The Current Limitations of Blockchain Traceability: Challenges from Industry

N. Sánchez-Gómez\textsuperscript{1}, M. Mejias Risoto\textsuperscript{1}, J. M. Ramos-Cueli\textsuperscript{2}, T. Wojdyński\textsuperscript{3}, D. Lizcano\textsuperscript{4} and J. Torres-Valderrama\textsuperscript{1}

\textsuperscript{1}University of Seville, ETS de Ingeniería Informática, Web Engineering and Early Testing Research Group, Seville, Spain
\textsuperscript{2}Soltel Group, Department of Research and Development, Leonardo da Vinci, 13, 41092 Seville, Spain
\textsuperscript{3}School of Banking and Management, Cracow, Poland
\textsuperscript{4}School of Computer Science, Madrid Open University, UDIMA, Madrid, Spain

Keywords: Blockchain, Traceability, Model-based Software Engineering, Common Information Model.

Abstract: Blockchain technology is a chain of cryptographically linked blocks. It was designed to be immutable, so that the identity and traceability of the information entered would be guaranteed. After analyzing several traceability solutions, in the context of a Spanish company project, it was found that in order for a traceability solution to be efficient and agile, an additional layer is necessary in the blockchain. Since this need originated in the industrial sector, the subject has awakened considerable interest in the research community. This paper explains why the extra layer is essential and why it should ideally be totally independent of the information that is recorded on the blockchain network. Although data in a blockchain network is immutable, the paper also outlines the need for additional verification mechanisms capable of determining whether the raw data was correct. Finally, it includes planned future work.

1 INTRODUCTION

Blockchain technology has evolved a lot since the introduction of Bitcoin in 2008. This cryptocurrency invented by an unknown person, or group of people, using the name Satoshi Nakamoto (Nakamoto, 2008), first appeared in 2009 (Joshua, 2011), when its implementation was released as open-source software. Engineers and researchers are now realizing the benefits of blockchain technology and looking for ways to integrate it into their infrastructures in numerous sectors, from finance and supply chains to health (Al-Saqaf and Seidler, 2017). Due to its decentralization and reliability, blockchain technology is potentially beneficial to businesses, its enhanced security, greater transparency, and easier traceability opening up many new opportunities (Song et al., 2019).

Traceability is one of blockchain’s biggest advantages over other solutions. Blockchain networks have intrinsic mechanisms that facilitate both external and internal audits (Westerkamp et al., 2020). Moreover, blockchain networks were designed to be immutable, so that the identity and traceability of the information entered would be guaranteed. Although several different techniques can be applied (Cleland-Huang et al., 2014), blockchain offers additional functionality in terms of data security, and this is very interesting for certain businesses where it is a prerequisite to have a secure, reliable record in which information remains unchanged and traceable.

After analyzing several product/service traceability solutions, in the context of Common Information Model for Traceability Project (Soltel Group’s CIMT Project is currently under development), it was found that a flexible traceability module is required: a module totally independent of the information registered in the blockchain network and one which, depending on the user case, does not need to be reimplemented and allows the tracking process to be applied to different components at the same time.
The CIMT Project is part of the SERVICECHAIN initiative, which is co-financed with FEDER funds. The main objective of this initiative is to solve the challenges of the Spanish companies in the management of digital identity, reliability and traceability of the goods and services transactions against the requirement of digital transformation and existing regulation. Soltel Group is part of the consortium and has the responsibility to develop the technology that allows the traceability of the product’s life cycle, from its manufacture to the acquisition by the final consumer and subsequent transactions, until its disposal (at the end of its useful life). This process is mainly oriented to Industry 4.0 and energy.

This need, first identified in industry, has aroused the interest of the research community. Also, although data on the blockchain network remains unchanged, industry also recognizes that additional verification mechanisms should be introduced in order to show whether raw data is correct.

Blockchain technology can be applied in many different healthcare contexts. It can be used, for example, to manage Electronic Health Records (EHRs) or in the tracking research methods used in clinical trials to make drugs safer (Shahnaz et al., 2019). At a high level, blockchain application can be thought of as a two-layer EHR. One layer handles all changes: it is a live medical record that is updated, for example, when the patient visits the doctor. The second layer just monitors and tracks changes, creating a note of the “who, when and where” associated with all the changes that take place as the patient’s records evolve. Blockchain networks add changes to this layer and prevent any editing or changes, creating a secure, permanent record that is validated by all stake-holders with access to the data. Although the data stay immutable, the blockchain does not have a verification mechanism to prove whether the raw data were correct, making it necessary to have an intermediate layer to facilitate such mechanisms.

Since blockchain has the potential for many use cases and is applicable in numerous industries, using Blockchain-as-a-Service (Singh and Michels, 2018) would also allow companies to easily integrate blockchain technology and additional functions into their businesses without disruption to their daily processes. Therefore, in future work it will be necessary to address the proposed architecture.

The rest of this paper is organized as follows. Technical background is introduced in Section 2. Section 3 discusses the suitability of using blockchain in the context of product or service traceability. Section 4 analyses and proposes an additional layer for blockchain traceability, and Section 5 provides a set of conclusions and outlines future work.

2 TECHNICAL BACKGROUND

2.1 Blockchain and Smart Contract

Blockchain technology, based on distributed ledger technologies (DLT), is a chain of cryptographically linked blocks. It provides a distributed shared data store and a computational infrastructure. As a data structure, blockchain only allows data to be inserted. It does not allow any existing data on the blockchain network be updated or deleted. This is to prevent tampering and revision. Blockchain technology has only a very limited capability to support programmable transactions and only allows small pieces of data to be embedded in transactions for other purposes: for example, to represent physical or digital assets.

This technology currently provides a general purpose programmable infrastructure, and a ledger that stores the computational results generated from that infrastructure. Smart contracts (Alharby and Van Moorsel, 2017) are programs deployed and run on blockchain networks. These programs can execute triggers, rules, and business logic (Mohanta et al., 2018) to enable transactions. In such a decentralised environment, blockchains can also assure traceability of transactions, and thus achieve a high level of transparency and reliability in the ledger.

2.2 Blockchain Applications

Financial entities, governments, enterprises and startups are now exploring the applicability of blockchain in their domains. Application examples include, but are not limited to, digital currency, international payments, securities registration and financial sector settlements, registries, identity and taxation as government services, Internet of Things (IoT) storage, computing and management, and industrial supply chains (Zheng et al., 2018). Supply chains and health-care are considered a particularly promising area for the application of blockchain (Wang et al., 2019) (Bell et al., 2018).

In all these sectors, traceability is important for one reason or another (Dessureault, 2007):

- It makes it easier to prove compliance with regulations in audits, thereby reducing exposure to regulatory risk.
- It improves stock control in terms of reducing wastage, managing expired stock, safety issues,
and increasing profitability.

- It helps prevent counterfeit goods or fraudulent products from entering the market.
- It is beneficial for brands in that it helps build trust with customers, suppliers and other stakeholders.
- It improves efficiency in supply chains, international trade, chains of custody, healthcare chains, etc.

2.3 Blockchain Oracles

Blockchain networks, or more specifically smart contracts, are limited insofar that they cannot access the external data which might be needed to control the execution of business logic. Smart contracts are stored on the blockchain network and can only receive data from external services through so-called oracles (Beniiche, 2020). These can collect information from different sources. For example, they can monitor the status of a product to determine whether it has arrived and then write that status on the blockchain network. The change in product status could then be detected by the smart contract, which would trigger payment following the purchase of the product.

Some blockchain-based applications have recently become more complex, incorporating concepts like smart contracts, oracles and Decentralized Autonomous Organization (DAO). A DAO is a distributed application implemented to make it possible for multiple parties, either human or machine, to interact with each other (Buterin et al., 2014).

In practice, blockchain oracles are crucial to the correct execution of a smart contract, since the insertion of incorrect information may lead to actions that are not easy to revert (e.g., certain types of money transfer). The blockchain oracle workflow is typically executed between three types of participants: data feed providers, oracle nodes/network operators, and blockchain operators (Al-Breiki et al., 2020). Data feed providers enable different APIs (Application Programming Interfaces) to read and provide data from different data sources such as sensors, stock markets or even ad-hoc solutions to oracle nodes.

To illustrate this idea, Figure 1 shows an API server made up of different REST APIs (the most widely used web service technology).

2.4 Model-based Software Engineering

Model-Based Software Engineering (MBSE), also known as Model-Based Engineering (MBE), Model-Driven Engineering (MDE), Model-Driven Development (MDD), Model-Driven Software Development (MDSD) etc., is a software development paradigm focused on the application of visual modeling principles and best practices throughout the Software Development Life Cycle (SDLC) (Völtter et al., 2013). MBSE commonly uses multi-user repository-based modelling tools (Friedenthal et al., 2007) which provide an environment where a precise, unambiguous view of a system’s components, including their behavior and interactions, can be defined and managed.

The MBSE paradigm is model-based to the extent that the visual modeling artifacts it generates are sufficiently precise and complete to serve as a software or systems blueprint for improving SDLC efficiency and productivity. It is model-driven to the extent that it at least partially automates - that is to say, it drives - the SDLC via requirements that are precisely and completely specified as part of the system model, and which can be fully traced across the SDLC. MBSE is therefore the formalized application of modeling to support requirements gathering, analysis, design, verification and validation activities.

To illustrate this paradigm, Figure 2 shows the major MBSE activities.

Models can be either abstractions or representations of reality that facilitate the understanding of its complexity. The MBSE approach has four main activities. The primary activity is to specify software requirements. Unified Modelling Language (UML) is commonly used to do this (Schumacher, 2018). After specifying requirements, different model transformation techniques have been used to obtain the output model of choice for further verification and validation. The two most common transformation approaches are Model-to-Model (M2M) transformation and Model-to-Text (M2T) transformation (Zhu et al., 2019). The verification activity is carried out to evaluate the correctness of the model-system. If a model fails to satisfy the verification requirements, alterations are introduced to rectify the design errors, as shown in Figure 2. The model-system is validated through simulation. It is common practice to generate the source code using the model transformation technique, which is then used for simulation (Rashid et al., 2019).
2.5 Common Information Model

The Common Information Model (CIM) (Uslar et al., 2012) is an open standard that defines how managed elements, in an IT environment, are represented as a common set of objects and the relationships between them. This standard formalism was developed by the Distributed Management Task Force (DMTF), as part of its WBEM (Web-Based Enterprise Management) proposal (Arora et al., 2004) (Alexander et al., 2004).

The CIM standard includes the CIM Infrastructure Specification that defines the architecture and concepts of CIM, and CIM Schema, a conceptual schema which defines the specific set of objects and relationships between them that represent a common base for the managed elements in an IT environment. Some current proposals for formalizing structural diagrams through UML are easily adaptable to CIM diagrams (McMorran, 2007).

3 BLOCKCHAIN TRACEABILITY INFORMATION

Product or service traceability, for example in supply chains, is the connection of all business processes involved in the commercialization, generation and distribution of goods, from raw material to finished products and end consumers (Xu et al., 2019). Traceability enables consumers to track products during production and distribution (Grover et al., 2018).

To illustrate this concept, Figure 4 shows a supply chain traceability system for blockchain technology.

In brief, MBSE is a term that predicates the use of modelling to analyse and document key aspects of the SDLC. This paradigm is a model-centric approach providing a single pint of truth which is reflected in a set of living artifacts (Hart, 2015).

It is important to emphasize that technological progress requires increasingly flexible, sustainable architectures. In recent years, this has triggered greater interest in such new model-based paradigms. These alternatives have proved to be enormously robust, flexible and scalable compared to traditional strategies. Moreover, they feature concepts like reuse and platform independence (Rodrigues et al., 2012). MBSE designs for platform architecture can be reused to add more functionality and/or change the target architecture.

Figure 4: Supply chain traceability system for blockchain technology.
Governments also use traceability systems to display relevant information (e.g., origin, location, etc.) and issue certificates. In healthcare, traceability acts as a point of convergence for numerous needs (Lovis, 2008). In recent years, not only has traceability or tracking become increasingly important in the healthcare sector, but the application of blockchain technology is also being explored to improve the interoperability of patient health information between healthcare organisations while safeguarding data privacy and security (Eryilmaz et al., 2020).

Traceability systems used in these sectors typically store information in databases controlled by the same entity, but with such centralized storage there is a risk that data will be interfered with. In a collaborative environment like blockchain, the security of this type of system is important, especially in fields like accountability and the processing of forensic information.

Traceability is particularly interesting for those entities that want to adapt to new production models currently in development or pilot phases, and also for entities with quality certification, where it is an essential requirement to have a reliable, secure logging system, in which information remains unchanged. Once data is stored in the blockchain network, it can be considered secure and immutable (Sartori, 2020). Research conducted so far suggests that using blockchain technology is advantageous for achieving traceability (Aung and Chang, 2014). Although the same activities can be carried out using other techniques, blockchain technology adds additional functionality to data security (Li et al., 2020), something that entities find very appealing.

Data transparency is desirable because stakeholders want to check the authenticity of information. And it appears that traceability systems can benefit from the digital nature of blockchain without being affected by blockchain’s current limitations.

Basically, any workflow can be fully reflected in blockchain technology and can operate across multiple entities, similar as a collect data. Each step of the process can be registered and, depending on the intervals at which new blocks are added to the blockchain, the block creation timestamp shows the time and date when the information was recorded. In fact, one of the main advantages of using blockchain traceability instead of traditional solutions is that it is a mechanism that facilitates external and internal audits (Suzuki and Murai, 2017). Thanks to its immutable recording of information, the authenticity of the data can be irrefutably guaranteed for the different stakeholders. In other words, anyone can easily check what each stakeholder has registered and when. This also facilitates the detection of possible incidents (even automatically).

It appears, therefore, that an entity no longer needs to do anything else to guarantee the traceability or inviolability of data in the face of possible audits (it would have to do a great deal of monitoring if the information were stored only in a standard database). However, although its data remains immutable in time, the blockchain does not have a verification mechanism to prove whether the raw data were correct (Galvez et al., 2018), making it necessary to have an intermediate layer to facilitate such verification or validation mechanisms. If a piece of data is altered, the blockchain will not detect it. Also, as will be detailed below, it is essential for the traceability solution to be completely independent of the dataset recorded in the blockchain network. This way, depending on the user case, it would not be necessary to implement it again and it would be possible to apply the traceability process to different components at the same time. That is to say, a product or service could be monitored using a common dataset and the extensions needed for each business case.

4 DESIGN OF THE TRACEABILITY MODULE

As already mentioned, there are two aspects which need to be improved in blockchain traceability. Verification mechanisms need to be adopted to check whether the raw data is correct and the traceability solutions need to be completely independent of the dataset recorded in the blockchain network.

From our point of view, for a blockchain traceability solution to be efficient it must take into account the possible granularity of the information to be recorded, the cost of operations and, above all, the agility of the traceability monitoring. In view of these requirements, it is therefore necessary to have an additional layer in the blockchain, so that the best possible balance may be obtained between them.

One of the challenges involved is that for each application it is necessary to develop specific smart contracts that provide not only the traceability service, but also all other services associated with the blockchain platform. It is also essential for the new layer to have a flexible design, offering added value to stakeholders, and to be totally independent of the information to be recorded in the blockchain network. In other words, it must be a layer that needs no redeployment to be able to apply the tracking process to different components at the same time, taking advantage of the tracking properties intrinsic to any
To address these issues, the first objective was to design an architecture for blockchain traceability solutions with pre-established smart contracts, as shown in Figure 5.

The pillars of this puzzle are the blockchain oracle and the external APIs. Thanks to the blockchain oracle, smart contracts can be invoked in different ways, using specific parameters or information in a data interchange format such as the JSON format (REST API).

The second objective was to design a base meta-model which, thanks to extensions, is capable of accommodating new data models without the need to re-implement the platform.

Here, MBSE and CIM would allow independence from the destination platform and facilitate the establishment of a base model that would make the solution independent of the information to be processed, thus making this service transparently and reliably accessible to any client.

For the definition of the CIM model, it is therefore necessary to define a series of types, characteristics and basic operations which allow native types of information to be established (at the meta-model level, interrelations, classes, values, etc. are also indicated). In other words, this meta-model is a meta-structure that can be used to define the information model, establishing a series of general restrictions that allow the model to be structured and its scalability managed. In this case, the service layer could be defined at the meta-model level, and would therefore not need to be modified if the information model is an extension of the meta-model (since the service layer exchanges information at the instance level, regardless of the type of object involved).

It is important to note that the models must comply with a number of requirements. On the one hand, specific data types, properties and methods must correspond to the needs of the use cases and programming languages involved in the development. Once these have been defined, correspondence with those defined in the meta-model must be indicated. On the other hand, the extensions must be derived as classes of subclasses, preserving the predefined scheme, and must therefore always provide a detailed description of the elements added.

In brief, if the CIM model is used as a starting point (as illustrated in Figure 6), its extension will make it possible to cover one of the needs identified. For example, entity class "Address" represents the address information of the user or organization. In blockchain, an address is similar to an email address. It is used to receive and send data on the blockchain network (just as an email address is used to send and receive messages). If "Address" is an extension of entity class "Class" (this entity class is shown in figure 6), "Address" will have all the same features as the original, plus something more than "Class". That is to say, the CIM meta-model diagram can be easily extended without modifying the original metamodel.

Since the blockchain networks can support many use cases and are applicable in numerous industries, many companies are beginning to use Blockchain-as-a-Service. One example is the CIMT (Common Information Model for Traceability) Project. In this Project, developed by Soltel Group, a hybrid blockchain platform called "Blockchain platform as a Service for Traceability (BaaS-T)" is used to overcome all the limitations (see Figure 7).

The CIMT Project uses the Ethereum network of Alastria (an association that promotes the digi-
tal economy through the development of blockchain, and Oracize (Ethereum) to bring data from the outside world into an Ethereum smart contract. Soltel Group Project is defining tracking processes, stakeholders, and traceability data models in order to apply specific extensions to any business case. On the other hand, MBSE and M2T transformations, are being applied to define three generic smart contracts (services) that (i) will initialize any process of product or service traceability, (ii) register possible changes, both common data and possible extensions, and (iii) perform an advanced search to determine the evolution of a product or service in each of its stages and to be able to check if the raw data is correct.

5 CONCLUSIONS AND FUTURE WORK

Blockchain technology allows all stakeholders to check the entire history and (for example) the current location of a product. This technology also creates transparency for all stakeholders. In fact, by irreversibly storing data, it creates a unique level of credibility, and allows stakeholders to strengthen their relationships.

However, two aspects of blockchain traceability need to be improved. Verification mechanisms need to be adopted to check whether the raw data is correct and traceability solutions must be completely independent of the dataset recorded in the blockchain network. This paper describes a possible way to overcome these limitations.

Firstly, it proposes an additional layer or module of traceability that takes advantage of the characteristics of blockchain oracles and APIs. Secondly, it defines an extendable meta-model that can store traceability information independently of the information that is stored in the blockchain.

Planned future work includes implementing the proposed architecture and designing the meta-model that will facilitate the agile monitoring of product or service traceability in blockchain networks.

ACKNOWLEDGEMENTS

First of all, we would like to thank all the experts for their participation and for sharing their valuable knowledge. We would also like to thank all the participants in our pre-tests for their collaboration. This research was partially supported by the NICO Project (PID2019–105455GB–C31) of the Spanish Government’s Ministry of Science and Innovation and Trop@ Project (CEI-12-TIC021) of the Andalusian Regional Ministry of Economy, Knowledge, Business and University.

REFERENCES


