed by developers who are running Bitcoin nodes with pre-installed Counterparty or Ethereum (Figure 3). Thus, each user calculates the subjective score and decides whether to engage in the transaction. If the transaction occurs, the user runs the proprietary contract to store the outcome in the Ethereum or Counterparty blockchains.

[17] https://ieeexplore.ieee.org/document/7412073

[18] https://link.springer.com/chapter/10.1007/978-3-319-33630-5_27

[19] https://www.seas.upenn.edu/~cse400/CSE400_2014_2015/reports/07_report.pdf

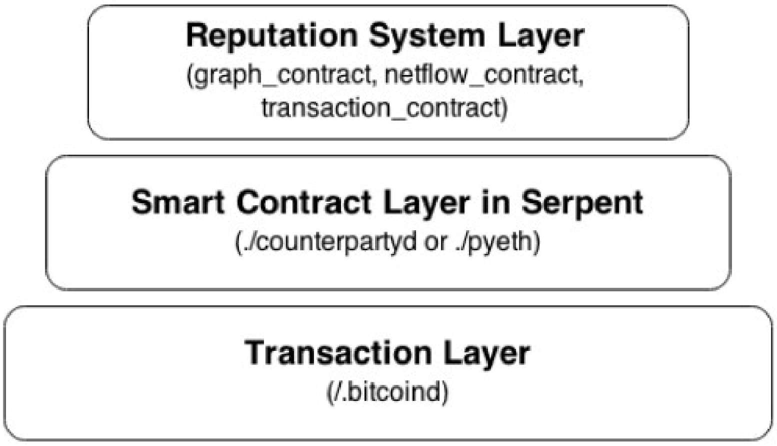

Reputation System Layer
(graph_contract, netflow_contract, transaction_contract)
Smart Contract Layer in Serpent
(./counterpartyd or ./pyeth)
Transaction Layer
(/.bitcoind)

FIGURE 3: TECHNOLOGIES USED IN THE ARCHITECTURE.[20]

Additionally, in the crowd-sourcing domain, there are important works as well. In particular, PaySense[21] constitutes a general framework that promotes user participation and provides a mechanism to validate the quality of collected data based on the reputation of the users. This approach adopts bitcoins as a reputation annotation system, and as a result, unifies the concepts of reward and reputation in a unique value. In this context, the balance in a specific Bitcoin address represents both the total awarded bitcoins for the sensing tasks reported with such Bitcoin address and the reputation obtained for the tasks. In another crowd-sourcing solution, named CrowdBC[22], a new consensus protocol called Proof-of-Trust (PoT) is suggested. The authors present a novel approach that separates the transaction validation and block recording in two different groups. The goal is to achieve a better trade-off between centralization/decentralization and security/fairness. The application proposes a hybrid blockchain solution that utilizes a permissioned blockchain as the underlying deployment architecture, while the transaction validation of the consensus protocol is performed through an open, public network environment, which exhibits the fairness and impartiality properties of a public blockchain (Figure 4).

[20] https://www.seas.upenn.edu/~cse400/CSE400_2014_2015/reports/07_report.pdf

[21] https://pubmed.ncbi.nlm.nih.gov/27240373/

[22] https://ieeexplore.ieee.org/document/8540048

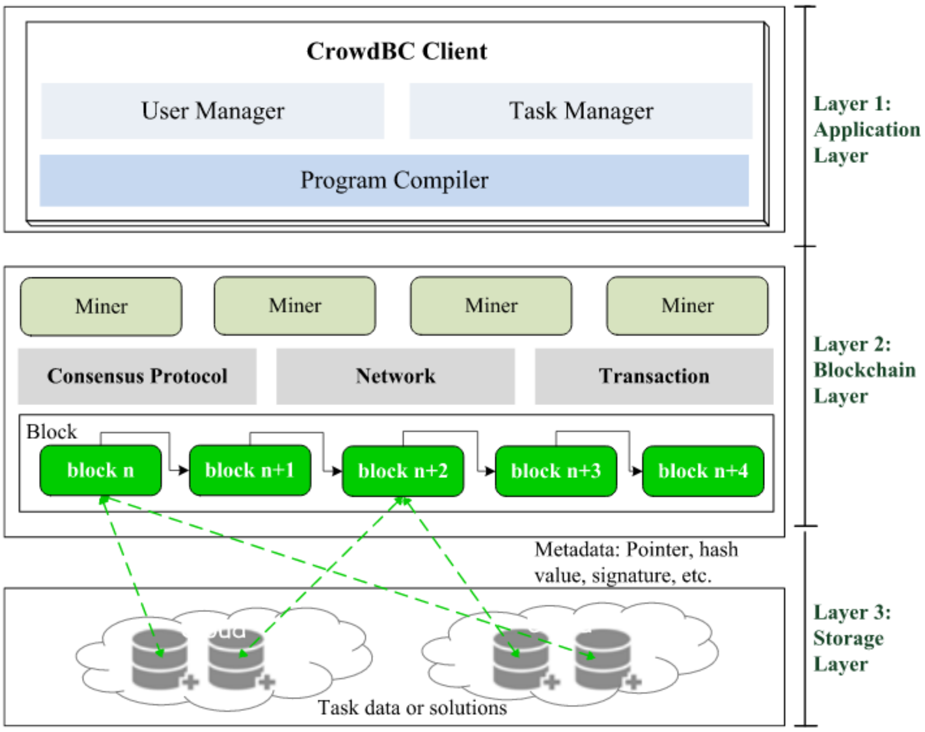


FIGURE 4: CROWDBC ARCHITECTURE.[23]

Blockchain-based trust and reputation systems have been used in IoT and sensors networks as well, to guarantee security and consistency. For instance, BATM[24] is an approach where blockchain is used as generic storage to manage trust and authentication for decentralized sensor networks. In this approach, a properly sized payload for storing essential security and trust information in a Bitcoin-based blockchain is implemented. The payload contained in the blockchain is used as an indication of a node's behaviour over time, while reputation and trust are derived by the event analysis saved in the blockchain. Another work focuses more on the substantial lack of trust between devices in IoT and creates a custom blockchain, named "obligation chain", where anyone is allowed to consume services by providing a public obligation for fulfilling the terms of use as specified by the SP (Figure 5) [25]. The trust that users already have with their mobile operators is leveraged to provide a complete path of trust between any customer and the SP. In this context, blockchain is used to create an obligation chain, which is a new platform for a distributed credit-like system, which has a built-in reputation mechanism allowing peers to decide whether or not to accept obligations based on the credit history of a consumer.

[23] https://ieeexplore.ieee.org/document/8540048
[24] https://arxiv.org/pdf/1706.01730.pdf
[25] https://dl.acm.org/doi/abs/10.1145/3205977.3205993

Funded by the Horizon 2020 Framework Programme of the European Union

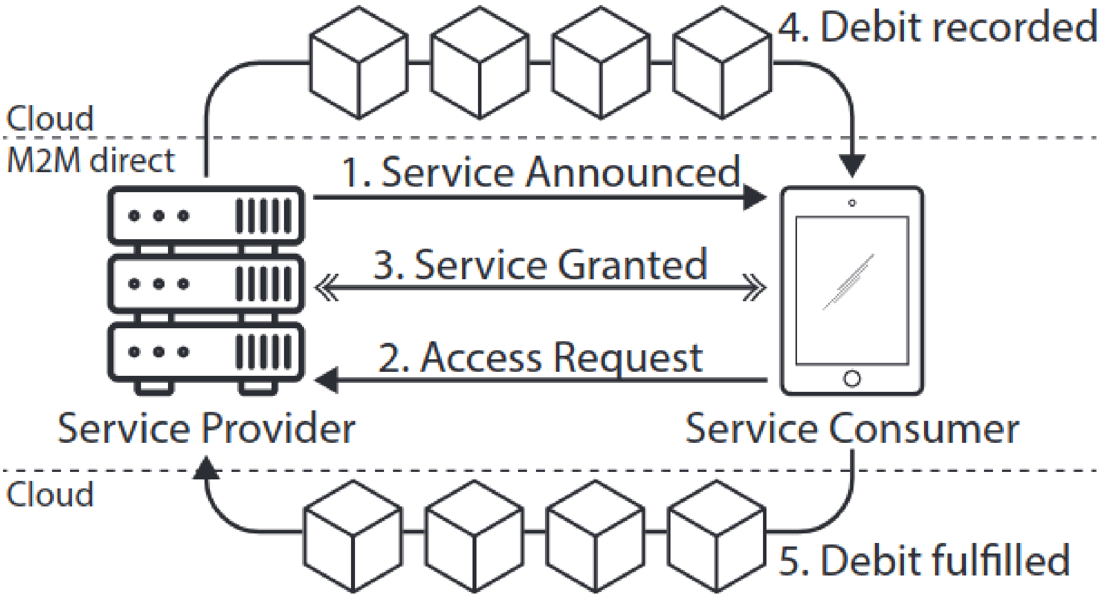


FIGURE 5: ARCHITECTURE INCLUDES OBLIGATION CHAIN.[26]

Moreover, in environments such as vehicular ad hoc networks (VANETs), the high topology variability complicates the timely evaluation of the trustworthiness of the received messages. Blockchain-based trust and reputation systems are addressing such challenges with sophisticated solutions. In particular, vehicles inside a VANET are using such systems in order to calculate the credibility of each event received from other vehicles as a message. All the messages received are aggregated based on the event they refer to and a decision is made for whether the event occurred or not. Once the in-vehicle assessment is performed, a vehicle receives a reputation score based on its behaviour[27] [28].

Blockchain-based trust and reputation systems are used in the swarm robotics field where legit information sharing is important and where the quality of the generated knowledge may be affected by malfunctioning robots. In the following approach, the Ethereum-based trust and reputation system is detecting robots that perform arbitrarily faulty or malicious behaviour[29]. Each robot represents a node in an Ethereum private network. If the distance between any two robots is smaller than 50 cm, the robots can exchange their blockchain information (blocks and transactions). The absolute difference between the value sent by the robot and the mean of all sent values of all robots is calculated and stored. This difference is then used to update the robot's reputation value.

In addition, there are important examples of permissioned blockchain-based trust and reputation systems in the Autonomous Systems (AS) domain. For instance, for

[26] https://ieeexplore.ieee.org/document/8540048

[27] https://ieeexplore.ieee.org/document/8358773

[28] https://ieeexplore.ieee.org/document/8455893

[29] https://iridia.ulb.ac.be/IridiaTrSeries/link/IridiaTr2018-009.pdf

the selection of an SP among possible providers, a solution characterizes the quality of an SP based on the conformance of its network performance with Service Level Agreement (SLA)s of interconnection agreements[30]. Dedicated smart contracts compute the SLA score for each AS and identify false testimonies about forwarding performance.

Furthermore, other permissioned blockchain-based trust and reputation systems provide solutions for multi-agent systems. For instance, in order to attain a trusted environment, a solution allows the network agents to interact with each other and enables tracking how their reputation changes after every interaction. The reputations are computed transparently using smart contracts and the underlying blockchain stores reputation values, as well as services and their evaluations, in order to ensure trustworthy interactions between the network agents (Figure 6) [31]. This kind of solutions offers a trusted environment with a legit reputation model in a distributed multi-agent system that is part of safety- and information-critical domain.

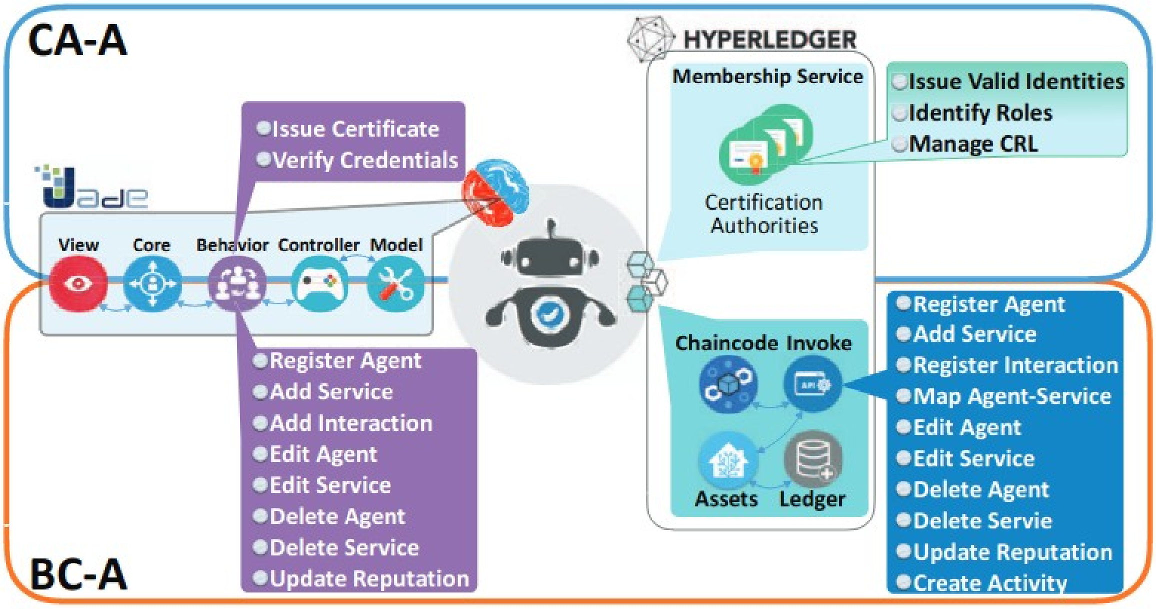


FIGURE 6: SYSTEM COMPONENTS OF CONCEPTUAL DESIGN.[32]

Finally, there exist general-purpose solutions of blockchain-based trust and reputation aiming to address the matter on a global scale. For instance, a blockchain-based authorization and access control system exploits a hybrid design for scalable and secure trust management and experience-derived reputation on a global scale, so that access delegations and trust assessments

[30] https://sands.kaust.edu.sa/papers/picking.anrw 18.pdf
[31] https://ieeexplore.ieee.org/document/8609678
[32] https://ieeexplore.ieee.org/document/8609678

are exchanged through a blockchain[33]. A global layer represents the backbone of the system consisting of miners that maintain a public blockchain, which can be instantiated upon an existing public blockchain (e.g. Ethereum). Miners are incentivized to invest computational power with fees paid for each operation on the blockchain, however, the decision on whether or not to request or delegate access to the resources is based on experience-derived reputation (Figure 7). Therefore, incorporating ratings by the interacting parties is an important part of the access delegation process that is achieved through the blockchain.

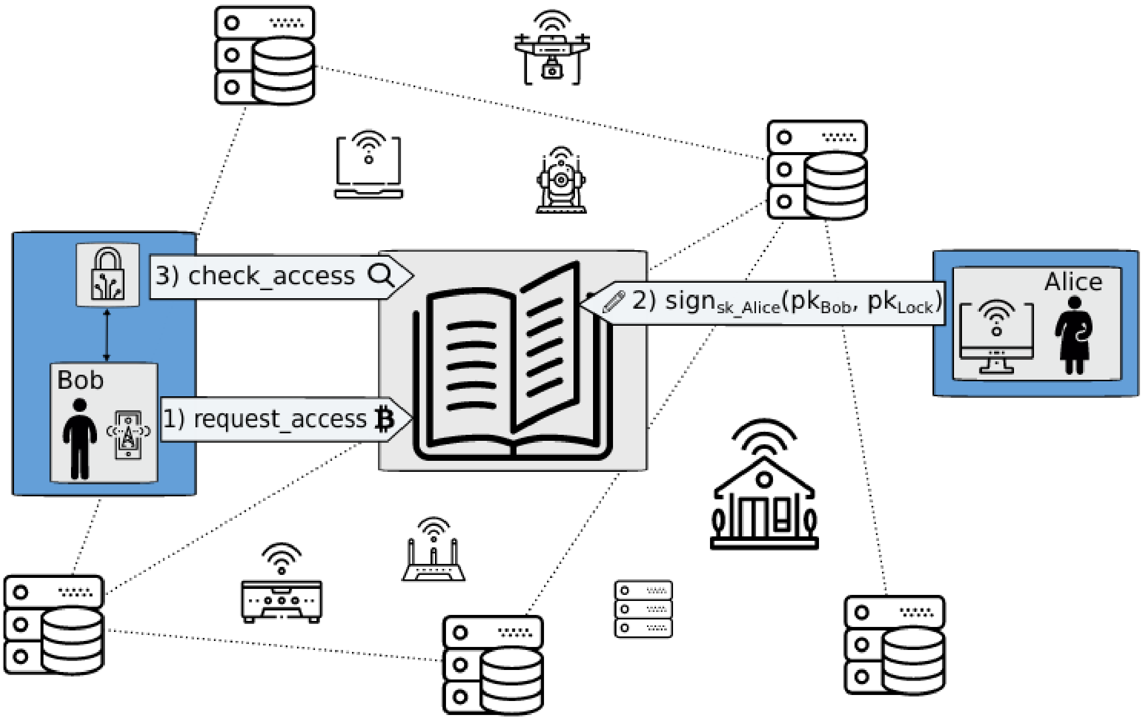


FIGURE 7: EXPERIENCE-DERIVED REPUTATION HOLISTIC SYSTEM.[34]

[33] https://dl.acm.org/doi/10.1145/3229565.3229569

[34] https://ieeexplore.ieee.org/document/8540048

2 Content transparency in blockchain technology

2.1 TRANSPARENCY STATE OF THE ART

One of the main attributes of blockchain technology is transparency. We usually refer to it as the auditability property that allows us to store information in a way that cannot be altered without a record of any change made. Through information traceability, any user in the network can audit and validate any transaction made between two peers or the history of all transactions. Cryptography and control mechanisms assure data integrity, being clear and trustful on the stored information. It ensures the constant availability of data about all the transactions that have happened in a context. Transparency in the blockchain is used in a myriad of use cases, from supply chain traceability (e.g. food origin, product components, fair trade), public tenders governance or accountable payments between users.

Three types of transparency are identified according to Bannister and Connolly [6]:

- Data transparency: what information is needed, who is involved, when and where it happens. Accessibility, understandability and accurate information enable this type of transparency for users. Data openness to blockchain users is normally served through block explorers where transaction data can be accessed and consulted.

- Process transparency: this response to the different interactions and events produced in the processing of data, the when, how and where something is performed. The validation process in nodes at storing hashed data and content and linking consecutive blocks enables data integrity and traceability. Process transparency is assured by these cryptographic functions and mechanisms.

- Decision transparency: concerns with the purpose (why and how decisions are made). Transparency applies to other components of the on-chain system that enable automated, pre-defined rules and agreement policies: smart contracts, algorithms, etc.

Transparency can be defined as a set of properties providing information about:

- Accessibility, the quality of being easy to deal with.

- Usability, the quality of providing good use.
- Informativeness, the quality of providing or conveying information.
- Understandability, the quality of comprehensible language.
- Auditability, the ability to examine with the intent of verification.

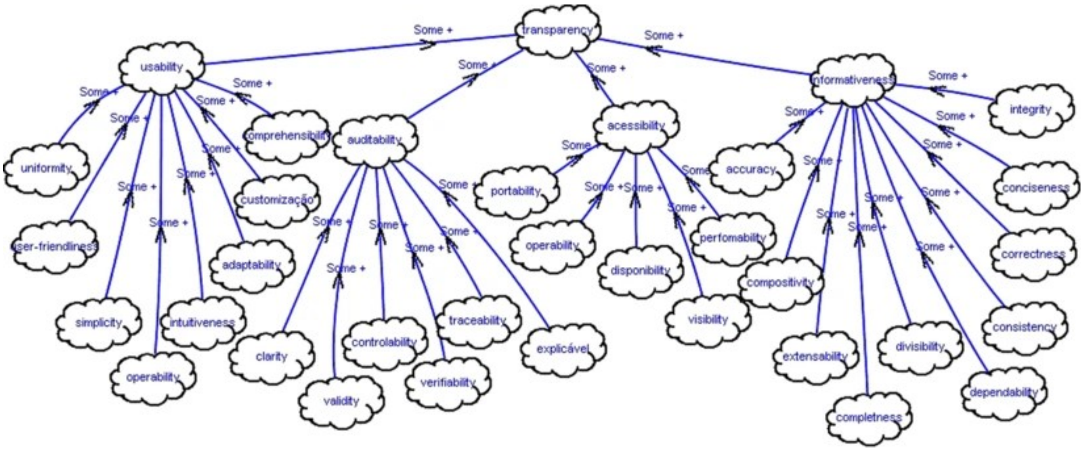


FIGURE 8: TRANSPARENCY SOFTGOAL INTERDEPENDENCY GRAPH[35].

These five characteristics shall be considered to design a system with implemented transparency. In the blockchain domain, the implementation is not simple and some of these interdependencies above have not yet been identified. How transparent is the blockchain?

Table 1 collects the set of properties and sub-properties around transparency that might be supported (fully or partially) or not by a blockchain system [36]:

TABLE 1: TRANSPARENCY CHARACTERISTICS AND BLOCKCHAIN SUPPORT.

Characteristic	Sub-property	Identified in Blockchain
Accessibility	Availability	Partially
	Portability	Yes
	Publicity	No
Usability	Adaptability	Partially
	Intuitiveness	No
	Operatibility	Yes
	Performability	Yes

[35] https://link.springer.com/article/10.1007/s12599-010-0102-z

[36] https://www.researchgate.net/publication/336345730_Transparency_Challenges_in_Blockchain

Funded by the Horizon 2020
Framework Programme of the European Union

	Simplicity	No
	Uniformity	Yes
	User-friendliness	Partially
Informativeness	Comparable	Yes
	Accuracy	Yes
	Clarity	No
	Completeness	Yes
	Consistence	Partially
	Correctness	Partially
	Current	Partially
	Integrity	Yes
Understandability	Composability	Partially
	Conciseness	Yes
	Dependability	Yes
	Decomposability	Partially
	Extensibility	Partially
Auditability	Accountability	No
	Controllability	Yes
	Traceability	Yes
	Validity	Yes
	Verifiability	Yes

This analysis provides some relevant conclusions for the implementation of full transparency on the blockchain:

- Information correctness is critical to have a shared source of truth. Thus, we need to guarantee that transactions verification and smart contract are correct.
- Transparency goes beyond transactions between peers and refers to the blockchain infrastructure, how it is built and use resources, publicity of documentation, conditions and restrictions, that effectively create a democratic ecosystem. Algorithmic transparency is contained under this topic and its openness on the underlying actions, structure and purpose also enforces the accountability characteristic of the system.
- Usability and user-friendliness are key to facilitate the consultation of information and the view and use of smart contracts inintuitively and simple

for users. Access to the ledger resources enhances overall transparency, including documentation and interfaces.

- Accountability (e.g. the use of available resources, conditions for performing actions, information sources is only associated with the infrastructure itself and often not supported for the operations performed on the blockchain. Proof of accountability is a vital property of any system and shall be put at the same level of security, trust, privacy and decentralization as part of the system transparency, being mutually interrelated. It incentivizes all network participants to behave honestly. Some issues arise around it: Immutability is a double-sided weapon when it comes to deal with illegitimate yet valid transactions on the ledger. Anonymity and pseudo-anonymity can hinder the right identification of malicious users. Third-party auditors and regulators are now substituted by crypto-algorithms that enforce trust but even math-based systems can be flawed by human errors in their design (e.g. Ethereum's DAO issue in 2016 or Zcash bug in 2018 produced several losses of crypto-currencies on users).

2.2 TRANSPARENCY VERSUS PRIVACY

Blockchain transparency comes with different degrees depending on the application and the blockchain topology: many public permissionless ledgers like Bitcoin normally offer a high level of transparency because transactions are publicly available to all network participants, whereas permissioned blockchains offer a higher degree of confidentiality between authenticated users that exchange transactions as a trade-off with transparency, which are only available for those involved in the operation. The tension between privacy and transparency attributes in the design of a blockchain system can limit its potential and innovations are placed to balance the importance of both attributes in order to assure that confidentiality is not incompatible with being clear and transparent.

Data privacy issues and compliance with European legal policies like GDPR (General Data Protection Regulation) must be addressed on a mandatory basis. Specially those aspects related directly with the "right to be forgotten" since the immutability property of blockchain disables the possibility to modify or delete data. There are different alternatives to address these policies without jeopardizing personal data privacy and transparency. Transactions on the blockchain shall contain a hashed point to where the actual data and content is stored (normally in centralized or decentralized databases) or provide solid proofs of the existence and value of data without disclosing it.

2.3 ZERO-KNOWLEDGE PROOF

Zero-knowledge proof protocols allow data to be verified without revealing that data. Thus, we can have transparency with the power of privacy altogether. It uses cryptographic algorithms so that various parties can verify the veracity of information (transparency) without having to share the data that compose it.

Several improvements to the protocol have occurred lately. zk-SNARKs (Zero-Knowledge Succinct Non-Interactive Argument of Knowledge) there is no interaction between the two parties involved in the transaction and proofs become more complex but more efficient in size. In this case, the protocol requires the initialization of the process with a trusted set-up phase which requires both the prover and verifier to have access to some common knowledge.

In 2018, zk-STARK protocol was introduced offering transparency by removing the need of a trusted set-up. ZK-STARKs (Zero-Knowledge Scalable Transparent ARguments of Knowledge) is a type of cryptographic proof protocol that enables users to share validated data without the data being revealed to the third-party, also known as zero-knowledge proof, in a way that is publicly verifiable and claiming to be more efficient and secure than its predecessor.

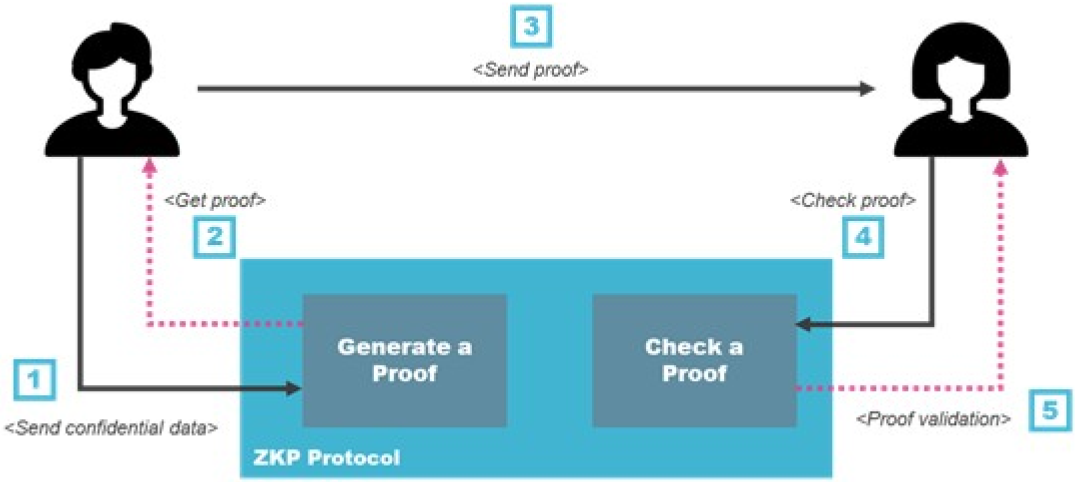


FIGURE 9: NON-INTERACTIVE ZERO-KNOWLEDGE PROOF SCHEMA.

Some novel blockchain-based architecture[37] for content delivery networks is considering how to exploit the advances of the blockchain technology to provide a decentralized and secure platform to connect content providers with users.

[37] https://ieeexplore.ieee.org/document/8885904

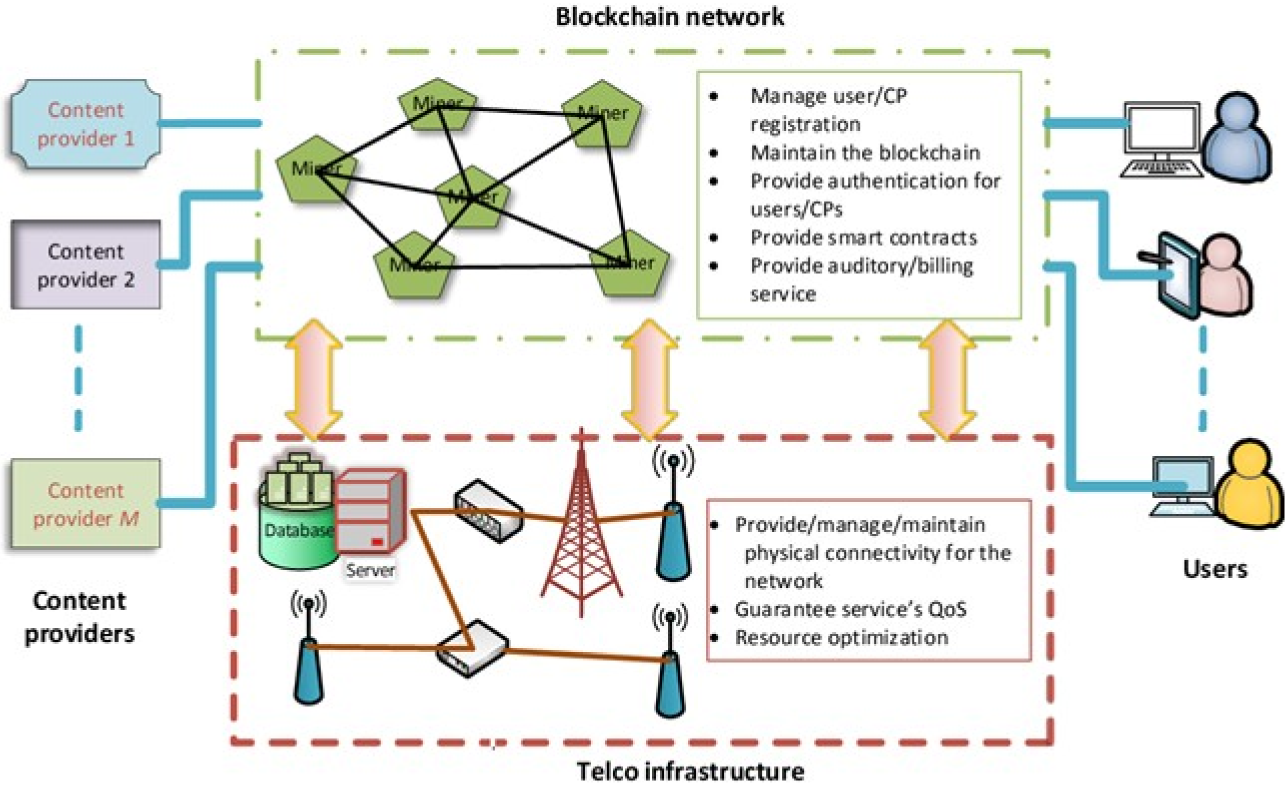
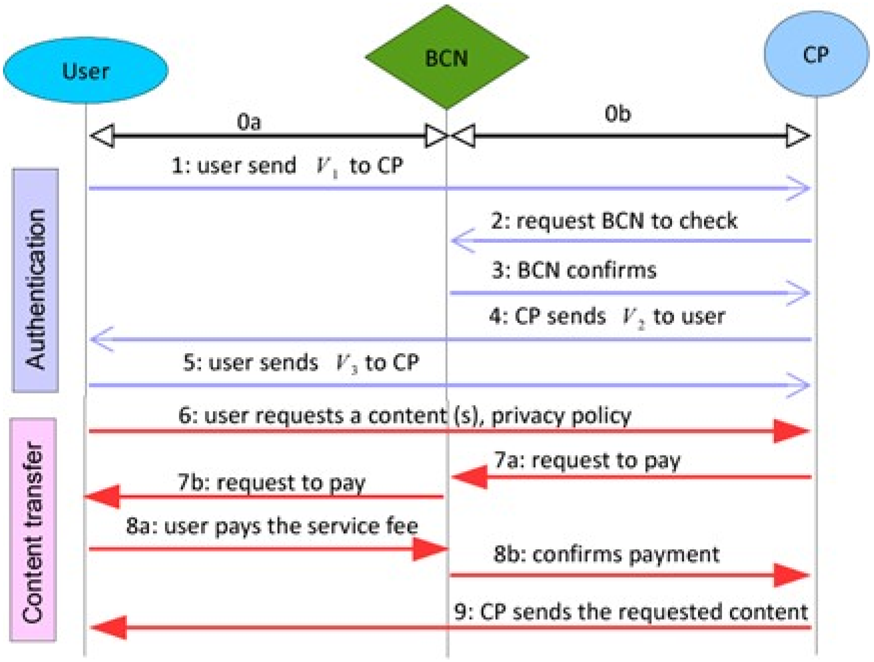


FIGURE 10: SYSTEM ARCHITECTURE OF A BLOCKCHAIN-BASED CONTENT DELIVERY NETWORK[38].

Blockchain is addressing user registration and subscription to different content providers and protecting user privacy using digital identities and cryptographic keys in the blockchain and enabling content access payments with smart contracts.



[38] https://ieeexplore.ieee.org/document/8885904

FIGURE 11: USERS AND CONTENT PROVIDERS IN A BLOCKCHAIN-BASED CONTENT DELIVERY NETWORK.

Digital signature mechanisms and required to verify identity but are not sufficient. Digital ID implementations through permissioned blockchain systems are necessary to enhance accountability and transparency on a content delivery network. If content providers are using personal data to personalize their offering to users according to their preferences, it is likely to use decentralized digital ID solutions like Self-Sovereign Identity (SSI) to ensure that personal data is not compromised and that the system is fully GDPR-compliant.

3 PRIVACY IN BLOCKCHAIN TECHNOLOGY

Blockchain and, in general, Distributed Ledger Technologies (DLT) have emerged as effective enterprise transformation tools. They provide capabilities beyond traditional databases to share data and manage workflow throughout an enterprise and across its ecosystem of customers, partners and suppliers in a trusted manner without central control.

Nonetheless, the adoption of Blockchain/DLT has been proceeding slowly in established markets. This is partly due to a high degree of legal uncertainty as to its compliance with data protection regulations. One of the reasons being that legal frameworks, such as the General Data Protection Regulation[39] (GDPR) initially apply to data processing in single server structures operated by a legally and technically tangible intermediary.

One of the challenges concerning the compliance with data protection regulation is determining whether remains particularly unclear under which circumstances the data processed in an IT system using Blockchain/DLT is considered personal data. According to art. 4(1) and recital 30 GDPR personal data is defined as "that information relating to a natural person that can be identified with that person". This definition is in line with the definition of Personally Identifiable Information (PII) used in ISO/EC 27701[40], where it is stated that "to determine whether a PII principal is identifiable, account should be taken of all the means which can reasonably be used by the privacy stakeholder holding the data, or by any other party, to establish the link between the set of PII and the natural person". At the end, personal data is the aggregation of a piece of information [about a subject] and an identifier [of such subject]. However, both the information and the identifier do not need to be explicitly present in the data. Having information [about a subject] that can be derived from the context or with the help of available external data sources could be enough to identify (obtain the identifier) of such subject.

A second challenge with Blockchain/DLT-systems comes from legal analyses that rely on popular scientific simplifications of Blockchain/DLT. This tends to disregard potential technical approaches to resolving legal provisions.

The legal uncertainty regarding Blockchain/DLT and data protection law leads to the widespread assessment that "personal data should not be processed on the blockchain". But (how) is this even possible? First attempts to resolve the issue by storing personal data "off-chain", "anonymizing" it or only feeding M2M data

[39] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=ES

[40] https://www.iso.org/standard/71670.html

(machine to machine) into a Blockchain/DLT-system in order to, seemingly, circumvent data protection law (which only regulates personal data) not always prove to be optimal.

Therefore, a key aspect to clarify could be the degree to which a natural person, in Blockchain/DLT-systems, is identifiable, providing developers with tools and methods (i.e., technical design patterns) that preserve privacy and reduce the risk of legal uncertainty and therefore raise the data protection level in an IT system. That is why current industry efforts look at the risk for the data subject trying to identify technical measures that increase the difficulty to recreate any personal reference to a data subject affecting her rights and freedoms.

In order for developers to address specific risks in a Blockchain/DLT-system, the relevant technical measures and the normative principles of data protection should be considered. German standard DIN SPEC 4997:2020-04[41] provides a quick summary of recommended technical measures to address common risks (Figure 12).

[41] https://www.beuth.de/en/technical-rule/din-spec-4997/321277504

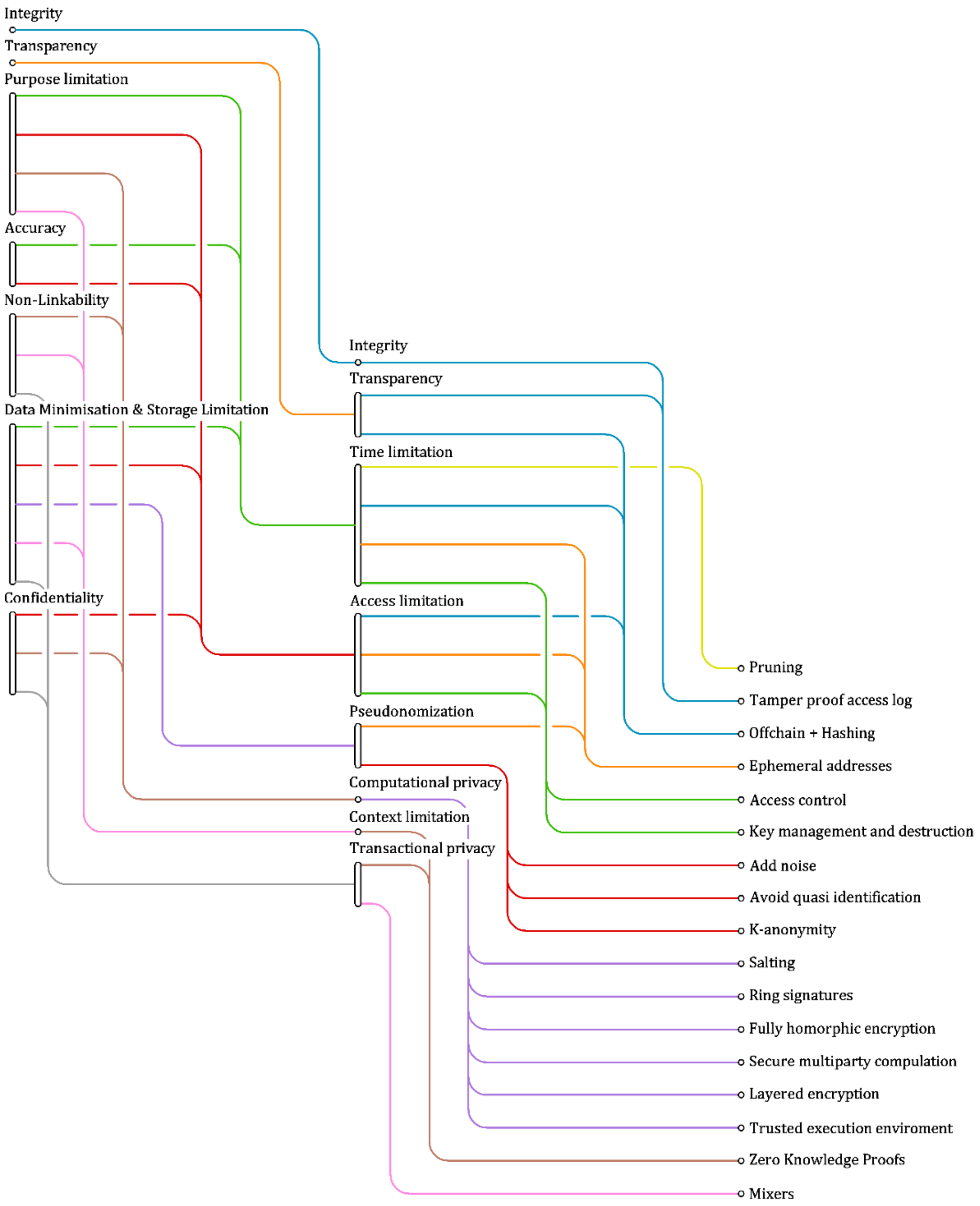


FIGURE 12: LEGAL PRINCIPLES VS. TECHNICAL MEASURES (SOURCE: DIN).

At the point of writing of this document limited research is available on the impact of technical measures on the risk of processing personal data using Blockchain/DLT.

That means that before facing any technical development on Blockchain/DLT, it is highly recommended to apply principles of proactive responsibility that guide any work on Blockchain/DLT, like the "Privacy by Design" principle (art. 25 GDPR). In that sense, DIN SPEC 4997, for instance, aims to lay a framework, accessible to lawyers and computer scientists alike, by providing:

- a description of the functional requirements for Blockchain/DLT systems to achieve compliance with GDPR requirements;
- a guide to the handling of personal data using Blockchain/DLT;
- architectural blueprints to illustrate the uses of Blockchain/DLT to improve privacy; and,
- procedures for business processes for the iterative maintenance and quality assurance of data protection.

Blockchain has sparked interest in its possibilities beyond cryptocurrency. The models that can be implemented using Blockchain/DLT are as unique as the imagination of the developers and can be implemented in many ways and with so many particularities. Trying to give a simple and generic answer to the benefits and problems posed by their use in treatment would be a mistake.

Among the wide range of possible uses in other areas and sectors beyond financial are supply chain management, asset tokenization, traceability and inventory of goods, management of digital identity, fraud identification systems, voting, property registries, development of financial services, etc. But one deserves to be highlighted among all of them: digital identity.

Recently released Spanish standard UNE 71307-1:2020[42] defines the concept of "self-sovereign identity" (SSI) and the more general concept of "decentralized identity" based on the idea that identity-bound subjects control the administration of their own digital identities. This requires the user's ability to create and use a digital identity across multiple scenarios and to have sole control over how, when and where such identity is used, thus ensuring user autonomy (user-centric).

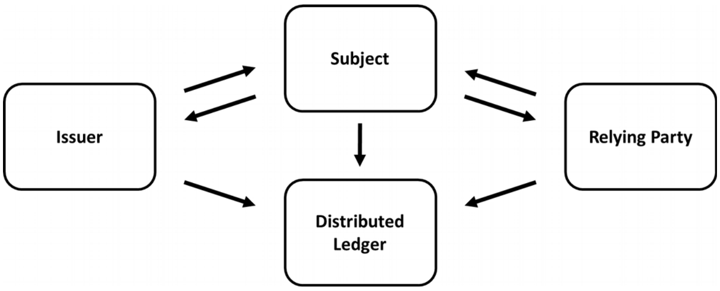


FIGURE 13: SELF-SOVEREIGN IDENTITY MANAGEMENT ACTORS (SOURCE: UNE)

As depicted on Figure 13, Subjects needs to gather [personal] information (Credentials) about themselves, issued by one or more Issuers, in order to share such [personal] information (Presentations) with the Relying Parties. At the same

[42] https://www.en.une.org/encuentra-tu-norma/busca-tu-norma/norma?c=N0064986

time, these latter must be able to authenticate the Subjects and verify the information received, to provide services or to get engaged in a transaction. To be valuable and admissible, all identity-related actions performed by the entities should be legally attributable to them. Such possibility is supported by registering digital evidence about the actions performed by the participants on a Blockchain/DLT (such as issuance, acceptance, presentation, revocation, withdrawal, etc.).

Any solution implementing a decentralized identity management framework has to be also compliant with the regulations applicable to Personally Identifiable Information and to keep the privacy of the different parties involved. Therefore, the framework must be designed according to security and privacy-by-design and by-default principles, ensuring that any relation between Subject and Issuer, and subsequently between Subject and Relying Party, is not traceable by any third party using public information registered on the Blockchain/DLT.

To ensure sole control by and autonomy of Subjects concerning their identity attributes and Credentials, actions performed by Subjects should not require the intervention or approval of the Issuer or Relying Party.

While it may not yet be possible to solve all of the challenges posed by the GDPR to the implementation of Blockchain/DLT solutions, progress can be made if the interested parties work together openly and pragmatically.

A Blockchain solution that respects the fundamental principles of data protection and privacy is achievable, and four key elements necessary to achieve that aim are:

- use of permissioned Blockchain/DLT;
- avoid the storing of personal data on the Blockchain/DLT;
- implement a detailed governance framework; and,
- employ innovative solutions to traditional data protection problems.

Innovative solutions to data protection challenges will only succeed with the understanding and support of regulators and lawmakers. Regulatory authorities should take the steps necessary to address the outstanding privacy challenges posed by Blockchain technology, most importantly, in relation to the use of encryption as a means of anonymization and deletion of personal data. There is a risk that, if steps are not taken by regulators and lawmakers to bridge the gap between data protection law and Blockchain/DLT, a slowing in (or even end to) advancements in the area of Blockchain/DLT solutions might be seen. Such an outcome would ultimately be detrimental to technological developments that may have the capacity to deliver substantial benefits to the world as a whole.

According to the Center for Global Enterprise (CGE)'s Digital Supply Chain Institute[43] there are a series of "do's" and "do not's" that one should take into account when determining if and/or how to apply blockchain in an environment of GDPR:

- do determine what type of data you are processing;
- do determine what regulations this processing triggers;
- do follow industry best practices for data processing;
- don't use a public blockchain to implement a solution that uses personal data;
- don't permanently store personal data on-chain; and,
- don't ignore regulatory guidance documents.

Additionally (Figure 14), DSCI provides a procedure in order to help enterprises adopt the right lane – to apply Blockchain or not – when trying to devise the right solution in response to their business needs, within a regulated data protection environment.

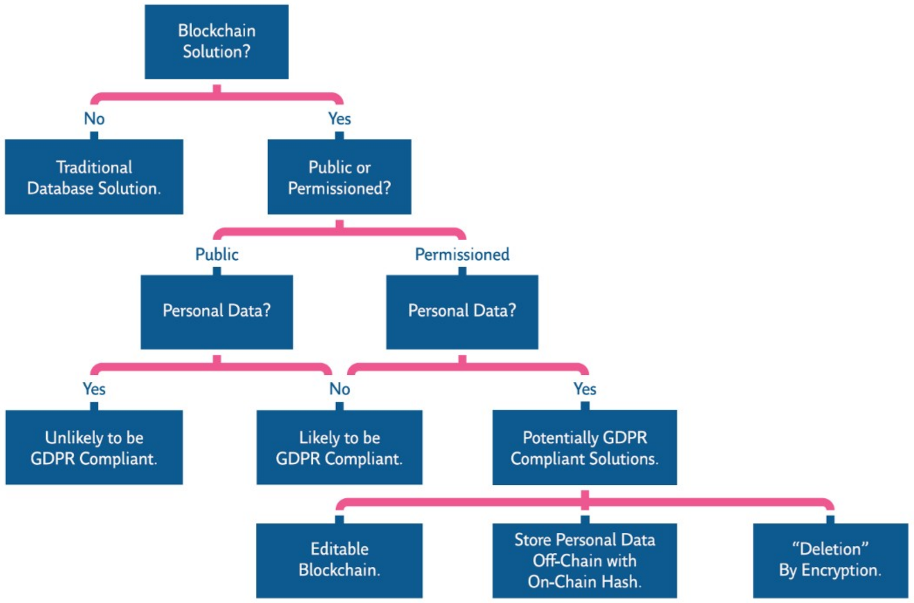

Blockchain Solution?
No
Yes
Traditional Database Solution.
Public or Permissioned?
Public
Permissioned
Personal Data?
Personal Data?
Yes
No
Yes
Unlikely to be GDPR Compliant.
Likely to be GDPR Compliant.
Potentially GDPR Compliant Solutions.
Editable Blockchain.
Store Personal Data Off-Chain with On-Chain Hash.
"Deletion" By Encryption.

FIGURE 14: BLOCKCHAIN IMPLEMENTATION DECISION TREE (SOURCE: DCSI)

[43] https://www.marinetransportint.com/wp-content/uploads/2019/02/GDPR_Blockchain_Short_Final.pdf

4 Proof-of-Validity, Proof-of-Location, Proof-of-X

4.1 Traditional Proof of Location Approaches

In recent years, the wide use of mobile devices has led to increased usage of applications based on the current location of users, the Location Based Services (LBS). Indicative examples of such location-based applications are Foursquare [7], and Yelp [8].

Proof of location can be seen as a digital certificate that attests someone's presence at a certain geographic location, at a certain time. Different approaches have been examined, which depend on infrastructure, named "infrastructure-dependent", and others which have recently emerged and are not based on reliant on the underlying infrastructure.

It has been noticed that different award incentives are given to mobile users by LBS based on their activity at a particular location, for example to users who check infrequently at a specific place. The development of additional services has been examined in the access control system where proximity detection is required [9].

For instance, in a critical health-care system, the authorized personnel e.g., doctors may be allowed to access the complete information/records of a patient if present in/around the vicinity of a hospital, else may have limited access to the records. Similarly, for entertainment applications, an online movie downloading system, such as Netflix, may provide the content only to the users if present in/around an area of interest, and surcharge other customers present at a different location/place or deny access to multimedia content because of copyright rules.

In the context of all these applications, it is important to verify the correct location of a user and prevent the false claim of a location to gain benefits. An analysis of Foursquare indicated that a big number of users provide false evidence and lie about their location. Although GPS based solutions tried to fill this gap, the limited signal coverage in indoor environments is a drawback to implementations based on this technology. Other solutions use wireless networks based on cell towers or Wi-Fi access points. The Android Network Location Provider (ANLP) uses both cell tower and Wi-Fi to determine the location.

Bluetooth location systems are based on beacons spread over the area of interest. Since the range is limited (below 10 meters), it is assumed that, if a user can detect a beacon, then she is near the location. To cover a large area, many beacons are needed which imposes high hardware and installation costs. To this direction Ferreira et al. presented a location proof system for mobile devices which allows the use of geographical coordinates, Wi-Fi fingerprinting and Bluetooth beacons, exhibiting bigger effectiveness in crowded locations where a user can obtain location proofs with a diversity of witnesses. Proximity systems can verify for example maximum communication latency to assert proximity. However, they are vulnerable to relay and signal amplification attacks.

Li et al. [16] examined the location cheating attack in database-driven cognitive radio networks, where the lack of a mechanism for the location's verification could lead to vulnerabilities. In this case, an attacker could spoof other users to another location and make them query the database with the wrong location or allow a malicious user to forge location arbitrarily and query the database for services. To thwart this attack, an infrastructure-based approach was proposed that relies on the existing WiFi AP network or cellular network to provide secure and privacy location proof.

Users themselves create content and share it through other services, with the determination and verification of the location of the users being, in many cases of utmost importance. Information regarding a user's location can be gathered in two main ways: self-location and remote-location. In the case of a self-location, the location data is supplied to the service by the user or automatically determined by the user's device before being transmitted to services, and the validity of the location information cannot be guaranteed to be accurate, since the users may provide false data, tamper their devices or find other weak points of the positioning technique used. On the other hand, in the situation called remote-location, an external entity is responsible to position the user, rather than the user providing their location directly to the service. The external entity might not only be a single entity but a group of collaborating entities exchanging information to determine with precision the position of the user. An example of such a situation would be a mobile telephony network, where, through the use of triangulation and other techniques, mobile phone companies can compute the location of a user within their network, provided that the user remains within the area of the network. The primary weakness of remote location is its reliance on the infrastructure to provide a location. In general, various solutions to the problem of location verification exist, ranging from distance-bounding protocols to location proof systems.

4.2 Proof of location-based Blockchain

The recent years, there has been a research effort to exploit the features of decentralised systems and propose a novel and more effective ways to prove the location of a user and avoid false evidence and inaccurate location claim. To this direction, Blockchain Proof of Location systems has been examined, which based on the decentralised nature of peer-to-peer networks assure high levels of privacy, since a central authority for proving the identity of a person is not required.

Therefore, Proof of Location has been revised in conjunction with blockchain technology to achieve a distributed and decentralised consensus about the position of events or agents in space and eventually in time. This concept acquires particular importance in the context blockchain smart contracts, with the aim to describe constraints related to spatial properties of the involved parties. The goal of PoL is to have consensus on whether an event or agent is verifiably at a certain point in space and eventually in time. Smart contracts which handle interactions with external services and sensors may force that a performed activity is valid only if it is executed by an agent which is verifiably at a certain location, or that two or more agents can fulfil the contract only when they are nearby to each other and such a certificate can be built through a distributed, trustless and decentralised consensus mechanism.

Active projects to this direction such as FOAM [10] and Platin [11] try to provide an infrastructure which acts as a location layer to be used by Smart Contracts. In the research work of Brambilla et al. [12], the report of a false location is not feasible since low range technologies are used and the proofs of location are stored in the Blockchain, which makes sometimes the transmission of a proof impossible. In their approach, they produce proofs of location, i.e., digital certificates that attest someone’s presence at a certain geographic location, at some point in time whereby Location Based Services are responsible to validate user locations. Different Location Based Services related attacks [13] have been considered to examine the robustness of their method such as “cheating on own geographic location”, “Replaying proofs of location”, “Colluding with other peers” etc.

4.3 Proof of Claim/Challenge-Response Tests

Similarly, to the Proof of Location or Proof of Presence concepts, various technologies and techniques have been developed to determine the truthfulness of various other claims provided by users, related to the data or metadata provided by them (e.g. when sharing news online). A characteristic example is the concept of CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) [14], a challenge-response test used to determine whether or not the user is human. Crowdsourcing [15] is a sourcing model in which individuals or organisations obtain goods and services, including ideas and finances, from a large, relatively open and often rapidly-evolving group of internet users; it divides work between participants to achieve a cumulative result. As a mode of sourcing, crowdsourcing existed before the digital age. Advantages of using crowdsourcing may include improved costs, speed, quality, flexibility, scalability, or diversity. Mobile crowdsourcing specifically involves activities that take place on smartphones or mobile platforms, frequently characterized by geolocation technologies. This allows for real-time data gathering and gives projects greater reach and accessibility.

4.4 Security and Threats analysis

In this section, we examine some of the major threats, security aspects and risks which affect schemes, technologies and approaches of the previous sections. Malicious peers may attempt to provide false evidence to honest peers. In this case, a well-defined approach should be able to detect malicious peers and allow the community of honest ones to penalize them.

1. Malicious peer cheats to obtain false proof: a user could attempt to cheat to gain advantages or access to services. As an example, in [16] the authors identified the attack coined as location cheating attack in databased-driven CRNs, in which users can cheat their locations and query the database for services. Additionally, a malicious user could attempt to issue multiple identities and produce false proofs.

2. Malicious peer provides false evidence on another peer’s proof: in this case, a malicious peer could produce false claims about other peer’s evidence e.g. location. It should be noted that a well-defined blockchain-based approach could verify the true identity of a user thanks to asymmetric cryptography mechanisms and the exploitation of public keys, private keys and digital signatures.

3. Malicious peer attempts to reuse a proof: the attempt to replay and re-broadcast of an older/outdated proof should be in time detected. While this proof was correctly issued, the time aspect should not be neglected. Blockchain timestamps provide the means to detect these attacks before they are completed.

4. Malicious peers jointly attempt to falsify a proof: this kind of attack has similarities with Sybil attacks [17]. For example, two malicious peers agree upon producing a proof of location attesting that their geographic locations are different from the actual ones. Then, they broadcast the false proof of location into the network.

5. Identity reveals of peers: while some honest users provide their contribution into producing valid proofs, some attackers could attempt to extract their real identity. For example, in proof of location, an honest user verifies the location of others who are in proximity. As a result, an attacker could infer his/her location. Changing the identity as in [18] or covering the digital traces are some of the ways to prevent this kind of attacks.

6. Adjustment to dynamic environments: in the case of a dynamic environment, a mechanism should correctly detect and decide upon the suitable witness who will assist in a proof.

REFERENCES

[1] Joachim Hafkesbrink & Markus Schroll, 2011. "Innovation 3.0: embedding into community knowledge - collaborative organizational learning beyond open innovation," Journal of Innovation Economics, De Boeck Université, vol. 0(1), pages 55-92.

[2] Francesco Restuccia, Nirnay Ghosh, Shameek Bhattacharjee, Sajal K. Das, and Tommaso Melodia. 2017. Quality of Information in Mobile Crowdsensing: Survey and Research Challenges. ACM Trans. Sen. Netw. 13, 4, Article 34 (December 2017), 43 pages. DOI:https://doi.org/10.1145/3139256

[3] Abdelmuttlib Ibrahim Abdalla Ahmed, Siti Hafizah Ab Hamid, Abdullah Gani, Suleman khan, Muhammad Khurram Khan, "Trust and reputation for Internet of Things: Fundamentals, taxonomy, and open research challenges," in Journal of Network and Computer Applications, vol. 145, 2019, https://doi.org/10.1016/j.jnca.2019.102409.

[4] K. Salah, M. H. U. Rehman, N. Nizamuddin and A. Al-Fuqaha, "Blockchain for AI: Review and Open Research Challenges," in IEEE Access, vol. 7, pp. 10127-10149, 2019, doi: 10.1109/ACCESS.2018.2890507.

[5] Ciampaglia, G. L., Mantzarlis, A., Maus, G., & Menczer, F. (2018). Research Challenges of Digital Misinformation: Toward a Trustworthy Web. AI Magazine, 39(1), 65-74. https://doi.org/10.1609/aimag.v39i1.2783

[6] Bannister, F. and Connolly, R. 2011. The Trouble with Transparency: A Critical Review of Openness in e-Government.Policy & Internet. 3, 1 (Jan. 2011), 158–187. https://doi.org/10.2202/1944-2866.1076.

[7] Foursquare, https://foursquare.com/

[8] Yelp," http://www.yelp.com/

[9] Javali, C., Revadigar, G., Rasmussen, K. B., Hu, W., & Jha, S. (2016, November). I am Alice, I was in wonderland: secure location proof generation and verification protocol. In 2016 IEEE 41st conference on local computer networks (LCN) (pp. 477-485)].

[10] Foamspace Corp, «FOAM Whitepaper,». Available: https://foam.space/publicAssets/FOAM_Whitepaper.pdf

[11] https://eos.io/news/with-platin-location-location-is-more-proof-than-platitude/

[12] Brambilla, G., Amoretti, M., & Zanichelli, F. (2016). Using blockchain for peer-to-peer proof-of-location. arXiv preprint arXiv:1607.00174.

[13] Rasib Khan, Shams Zawoad, Md Munirul Haque, and Ragib Hasan. 'Who, When,

and Where?' Location Proof Assertion for Mobile Devices, pages 146–162. Springer Berlin Heidelberg, Berlin, Heidelberg, 2014

[14] von Ahn, L., Blum, M., Hopper, N. J., Langford, J., Biham, E.: CAPTCHA: Using Hard AI Problems for Security. In: Advances in Cryptology — EUROCRYPT 2003, pp. 294-311. Springer Berlin Heidelberg (2003)

[15] Brabham, Daren C. Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. In: Convergence, vol. 14, Issue 1, pp. 75 – 90, SAGE Publications Ltd (2008)

[16] Li, Yi, et al. "Privacy-preserving location proof for securing large-scale database-driven cognitive radio networks." IEEE Internet of Things Journal 3.4 (2015): 563-571.

[17] John R. Douceur. The Sybil Attack. In Revised Papers from the First International Workshop on Peer-to-Peer Systems, IPTPS '02, 2002

[18] Zhichao Zhu and Guohong Cao. Toward Privacy Preserving and Collusion Resistance in a Location Proof Updating System. IEEE Transactions on Mobile Computing, 12(1):51–64, January 2013.