cryptography

MDPI

Article

Process Authentication through Blockchain: Three Case Studies

Mario Ciampi, Diego Romano and Giovanni Schmid *

Consiglio Nazionale delle Ricerche, Istituto di Calcolo e Reti ad Alte Prestazioni, 80131 Naples, Italy
* Correspondence: giovanni.schmid@icar.cnr.it; Tel.: +39-0816139529

Abstract: In this work, we elaborate on the concept of process authenticity, which intuitively corresponds to the validity of all process steps and their proper binding. It represents the most exciting forefront of distributed ledger technology research concerning the primary challenge of reliably connecting distributed ledger networks to the physical context it must operate. More in detail, the paper describes a novel methodological approach to ensure the authenticity of business processes through blockchain and several security mechanisms applied to the digital twins of the actual processes. We illustrate difficulties and opportunities deriving from implementing process authenticity in concrete case studies in which we were involved as software designers belonging to three critical application domains: document dematerialization, e-voting, and healthcare.

Keywords: distributed ledger technology; blockchain; process authenticity; tokens; anchors; oracles

Citation: Ciampi, M.; Romano, D.; Schmid, G. Process Authentication through Blockchain: Three Case Studies. *Cryptography* **2022**, *6*, 58. https://doi.org/10.3390/cryptography6040058

Academic Editor: Kentaroh Toyoda

Received: 5 October 2022
Accepted: 9 November 2022
Published: 11 November 2022

1. Introduction

A *process* or *procedure* is a series of (elementary) actions which are carried out in order to achieve a particular result, often with the aid of tools and devices. Well-defined procedures with reliable results are the foundation of structured societies and determine their production, development, and resilience capabilities. Over time, human societies have shown a strong tendency to organize themselves in an increasingly structured and complex way thanks to the ever greater control of the environment and natural phenomena resulting from the development of science and technology. Our post-industrial societies are characterized by depending on highly complex processes, which can involve many different actors and technologies, be composed of many phases, and be implemented on a large scale. The more articulated and complex a procedure, the higher the risk that its outcome may not be the desired one. Therefore, the highly complex processes that make up the structure of post-industrial societies are also their "Achilles' heel".

Many efforts of modern science and technology have focused on the development of methods and tools to ensure the accuracy and reliability of devices and processes in almost all areas of human activity, from industrial production to commerce, from healthcare to finance. With the development of information technology, many of the above operations and processes are now carried out with the aid of computers and other digital tools, giving rise to entire new disciplines, such as information security and cyber security, in order to limit the risks of violations and their consequent damages. Surprisingly enough, however, none of these disciplines explicitly address in a general setting the problem of *process authentication*, which consists of obtaining evidence that all the process steps were carried out correctly and in the right sequence. As a matter of fact, the Google Scholar searches for "process authentication" and "process authenticity" returned about 1500 and 800 publications, respectively, of which just a dozen were actually related to the issue considered in this work but concerned specific processes. Still, one of the areas of IT research and development that is experiencing great momentum is *distributed ledger technology* (DLT), particularly blockchain technology. At first, these technologies spread with the promise of eliminating financial intermediaries and enabling the democratization of money management through the distribution of trust among many parties (*decentralization*) [1]. However, if there is one

feature common to all the requests for the application of DLT/blockchain technology, this is not decentralization, but rather the support for the integrity and monitoring of critical processes [2,3]. Many applications of these technologies clearly show these aspects, as described in [4]. In order to facilitate the application of DTLs, many enabling platforms have been proposed and implemented so as to provide a wide choice to DLT developers [5,6].

The opportunities that DLT and, in particular, blockchain provide in different application domains are evident and discussed in many research studies. In [7], a survey on the application of blockchain technology from the point of view of applications, challenges, and opportunities is presented, highlighting the tradeoffs regarding different blockchain consensus mechanisms and application areas [7]. A methodology for the application of blockchain to enable new sustainable business models related to supply chain cost reduction is discussed in [8]. An extensive investigation on the adoption of blockchain for creating value in a critical application domain such as healthcare is illustrated in [9], where four different approaches are identified: endogenous hedonistic value, public–private conflict mitigation, utilitarian/instrumental value, and social value. Finally, the results of the systematic study illustrated in [10] highlight that blockchain technology—as an external force—creates an intersection among diverse research areas such as accounting, auditing, accountability, business, management, computer science, and engineering fields.

In the present work, we first elaborate on the concept of process authenticity and the standard cryptographic tools at our disposal to enforce that property in different use cases. We show that no such single tool exists for generic cyber-physical procedures, but that the bulk implementation of process authentication stems from the tamper-proof, append-only data structure representing the core of DLT. Then, through retrospective accounts of designs we delivered for real customers, we illustrate similarities, differences, and caveats in implementing process authentication thanks to blockchain technology. We consider three relevant case studies in the industry, public administration, and healthcare, respectively, showing how we decided to deploy a blockchain ledger and implemented some DLT-related concepts to achieve our goal.

The main contributions of our work are:

- Broad notions of *process* and *digital twin of a process* that encompass many different cyber, physical, and cyber-physical procedures;
- An explicit definition of *authenticity* comprising all the use cases included in the above notions;
- A framework to enforce and monitor the authenticity of a process, rooted in the ledger data structure and related concepts;
- Concrete examples illustrating the design patterns followed to enforce authenticity for some relevant processes in the industry, public administration, and healthcare, respectively.

The goal we intend to achieve with this contribution is twofold. On the one hand, we provide notions, tools, and a theoretical context of reference to create, evaluate and monitor the authenticity of sufficiently generic cyber-physical processes with the help of computer systems. On the other hand, we give concrete examples of the application of these concepts and tools to real-world significant cases, showing how each of these cases underlies specific challenges, requiring ad-hoc solutions by following a careful analysis of the application domain.

The remaining part of this paper is organized as follows. Section 2 refers to scientific literature that we have found to some extent relevant to our approach or to the case studies presented below. Section 3 introduces our framework for process authenticity, showing that some DLT concepts and tools are critical to enforcing it, but also that they could be very difficult to implement, depending on the application context. Sections 4–6 concern the case studies presented in this work, relating, respectively, to document dematerialization, electronic voting, and the management of health processes. These sections briefly illustrate the application domain requirements and our design choices to meet them. We summarize our findings and concluding remarks in Section 7.

2. Related Works

Information security has introduced authentication procedures and mechanisms from the very beginning, recognizing cryptographic tools as the most suitable preventive ones. Indeed, many developments in modern cryptography concern concepts relating to different types of authentication and the tools for enforcing them, from message authentication codes (MAC) and digital signatures to protocols such as TLS.

The scientific literature is plenty of contributions on basic authentication tools and procedures for enforcing integrity protection in more complex use cases. Although these use cases often underlie actual processes as intended in this work, the notion of process authentication does not seem to have found an explicit and clear contextualization. Additionally, due to the scarcity of successful DLT projects outside the financial sector and the lack of step-by-step approaches to implementing this technology, there is no clear understanding of how and when to implement DLT concepts and the actual benefits of such implementations to preserve the integrity of a process.

Supply networks are undoubtedly the application field where researchers and technologists have more often strived to implement some form of process authentication. Besides the financial sector, it is also the industry with the most extensive body of work in blockchain technology. Document dematerialization, the first case study presented in this work, is one of the many DLT-based pilot projects in the context of SMART (Sustainable, Modern, Adaptive, Robust, and Technology-oriented) supply networks [11], a very active technological sector concerning which there have been some studies and scientific publications complementary to our work. The authors of [12] report a two-year design science research study of a smart contract initiative piloted by a consortium in the UK's construction sector. They explore how a group of supply chain actors collectively designs and pilots a blockchain solution that addresses the supply chain transparency and provenance problem, developing a set of design principles that can be applied and tested in different supply chain contexts. The paper [13] analyzes blockchain adoption drivers and barriers, applications, and implementation stages within food supply chains. The results are then used to develop a three-stage conceptual framework that can help managers establish the suitability of the technology for their organization. The study [14] highlights that blockchain technology can improve audit trail, transparency, and traceability in the context of *additive manufacturing* (AM), characterized by digital assets such as CAD files, which should be shared and processed respecting the rules of intellectual property. A major contribution of this paper is the identification and prioritization of both technical and non-technical barriers to blockchain adoption in AM supply chains, based on a ranking-type Delphi study of responses by experts in the field.

Regarding the blockchain e-voting case study, several publications help us understand the general setting of e-voting. The authors of [15] introduce a schematic of key concepts with a review of electronic voting systems to date. A practical example with reasoning on election organization, cryptographic tools, and voting schemes is available in [16]. Focusing on remote e-voting (also called online e-voting), an introductory critique is available in [17]. When going specifically into the setting of blockchain e-voting, the authors of [18] give a schematic overview of several initiatives present in the literature highlighting the pros and cons. Finally, in [19], the authors give a severe framework to identify risks in the adoption of blockchain e-voting systems, taking the voters' side and providing a set of critical questions to evaluate any new voting system proposal.

In the healthcare sector, numerous research proposals and prototype implementations have been made on the application of blockchain technology. In [20], the authors propose a decentralized system based on Ethereum for document management in Electronic Health Record (EHR) systems, capable of both i) ensuring access to medical data by allowing patients to manage their own data and ii) providing mechanisms to share medical data for research purposes. In [21], a blockchain system is proposed to manage access control to health data represented according to the standard HL7 FHIR format and guarantee data integrity. In [22], a blockchain system is integrated with a mobile health system

capable of managing data from wearable and non-wearable medical devices to ensure access control, privacy and data integrity. In [23], blockchain technology is adopted to provide secure storage of EHRs through granular access rules. In [24], a model based on an authorized blockchain is proposed for the management and storing of EHRs of registered patients, guaranteeing transparency and immutability. In [25], the authors propose a blockchain-based system capable of achieving confidentiality, authentication, and integrity of medical data and supporting fine-grained access control, using attribute-based encryption and identity-based encryption to encrypt medical data and identity-based signature to implement digital signatures.

3. A Framework for Process Authentication

It is interesting to note that very different processes in terms of application context and processing share a common core structure. In fact, in their essence, they do not differ from an assembly line or a supply chain: we can represent each one as a temporally ordered set of steps (processing phases) carried out on one or more classes of objects. Each phase is achieved thanks to the participation of one or more agents authorized to perform only specific actions. Furthermore, each phase has preconditions, a success final condition, and an error final condition. Error conditions give rise to branch points (exceptions) that can result in repeating some processing, performing alternate processing, or stopping processing for the objects involved. In any case, some agents may be in charge of checking the final condition and deciding whether it was a success or a failure.

The items under processing can be digital assets, tangible and intangible objects, or both. On the other hand, agents can be automatic machines or humans with or without the support of special devices. Therefore, this framework encompasses a wide variety of cyber, physical, and cyber-physical procedures. The reader can easily verify that the three case studies presented in this article, although related to very different application domains and problems, fall within the aforementioned abstract scheme, sketched in Figure 1.

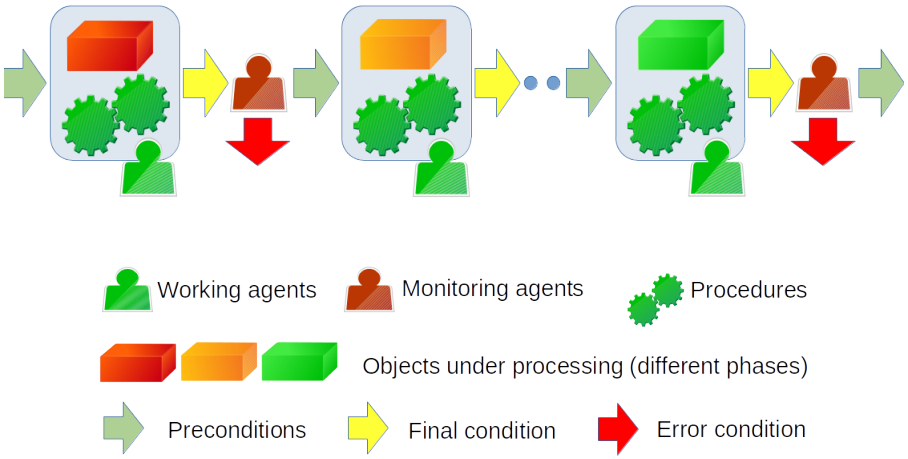

Figure 1. Schematic of a generic process. Each step represents a processing phase carried on by some physical or digital objects thanks to a group of working agents that can be humans or devices. Each phase is associated with a specific input (preconditions) and output (end conditions). The validity of end conditions can be checked by monitoring agents.

In e-voting procedures, which are discussed in Section 5, the unique kind of items under processing are the intangible votes of citizens, whereas medical workflows (see Section 6) can encompass a plethora of physical documents and digital formats that require a data model to identify the types of data produced and consumed in each activity. In the case of document dematerialization (see Section 4), the objects of the processing are of two types: archival units and their corresponding digital archives, which consist of copies

obtained by optical scanning of the first ones and the related files and metadata useful for archival purposes. For a generic production chain, they are the object to be produced P and those from which it is obtained. For example, we can think of assembling a car from all its components or producing food from its ingredients. These last two examples are only apparently more complicated than the one relating to dematerialization. We can assume that each of the constituent components of our product P is the final successful outcome of an independent production chain, which does not fall under the responsibility of the company that produces P, whose only duty will be to ascertain that the quality of P's components (as corroborated by third-parties) is capable of producing P. Something similar also occurs in the dematerialization process: for example, during the physical treatment phase of the paper archives, suitable chemical preparations from third-party companies must be used.

The previous property can be observed in many other different contexts than production chains, including those concerning public consultations and medical workflows. It establishes that different processes can be composed in a more complex one or, conversely, that a process can be decomposed into simpler, independent sub-processes. The *composability* of processes is often exploited by industry and service providers because it allows splitting work charges, checks, and responsibilities into several phases or between several companies. In some cases, composability allows splitting a process with branches and loops in an ordered set of pipelined or parallel linear processes such as that in Figure 1.

3.1. The Digital Twin of a Process

The differences among the process descriptions resulting from the previous schematization become even more nuanced if we consider the digital representation of these processes. Introduced for manufacturing [26], the term *digital twin* denotes a set of concepts and methodologies for creating digital models of real-world objects and processes that serve as their virtual counterparts or analogs [27,28]. Through a comprehensive physical and functional description of a component, product, or system, its digital twin includes all information that could be useful in all the current and subsequent life-cycle phases [29,30].

Abstracting from manufacturing, the *digital twin of a process* can be regarded as an informational structure describing its key parameters and outcomes. In the following, we will refer to the digital twin obtained during the execution of a process, denoted in [28] as β digital twin to distinguish it from the α digital twin introduced for describing and simulating the process before its implementation. The distinguishing features of the digital twin as intended thereafter are that: (i) it contains the actual parameter values, outcomes, and other associated resources determined during process execution; and (ii) after process completion, it is completely defined and no further changes may be introduced.

But what does it actually mean to obtain a digital twin $\tilde{P}$ of a given process P? What real benefits does this entail?

It should be clear that the way we can obtain $\tilde{P}$ and the accuracy with which it represents P strictly depend on P's nature. Some processes are natively digital, or it is very easy for them to obtain a digital twin. In the editing of a file through a computer application, for example, the person acting as a working (and monitoring) agent is represented by an operating system user, the object under processing is digital data, and the performed actions correspond to program instructions executed by the computer. Several multi-party processes, in different contexts, also share this feature. A notable example is that of games: card games, chess, and roulette are all processes for which it would not be difficult to design a digital twin, albeit its implementation could be cumbersome. In fact, there have long been online versions of such games on the Internet. Many other processes, however, are much more difficult to represent digitally. Whether it is the production of a type of vegetable, the assembly of the components of an appliance, or the care of a diabetic patient, it will be necessary to digitally code in some way material objects and physical procedures affecting them. Clearly, this can be extremely difficult. Humans and devices carrying out or monitoring the process can make use of computer programs, where they can easily be

associated with numerical identifiers. On the other hand, modern cryptography has long provided algorithms and protocols to ensure the proper correspondence of such identifiers to the subjects they represent and the fact that a particular subject is present in a specific circumstance or time frame.

But what about the growth of a plant as a function of its cultivar and the health of a patient following a drug treatment? The critical point is to understand if and how the object being processed can be characterized as a set of alphanumeric parameters and which variations of these parameters occur due to correct and wrong procedures, respectively. In this way, it is possible to express each process stage's preconditions and final conditions in terms of data and computer programs. However, it will be necessary to have devices capable of measuring the parameters above, correlating them to the various stages of the process.

The approach to obtain a digital twin $\tilde{P}$ for many processes in different application areas is to have one or more data structures capable of representing the inputs and outputs managed by the process P and, through appropriately programmed logic, describe P as a series of state changes for those data structures. As should be clear from the previous discussion, the data structures and programs that make up the digital twin $\tilde{P}$ of a process P are strictly dependent on the process under consideration, as well as the adherence of $\tilde{P}$ to P. As a general rule, $\tilde{P}$ must give a digital representation of the objects under processing, precisely reflecting their state changes in each processing phase until the final process outcome. The reader can verify that we rely on minor variations of this approach for all the case studies described in this work.

The reasons that induce industry and other sectors to consider the digital twin of a process P derive above all from the need to monitor P at low cost and efficiently so as to correct some anomalies and guarantee over time and on a large scale that the outcomes from P meets specific criteria. Furthermore, in some cases, there is the need to protect brands and consumers from *counterfeiting*, which is pejorative and undeclared changes to a production process for profit. Although a process could be tracked, monitored, and certified without resorting to its digital twin, this usually makes it possible to simplify controls and reduce related costs.

3.2. Enforcing Process Authenticity

The usefulness of using a digital twin $\tilde{P}$ for a process P is even greater when the company or organization running P wants to ascertain the absence of accidental or voluntary alterations in P, proving this circumstance to others if required by law o suggested by the market. Indeed, as shown in Figure 2, each processing phase of P results in a $\tilde{P}$ step, but now these last are composed just of data and computer programs, for which cryptography provides very effective tools to detect their forgery. Precisely, each $\tilde{P}$ step is made up of programs that check preconditions and generate conditions with the support or under the control of digital identities. Cryptography can offer protection for all these kinds of data, regardless of their function and format, also assuring their correct provenance (authenticity).

In this regard, we should note that the correct outcome of P is a necessary but insufficient condition for its authenticity. Even if the result appears correct, some phases of P may have taken place differently than expected, with consequences not present on the primary outcome but in a derivative product, or not easily detectable following routine checks. Furthermore, even if each phase of P is performed correctly, their different order could produce unexpected effects or results. For the proper enforcement of P's authentication, it is therefore imperative that its digital twin $\tilde{P}$—in terms of both data and programs—represents all P phases and their relative succession in a unique and unforgeable way, respecting the overall ordered series of preconditions and final conditions.

It should be clear at this point that the most suitable data structure to encode the evolution of an unforgeable process corresponds to the ledger concept implemented in various flavors in DLT. In particular, a digital twin such as the one depicted in Figure 2 can

be effectively represented using a time-oriented, tamper-proof linked list of digital records; that is, the quintessence of the *blockchain* data structure.

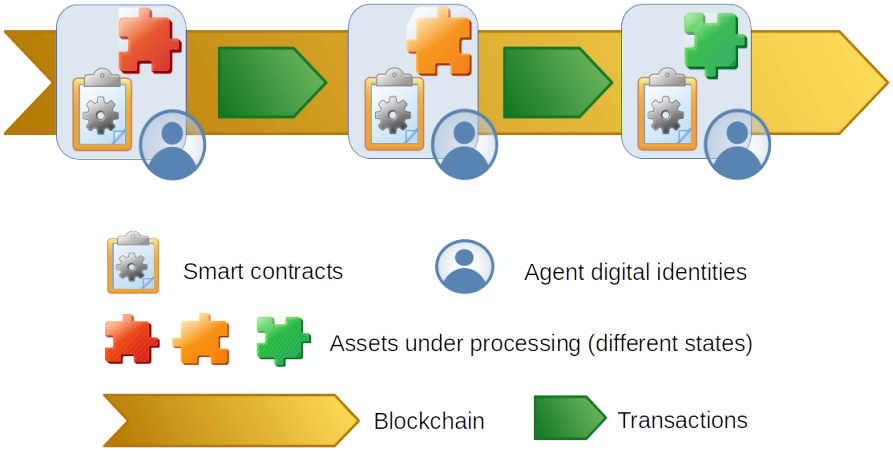

Figure 2. Schematic of the digital twin of a process. Thanks to DLT concepts and tools, each step encodes a corresponding processing phase in the original process. The underlying blockchain ledger guarantees the unforgeability of data and programs.

Some other DLT concepts and tools turn out to be critical for an adequate representation of process authenticity. First and foremost, the actions performed during each single processing phase of P must be represented in $\tilde{P}$ through unforgeable programmed logic enforcing a success end condition only if some processing on valid preconditions occurs. This is precisely why *smart contracts* were introduced in [31].

Depending on the DL platform in use, the notion of *asset* or *token* [32] is another major concept useful in implementing process authentication. In Ethereum and other systems, tokens were introduced to encode the lifecycle of both digital assets and tangible objects. Similarly, Hyperledger Fabric assets allow us to virtually encode any tangible and intangible real-world asset. Last but not least, we can grasp the actions performed by the different agents in the different phases of a process through the notion of *transaction*, since it represents a change of state for an asset or token. This way, a transaction chain in the ledger encodes the cause-and-effect correlation among different actions performed on an object under processing.

All the above notions can be implemented using conventional cryptographic techniques that are mainstream in developing DL systems. However, the digital twin of a process also requires the correct binding of the digital identities representing agents and objects involved in the process with their corresponding entities in the physical realm. We can accomplish that relatively easily for people and devices that can use computer programs, thanks to digital signatures and infrastructures for managing the trust, as described in the following section. However, for the case of a peeled tomato, we need to resort to the more advanced and contrived approaches sketched in Section 3.4.

3.3. Identity Management and Access Control

Processes such as those we have previously introduced require that heterogeneous agents (human beings and computer procedures) and devices (sensors, actuators, computers, etc.) operate correctly and coordinated to obtain the desired results. Precisely, the success of a process requires that each agent or device acts following a precise timeline and carries out only specific actions that depend on certain initial conditions.

It follows that the creation of mechanisms for the enforcement and verification of the correctness of a process requires, in the first place, management of the identities of the

subjects involved in the process and their roles, in order to be able to implement reliable access control.

Typically, the deployment and execution of a process correspond to the delivery of a service by a single company or organization, which could deploy parts of the process in outsourcing or rely on some preprocessing and its related artifacts as provided by other companies. In this case, all the functional divisions responsible for carrying out the various process phases belong to a single administration, simplifying the management of the digital identities of the actors involved and, more generally, the security of the entire process. Nevertheless, the implementation and deployment of the digital identity management component must be well thought out. In purely centralized trust management, the IT system and the database represent single points of failure. On the other hand, it could be challenging to manage many actors and resources without adequate coordination and monitoring among teams and IT systems relating to the various company divisions. In the provision of the service that involves subcontracting, the supplier can contact external companies while being the only one responsible, even in terms of law, for the service offered.

A less common case of a process deployment is through a company consortium, where some companies agree to pursue a common product or service but with a clear division of tasks and responsibilities. For example, a consortium aimed at creating a large-scale document dematerialization service adhering to the highest standards, as described in Section 4, could include a leading logistics company, a company specializing in standardized digitization processes, and a third company specializing in archival management.

All such methods of providing the service underlie a modular and hierarchical structure for trust transmission and management that can be implemented through a public key infrastructure (PKI), such as that supported through the X.509 standard [33]. An X.509 certificate is a digital signature binding an identity (hostname, organization, or individual) to its public key, either signed by a certificate authority (CA) or self-signed. A party A holding a valid X.509 certificate of another party B can use the public key it contains to establish secure communications with B, or validate documents digitally signed by B. X.509 also defines certificate revocation lists (CRLs) and a certification path validation algorithm. CRLs distribute information about certificates that have been deemed invalid by a signing authority (e.g., because the private key was lost or disclosed), whereas the certification path validation algorithm allows for certification chains: certificates can be issued by intermediate CAs which are, in turn, certified by higher-level CAs, eventually reaching a trust anchor.

In centralized service delivery, the apex of trust is the CA for a single company providing the service and acting as a (single) source of authority (SOA). If the service or process is delivered through a consortium, identity management can be provided by a set of cross-certificate CAs corresponding to the companies that make up the consortium. Starting from the above trust anchors, the PKI of the company or consortium can articulate into intermediate CAs to reflect the operational divisions of an organization, each with its units of personnel and devices. Thanks to external CAs and intermediate CAs, a company CA is aware of all the public key certificates for human agents and devices participating in the process. This way, the system can track internal (company personnel and tools) and external actors (customer representatives, public officials, business partners) in various capacities in executing or monitoring the different phases of the process. By exploiting the properties of public key certificates and digital signatures, the supplier and its customers can have evidence of who participated in the process and their actions as a starting point for an exhaustive and reliable tracking of what possibly went wrong.

Within this framework, trust decentralization can be enforced by implementing consensus policies (e.g., the *endorsement policy* in Hyperledger Fabric [34]) at the consortium level and individual company level. We conclude this section by emphasizing that the aforementioned approaches to trust management concern the application (business) layer, which is independent of the management of the distributed ledger. This means that we can implement such approaches on both *permissioned* DL systems (e.g., Hyperledger Fabric)

and *permissionless* systems such as Ethereum. For a more comprehensive discussion on this topic, the reader can refer to [32].
Related work and references to the main concepts discussed in the present work are summarized in Table 1.

Table 1. Scientific papers and technical reports related to the topics and concepts discussed in the present work.

Authors and Year	Topics Related to the Present Work
Kawa and Maryniak (2019) [11]	DLT-based pilot projects in the context of SMART supply networks and lean manufacturing
Wang et al. (2021) [12]	A design science research study of a smart contract initiative piloted by a consortium in the UK's construction sector
Vu et al. (2021) [13]	An analysis of drivers and barriers, applications, and implementation stages of blockchain within food supply chains
Kurpjuweit et al. (2021) [14]	Technical and non-technical barriers to blockchain adoption in additive manufacturing supply chains
Wang et al. (2017) [15]	A schematic of key concepts with a review of electronic voting systems at date
Baudron et al. (2001) [16]	A practical example on election organization and voting schemes with related cryptographic tools
Gibson et al. (2016) [17]	An introductory critic on remote e-voting
Abuidris et al. (2019) [18]	Overview of several blockchain e-voting initiatives present in the literature
Park et al. (2021) [19]	A framework to identify risks in the adoption of blockchain e-voting systems
Azaria et al. (2016) [20]	A decentralized system based on Ethereum for Electronic Health Record management
Zhang et al. (2018) [21]	A blockchain system to manage access control to health data in the standard HL7 FHIR format
Liang et al. (2017) [22]	Integration of blockchain with a mobile health system capable of managing data from wearable and non-wearable medical devices
Shahnaz et al. (2019) [23]	Blockchain technology for secure storage of EHRs through granular access rules
Capece and Lorenzi (2020) [24]	A blockchain-based model for the management and storing of EHRs of registered patients
Wang and Song (2018) [25]	A blockchain-based system that uses attribute-based and identity-based encryption for the confidentiality, authentication, and integrity of medical data
Grieves (2014) [26]	Introduction of the "digital twin" concept for process monitoring in the molding industry
Cheng et al. (2018) [27]	Digital twins for cyber-physical integration and smart manufacturing
Kholopov et al. (2019) [28]	The concepts of α and β digital twins in industrial automation
Boschert and Rosen (2016) [29]	The "digital twin" concept as a comprehensive physical and functional description of a component, product, or system
Mandolla et al. (2019) [30]	A digital twin for additive manufacturing in the aircraft industry through the exploitation of blockchain
Szabo (1997) [31]	The introduction of the concept of smart contract
Romano and Schmid (2021) [32]	An overview of the most relevant and impacting DLT concepts. An explicit notion of "process authenticity" although informal

3.4. Binding the Ledger to the Outside World

The trickiest point in the context of process authentication is to correctly bind the digital entities managed through the digital twin to their counterparts in the physical world. Standard authentication methods (e.g., digital signatures and access control) can accomplish the above binding for people and, often, for processing devices, but the physical resources subject to processing usually require more advanced mechanisms. More generally, for the digital twin $\tilde{P}$ of a process P we have to face the *garbage-in, garbage-out* issue: if $\tilde{P}$

is fed with data other than those that characterize the process P, then $\tilde{P}$ will not be able to represent P accurately, much less P authenticity. On the other hand, it can be relatively easy for an attacker to alter some of the boundary conditions that represent the frontier between the physical and digital world in order to load inauthentic data into the ledger, defeating P modeling via $\tilde{P}$. *Anchors* and *oracles* are two recent concepts developed in the DLT field whose goal is to provide tools and approaches to prevent the problem above. Oracles are third-party services used to validate data generated by unreliable sources and pass them as input with authenticity proof to smart contracts, whereas anchors tie physical object identifiers to object properties that are hard to clone, forge or transfer to other objects. Anchors must enforce an advanced notion of authenticity, which is described formally through the *authentic data encoding for physical resources* scheme introduced in [32]. We are going to describe some of the challenges of implementing anchors in the next section.

4. Case Study: Document Dematerialization

For about a year, we participated in a project concerning the dematerialization of paper documents in compliance with the updated provisions on the subject issued by *AGID-Agenzia per l'Italia Digitale* (https://www.agid.gov.it/en, accessed on 3 October 2022). The project goal was to corroborate evidence of the authenticity of the dematerialization process thanks to an information system that monitors each phase of the process and uses techniques to ascertain the conformity of digital documents with their paper originals. Some of these techniques fall under DLT, and we illustrate their benefits and limitations after carefully analyzing the design requirements for this case.

4.1. The Document Dematerialization Business

Many companies and public administrations still use paper documents to store and share information with employees and users. In most cases, they implement only partial digital workflows, with output represented by paper documents that must be converted to digital format to be processed by subsequent systems. However, this turns out to be detrimental to productivity, quality of service, data security, and costs. The management of information through paper documents results in long times for searching and consulting the records. The French *Association Information et Management* (https://aim.asso.fr/, accessed on 5 September 2022) has estimated that approximately 7.5 h per week are spent searching for information without finding it. Moreover, high costs are required for processing and archiving paper documents. According to a study conducted by IDC (https://www.idc.com/, accessed on 5 September 2022) in April 2020, printing consumables (reams of paper, ink cartridges), operation and maintenance of printing peripherals, and the item related to the enveloping of documents and postage represent on average between 1% and 3% of a company's turnover. In addition to locations for machines and paper stocks, the storage of sensitive paper documents can be very cumbersome and expensive since it requires one or more dedicated deposits with monitored climate conditions and surveillance personnel.

Document dematerialization is the replacement of physical (usually paper) documents with digital files in a company. By centralizing all the documents in a digital collaborative space, dematerialization considerably simplifies the search and consultation of the documents. It also allows a better follow-up and traceability of documents, limiting the risk of errors thanks to automation in the data entry. Switching from physical deposits to storage information systems is a great source of savings in both costs and space. Last but not least, dematerialization improves global productivity by saving time and increasing the responsiveness of many processes, especially in a context where remote work and mobility are on the rise. The IDC study shows that reducing dependency on paper documents was in the top five priorities for IT investments of companies surveyed in the supply chain sector, with 29% of them designating dematerialization as a priority.

Going from paper to digital information carriers implies routing documents through several phases: their digitization, their indexing to be able to find them easily, even their versioning to keep track of all the successive versions of the modified documents, their

archiving, and their certification to attest their veracity. Digital solutions facilitate these steps and allow them to be processed automatically. Technical expertise, compliance with legislation, and technological and digital know-how are the keys to the digitalization of document processes with cutting-edge solutions. Typically, companies and public administrations do not have this expertise in-house and have to obtain their dematerialization in outsourcing. This is particularly true when the dematerialization concerns administrative, legal, or sensitive documents as per regulations in many countries. In the European Community, for example, the rules for processing sensitive data have been defined in the General Data Protection Regulation (GDPR) [35], and, based on them, the Italian *AGID-Agenzia per l'Italia Digitale* has updated the Digital Administration Code [36]. According to current legislation, only a notary or public official can guarantee the evidential effectiveness of a document copy by issuing and signing a *certificate of conformity*. Thus, a dematerialization process has to include the above step to return digital copies with legal value.

How a supplier provides a dematerialization service depends on marketing choices and technical-organizational needs. A single company often offers the service, but sometimes it uses subcontracting or belongs to a consortium. In any case, we fall in the framework of trust management described in Section 3.3.

4.2. The Document Dematerialization Process

The goal of a dematerialization process is to produce digital twins of paper documents with the same legal value as the latter but with the advantages—in terms of costs, performance, and scale—deriving from using computerized techniques to classify, store and retrieve information. Such a process requires the intervention of different heterogeneous entities that must operate and interact synergistically so that the process complies with the highest standards (e.g., [37,38]) and stringent regulations in force for the management of sensitive data (e.g., [39]). In particular, all the participating entities—whether human beings, IoT devices responsible for process control, or the same resources subject to the process (paper documents and related digital copies)—must be suitably identified and traced throughout the process. Specifically, four main categories of entities intervene during a treatment:

- **Processing agents**, that can be operators of the company providing the service, customer operators, and public officials.
- **Processed objects**, which are sets of related paper documents constituting single *Archival Units* (AU), where in turn each document composes of one or more *Digitizable Elementary Units* (DEU), typically consisting of a single sheet of paper each.
- **Digital artifacts** produced as a result of the AU processing (e.g., tracking data, digital copies obtained by scanning from paper originals, metadata for archival management).
- **Physical devices** for carrying out or controlling the treatment (e.g., scanners, smartphones, video cameras, containers).

All these entities participate and contribute in various ways to the realization of document dematerialization and its monitoring. This rather complex process composes the following phases and their related functional modules:

- **Process setup** An operator of the Sales and Deliveries division performs a series of actions to set up a document dematerialization process for a given customer. Based on a service contract stipulated with a customer, the operator produces a series of conditions and electronic documents to manage and monitor the overall process.
- **Taking charge at the customer premise** An operator of the Logistics division is at one of the customer's premises to collect the AUs that must be digitized. The operator must insert the AUs in suitable containers for their transport to the digitization facility while producing a *delivery report* by recording data concerning the AUs and their transportation into an information system. The information system was configured in the previous step, and it implements programming logic and data reflecting the requirements and constraints of customers according to their stipulated service contracts.

- **Document transportation** Operators of the Logistics division use vehicles equipped with a GPS tracking system to transfer AUs from the customer premise to the digitization facility, optionally returning the documents back to the customer after their digitization.
- **Taking charge at the digitization facility** Operators of the logistics division check that the AUs received (and their related DEUs) correspond to the delivery report drawn up in the previous step and that they have not suffered damage during transport.
- **Physical pre-treatment** Operators of the Archiving division inspect and classify the AUs received from the previous step. They possibly carry out physical treatments (sanitization, restoration, etc.) on some DEUs to improve their subsequent processing. They keep track of document existence and status using the information inserted by the operators of the Logistics division at the customer's premises and the digitization facility.
- **Digitization** Three operators and a supervisor participate in this module, which implements the process of digitizing documents. An operator takes care of the acquisition of digital copies thanks to scanner devices equipped with advanced functions. He is assisted in this task by a second operator, who is responsible for verifying the conformity of the copies obtained based on the type of documents and contractual conditions. Finally, a third operator is responsible for inserting all the metadata necessary for archival purposes in the files obtained following the scan.
- **Certification** A public official (e.g., a notary) issues a *certificate of conformity* which certifies the correspondence of the digital copies resulting from the previous step to their original paper documents.
- **Settings and data feed management** A system administrator sets up the various computer devices and applications required by the process (Databases, IoT devices, Identity and Access Management systems). Each supervisor manages parameters and inputs concerning the sub-process implemented by their division (e.g., the Optical division supervisor manages acquisition and compression parameters).
- **Incident management** A service manager collects and manages incidents that can occur during the process.

Given that digitizing a paper document turns out to trivially be a digital document, it is legitimate to ask whether we can track and monitor the aforementioned process thanks to a conventional approach.

The traditional solution to guarantee the authenticity and non-repudiation of a digital document consists of applying a digital signature algorithm, a solution which, in the case of a conventional ledger (i.e., non-blockchain type), could be used to satisfy the requirements for the *Taking charge* and *Digitization* phases by generating and verifying digital signatures on three separate documents. For example, the delivery report produced during the *Taking charge* at the customer premise must contain the list of AUs (and their related DEUs) taken in charge according to the contractual conditions. A digital signature on the report by both the customer and the operator can suffice to validate this step. The relative verification procedure should be carried out at the time of taking charge at the digitization facility and it would consist in verifying the validity of the signatures affixed and the consistency of the AUs and DEUs listed in the report with those physically received by the logistics division. Only documents in both lists should be considered for the next stage of the digitization process, while for the others, exceptions should be generated and separate management provided.

However, this approach does not explicitly link the process steps together and does not allow for direct handling of processing exceptions. By separately executing signing-verifying procedures for each phase of the process, we lack a unitary and comprehensive view of the dematerialization pipeline.

On the contrary, the "linked nature" of the transactions implemented through a distributed ledger allows for the digital representation of a whole process as an ordered set of cause–effect relationships. An in-depth analysis of the above process, in terms of

its use cases and their mutual dependencies, showed that it can be modeled as a pipeline of steps with few branch points so that a ledger with a linear chain structure, i.e., a time-oriented ordered list of append-only records is perfectly adequate to track and trace its evolving state.

As detailed in the following section, we exploited some core concepts of blockchain technology in order to obtain a faithful representation of the document dematerialization process thanks to a blockchain-type data structure and a series of smart contracts. By design, these latter describe all the processing steps necessary to obtain a digital twin of a paper document, equivalent to it in both legal and administrative terms. In this way, it is possible to automate a whole series of control procedures and obtain a quality guarantee for the entire process.

4.3. The Blockchain Model

For document dematerialization, it is appropriate to choose an *asset* capable of encoding, on the one hand, the paper documents that compose the original archival units, and, on the other, their corresponding digital copies and related metadata.

By designing *smart contracts* for processing the asset through suitable *transactions*, we are able to describe the dematerialization process as a series of different steps involving state changes in the AUs, considered as "elementary units of work". Indeed, transactions allow keeping track of the actions performed by the different agents in the different phases of the dematerialization process for a given AU. Thus, we can encode the cause-and-effect correlation among different actions on AUs as transaction chains in the ledger.

Finally, we can use the notion of *block* to associate multiple AUs with the package used to transport them from the customer site to the processing center. Block creation and verification correspond to assembling a package at the customer site and opening it at the processing center, respectively. This workflow is substantially different from that implemented in blockchain platforms currently on the market, where the transactions submitted by different agents get assembled in blocks only for efficiency reasons.

Document dematerialization is clearly a cyber-physical process since it involves the processing of paper documents through physical devices. Therefore, its proper digital encoding requires the adoption of some of the strategies and tools discussed in Section 3.4. The most compelling case concerns AUs and their packages since they require a technology that is easy to deploy and at a low cost. The most straightforward solutions consist of tapes and labels printed on special supports to provide visual evidence of their possible tampering. These unequivocally indicate the possible first opening thanks to unique graphics that can be checked both with the naked eye and through UV rays. The labels contain a QR code able to identify them uniquely, and the appropriate application of tapes and labels to the AUs or their packages can protect them against tampering or theft.

If replacement or cloning are plausible threats, through special techniques and the support of a remote server, we can obtain more sophisticated tags resulting in *anchors* [32].

If customers monitor AUs packaging at their sites, then an *anchor* implementation for the package IDs and tamper-proof packaging could suffice to ensure a one-to-one unforgeable association between physical AUs and their digital counterparts. We can use the hash digest obtained as the Merkle tree root of the transactions composing a block as the ID of the package. Alternatively, in a more severe threat model, we can assure a one-to-one unforgeable association between each paper AU and its encoding assets by anchoring it through the asset ID.

5. Case Study: e-Voting for Public Consultations

In many discussions about e-voting, the tacit subject is remote e-voting, which is a relevant and enabling idea when mobility is difficult, or gatherings are discouraged. Several systems exist [17] to express preferences via web or mobile app, and they are effectively adopted to consult opinions in limited assemblies or specific communities. Consequently, authors usually pay little attention to corruption or coercive pressures when narrowing

the idea of e-voting in such limited contexts. If this is the case, a proposal of blockchain-based e-voting systems [18] overstates actual risks while introducing an immature layer in the structure.

On the other hand, when thinking about public consultations involving a multitude of people on sensible themes (Referenda) or assigning democratically chosen governing seats, political forces come into play, and we need to consider both specific legal cases and technological issues. In [40], an EU Cooperation Group created a taxonomy to classify the origins and consequences of large-scale cyber-attacks on technologies related to elections. A list of possible threats regarding the solely digital voting procedure includes:

- Tampering or DoS of voting and/or vote confidentiality during or after the elections;
- Software bug altering election results;
- Tampering with logs/journals;
- Breach of voter privacy during the casting of votes;
- Tampering, DoS, or overload of the systems used for counting or aggregating results;
- Tampering or DoS of communication links used to transfer (interim) results;
- Tampering with the supply chain involved in the movement or transfer of data.

The above risks directly depend on the inherent properties of a large-scale networked e-voting system, regardless of the assumed technological layers. May the blockchain mitigate those risks? What are the legal points to consider in the e-voting process?

In order to investigate these questions, in 2018, we decided to answer a call of the municipality of Naples (Italy) addressed to associations, universities, research centers, and students to form a volunteer working group on the usage of blockchain technology in municipal referenda. The underlying idea aimed to directly involve the citizens in the democratic decisional process on town-related issues, limiting the costs of a traditional voting consultation.

Since the first day meeting, thanks to suggestions from officers embroiled in past voting sessions, the group focused on possible strategies to limit corruption, vote-buying, and malicious manipulation of ballots. Those generic issues, in the first instance, convinced us to abandon the idea of remote unattended e-voting. Whatever technology, secured protocol, or cryptographic algorithm we would introduce, we could not receive assurance about private voting: if a voter uses a privately owned unattended device to express a preference, unwanted third parties can record or assist that expression, opening to any corrupting, vote-selling, or menacing activity. Remote e-voting goes against the principle of anonymity requested by public consultations. For this reason, we imagined a hybrid approach, not considering a digital twin of the whole process, but rather a system including personal identification and voting in a polling station, with a digital recording of the votes on a publicly accessible ledger.

When discussing the protection of process authenticity in the context of e-voting, we need to consider many competing factors, and blockchain does not always work toward a reliable outcome. Starting from the involved actors, we considered all the components present in the traditional paper ballot process –resulting from century-long debates on modern vote expression– and their possible equivalent role in a hypothetical blockchain electronic system. The legal framework currently enacted for voting must be respected, and since each nation employs a different system, our reasonings were tailored to the Italian case for local referenda. However, the approach herein described can be reshaped to different situations.

As well described in [19], elections frequently have a diverse range of stakeholders and opponents. System designers and legislators conceived of balancing the different forces playing during the voting procedure by identifying each actor and assigning them roles, duties, and powers:

- **Voters**, attentive in the correctness of the voting procedure, but also in affirming their interest that could lead to malicious behavior;
- **Candidates** or **referendum promoters**, who want their position chosen by the electors;

- **Election officials** and **poll workers**, who are directly involved in the correctness of the procedure and could misbehave for interest;
- **Scrutineers** and **list representatives**, requiring access to all possible evidence to check for correctness;
- **Organizers** and **suppliers**, who are responsible for the actuation of the process and, therefore, the most exposed figures.

The current voting procedure explicitly or implicitly limits many possibly misconducted activities, and the designers of an e-voting system should keep existing good ideas while correcting, where possible, weaknesses and limiting the introduction of new fragilities.

Principal weaknesses in our local context regard corruption/vote-buying coupled with a tolerated sloppiness in the process execution, leading to errors or ease of misconduct. An unmodifiable ledger such as blockchain could positively impact problems such as missing ballot papers, tallying errors, inconsistencies, and transmission issues while giving more audibility options to scrutineers, list representatives, and the general public.

However, better audibility is an enemy of anonymity. One of the strengths of paper ballots is their tangible anonymity: we can count the pieces of paper, which should correspond to the voters, but we cannot understand who put that single piece in the ballot box or at what time. In a digital context, a voter who can use a record in the ledger to track her single vote will be happy to receive proof of counting her intentions, but at the same time, that record will expose her choices (Figure 3a). In general, registering single votes on a blockchain, even if expressed through a presided voting booth through a shared wallet, can expose a timestamp that malicious parties can track (Figure 3b). As a general rule, voters should never have a receipt identifying their vote, and a physical and presided ballot box is a reasonable compromise between verifiability and receipt-freeness in most cases.

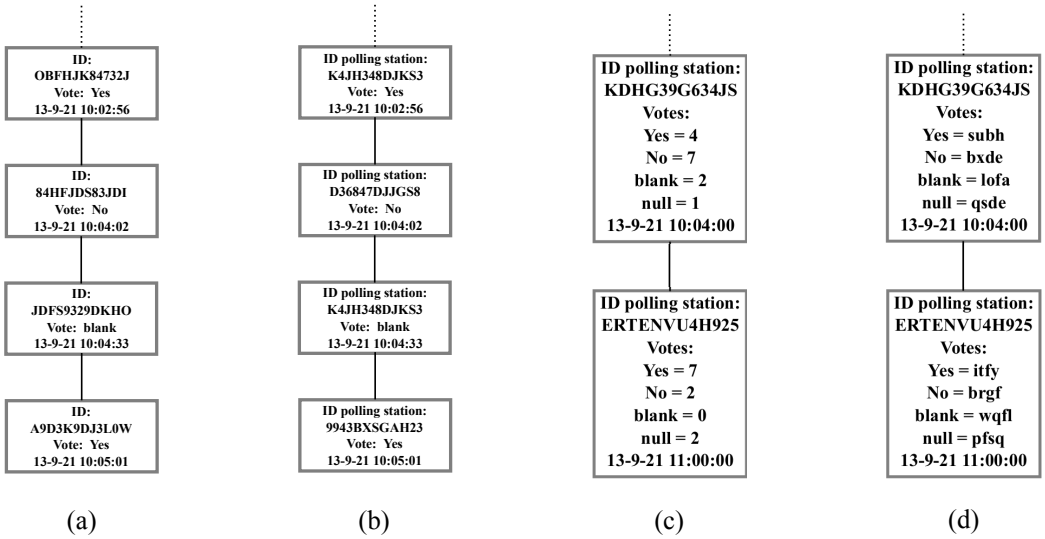

Figure 3. Schematic of possible text for e-voting transactions on a blockchain: (**a**) Remote voting with anonymization (privacy) by voter ID, vote and timestamp are exposed; (**b**) Polling station voting with anonymization by collective ID, vote and timestamp are exposed; (**c**) Polling station voting with anonymization by vote aggregation, votes and timestamps are exposed; (**d**) Polling station voting with anonymization by vote aggregation and ciphering, timestamp is exposed.

Can a blockchain help reduce the weaknesses described above without impacting the required anonymity in public consultations? We believe this is the focal question for any case study involving blockchain and e-voting, and no immediate answer is available.

Before proceeding with a tentative answer, we need to step back to the original questions on possible technological threats to e-voting systems. The highest risks in any system reside in the central controller, and for this reason, any successful attack on that controller would invalidate the outcome of the process at the whole scale. On the contrary,

one of the strengths of the traditional paper ballot in Italy is the distributed topology with local responsibilities. Each polling station has officers, workers, and scrutineers that run the local process, from preparing the premises to tallying and packaging of reports and records. If we want to mimic the actors of the traditional paper ballot in an e-voting system with polling stations, we should keep that topology in our networked system: any attack on a station would not interfere with the global system, and an eventual re-voting would be local and not complete. A blockchain ledger could fit in this idea.

At this point, we can sketch an initial draft of a blockchain e-voting system with polling stations. Let us consider a certain number of polling stations distributed on the territory based on the voters' population. Each voter receives authorization to vote from the state office in charge. We are agnostic about this procedure, but verification should involve officers in the polling stations, not some software. The polling station chair authorizes the voter by opening a session in a voting booth. Each station has a virtual ballot box, not yet well defined, collecting the electronic votes expressed in voting booths through proper voting interfaces (e.g., tablet, laptop, PC. Figure 4). When the voter completes her expression of preference, the chair closes her voting session, and her preferences get transferred from the polling hardware to the virtual ballot box. Votes can be aggregated and anonymized before sending the local tallying to a public blockchain address (Figure 3c). At the end of the polling period, anybody holding read access to the adopted blockchain can make the final tallying and declare the winner of the consultation.

Summarizing, during two years-long activity of the working group, we identified some fundamental principles to ensure that the blockchain e-voting process could mitigate traditional risks and improve some desired values for a correct democratic representation:

1. A distributed system can circumscribe possible attacks on local polling stations as long as the chosen blockchain infrastructure is solid and reliable.
2. Voters should go to a polling station and present a valid voting certificate to an officer who will consent to access a polling booth.
3. Votes must be aggregated and anonymized before recording on the blockchain.
4. Each polling station should have a blockchain wallet to transact with a central wallet and transmit aggregated votes. Those transactions should be periodic during the polling hours but with a minimum consistency in the number of votes to prevent time-related traceability. Periodicity helps report participation statistics during polling hours and improves resiliency by reducing local re-voting in case of problems or attacks.
5. Italian regulations prohibit releasing exit polls during voting hours, and therefore voting information within transactions should be encrypted (Figure 3d), and decryption keys publicly released at the end of the voting period.
6. The adopted distributed ledger should be public to allow free scrutinizing.
7. Anybody provided with suitable software, free read access to the adopted blockchain, and the decryption keys must have the opportunity to do the final overall tallying.
8. All software must be open source to let the public check functionalities. Local systems running the voting software must provide a mechanism for run-time authentication of running processes (e.g., as described in [41]).

An additional opportunity envisioned by the working group, not specific to blockchain but electronic voting in general, is the ability to implement revoting functionalities: if a voter can change her vote repeatedly, any visual record of her choice (e.g., a picture on a mobile phone) does not represent a final decision, and a malevolent third party will not have a proof for coercive finalities.

Regarding point 3 above, much discussion is still open. How can a voter trust the implemented system for the local virtual ballot box? Since receipt-freeness [16] is compulsory to prevent coercion and vote-buying, are end-to-end voting systems feasible to implement this task? Are zero-knowledge-proof or shuffle techniques feasible in this context? Using complex cryptographic protocols could deter skeptical voters from e-voting systems, obtaining a contrary effect from the desired increased democratic participation.

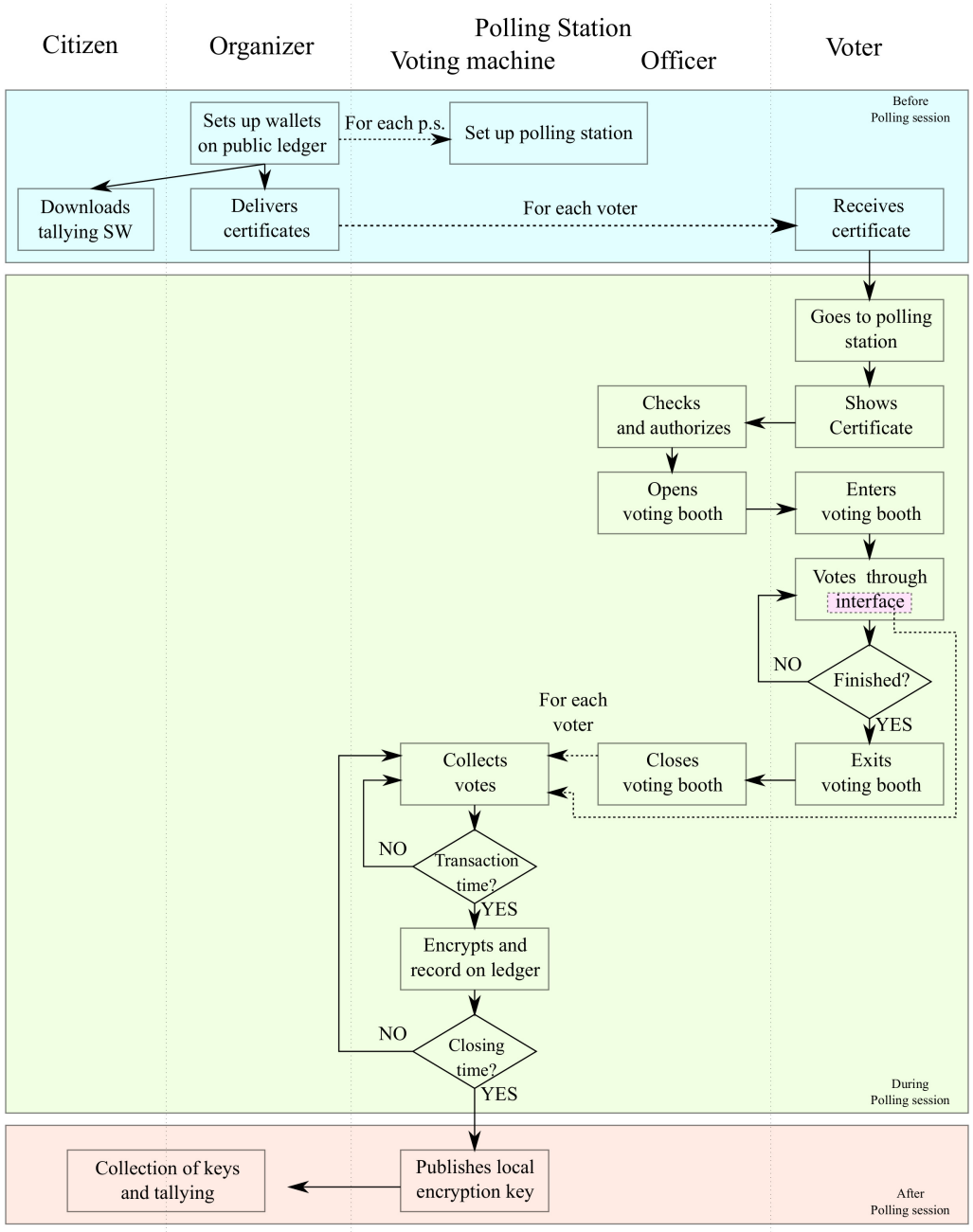

Figure 4. Draft of a blockchain e-voting system and algorithm with polling stations. The diagram is stacked for each actor.

In general, e-voting for public consultation is a challenging theme because digital anonymity is almost a chimera: blockchain looks like a promising tool. However, it introduces new risks linked to voter identification. Its anti-tampering features help design a system that improves the democratic aspects of tallying. However, solutions for the virtual ballot box must be tested in the real-world context to check acceptance from the public. Considering the political uncertainty of funding on this theme, we can expect a stable solution after several iterations of research initiatives.

6. Case Study: Secure Sharing of Electronic Health Records

Over the past two decades, many efforts have been made by researchers, policymakers and standards development organizations to propose and implement new solutions that can improve health services using ICT. Electronic booking of doctor visits, electronic prescriptions, exploration of digital medical images (such as CT, MRI, PET, etc.), and clinical

decision support systems are just a few examples that show how the digital transformation of the healthcare sector is providing powerful tools to doctors and decision-makers to make clinical services more efficient.

6.1. Interoperable Electronic Health Records

One of the most important ICT investments in healthcare is undoubtedly the EHR, which can be defined as "digitally stored health care information about an individual's lifetime with the purpose of supporting continuity of care, education and research, and ensuring confidentiality at all times" [42]. The development of an EHR system is of fundamental importance for the digitization of data and health services, and not only for the significant functions it provides regarding immediate, more accurate and effective patient history. Indeed, the implementation of an EHR requires a series of prerequisites: medical documents produced by multiple healthcare organizations must comply with the same format and share the same types of metadata, the clinical content must be represented using the same coding systems, all computer systems must be able to use the same communication protocols and security mechanisms, the software applications used in healthcare facilities, medical laboratories and general practitioners must be compatible with the same technical specifications. This means that the EHR can be seen as a driving force towards a digital health revolution in a broad area, as well as an alignment of the IT services provided by facilities located in different regions.

To achieve this aim, standardization bodies around the world have published numerous technical standards, often localized to a specific country, which define many technical aspects. The most important health IT standards are as follows:

- Health Level Seven (HL7) with the Clinical Document Architecture (CDA) standard defines the structures of the clinical document and, with the new Fast Healthcare Interoperability Resources (FHIR) standard, the data formats and REST operations;
- Clinical content is typically encoded using LOINC (laboratory observations), ICD (diseases), SNOMED CT (pathological anatomy and much more) or ATC (active principles) coding/classification systems;
- Many communication protocols are based on IHE transactions specified in integration profiles.

The adoption of international standards is also leading institutions to issue regulations regarding the interoperability between EHR systems of different countries, such as European or United States countries, in clear compliance with their respective privacy regulations, such as GDPR or HIPAA.

Despite these important advances, due to its native decentralization, the healthcare sector still presents some significant issues that need to be adequately addressed, particularly concerning the authenticity of the processes. Indeed, health processes, such as care plans, e-prescriptions, medical examinations, etc., are composed of several tasks—performed partly in sequence and partly in parallel—which typically produce relevant data. To ensure good data quality, it is essential not only that the data comply with standard formats, but also that it is produced following the correct clinical processes. In this way, all health facilities are able to perform homogeneous functions, ensuring the same quality of health services to all citizens, regardless of their residence.

Blockchain technology provides intrinsic characteristics capable of correctly addressing the current limitations. It is worth emphasizing that, in order to ensure the authenticity of the process, it is important to guarantee the security of the health record, which means protecting privacy, data ownership, integrity and sharing. With this precondition, blockchain technology, together with appropriately designed smart contracts, permits us to monitor, control and facilitate the correct execution of healthcare processes.

The next subsections organically illustrate a series of experiences that we have had to address existing problems in the healthcare sector using blockchain technology.

6.2. Health Data Management

As illustrated in [43], data privacy should be ensured after achieving secure storage. Blockchain technology makes it possible to ensure that data is stored according to high security measures. In fact, the use of hashing and other cryptographic mechanisms—along with the validation of blocks through consensus algorithms—allow for strong authenticity of health data.

A problem to be addressed in this context with the adoption of the blockchain concerns the enormous volume of data produced by healthcare facilities and patients. We analyzed this issue in [44], where we came to the solution that only the metadata containing the main relevant information about the generated documents/data should be stored in the blockchain (thus representing the asset), while the entire clinical document/data should be stored off-chain in health information systems. This approach is in line with the current IT architecture used for EHR systems, where documents/data are stored in repositories distributed at healthcare facilities while their metadata are stored in central registers in order to index the documents/data to which these metadata refer to. In this way, the patient's privacy is preserved, as no personal data are stored in the blockchain (in particular by applying pseudonymization).

More specifically, we used a private and permissioned blockchain, in which the network nodes are organized into:

- *Validating peers*, which are healthcare-related organizations with a copy of the ledger that verify the flows of transactions;
- *Endorsing peers*, which are validating peers that verify the correctness of transactions on the basis of consent at the application level;
- *Ordering peers*, which assemble and manage the transactions in blocks.

In this context, we have experimented with the adoption of blockchain technology for the management of data represented and exchanged according to the HL7 FHIR [45] standard, which in recent years is spreading all over the world for its ability to produce computational specifications, facilitating so a homogeneous implementation of software applications [46]. In particular, we designed and developed a prototype using both (i) the FHIR server, able to manage FHIR data structures (named resources); and (ii) the Hyperledger Fabric blockchain framework. Our experimentation was performed by defining a two-tier architecture system capable of exchanging messages at the application level (FHIR REST operations) and messages at the network level (Fabric transactions). To integrate these two layers, we have coupled the FHIR servers to the Fabric peers via software components (which act as Fabric clients), which are intended to intercept the high-level messages sent to the FHIR servers and forward the related transaction requests on the Fabric channels. In this way, the interceptors allow the blockchain located at the network level to always be informed of the interactions that occur at the application level, ensuring consistency between the two levels.

6.3. Health Process Authenticity

Blockchain and smart contracts are enabling technologies and protocols capable of guaranteeing the integrity and authenticity of processes, a key factor in the health sector. This condition, in fact, makes it possible to trace the correct execution of healthcare processes, helping doctors, healthcare facilities and, above all, patients to correctly follow treatment plans, clinical paths and guidelines. Furthermore, all documents and clinical data can be appropriately correlated with each other, allowing healthcare professionals to have all the clinical vision of the path that a patient is following.

To achieve this objective, we claim that it is necessary to design a system architecture based on the following approach:

1. The identified health processes must be adequately analyzed and formalized. Standards such as OMG BPMN make it possible to represent a business process in a graphic and formal way. The tasks that make up a process, their associations, as well as the actors responsible for carrying them out can be intuitively specified for both humans and machines. Collaboration between different business processes can also be represented. The use of formal process notations is crucial to ensure that the process managed by the architecture is a true *digital twin* of the actual health process carried out by the end-users.
2. The data model of the system architecture must be defined. This model has to be specified by identifying the types of data produced for each activity of all business processes considered. Particular attention must be paid to the exchange of data between different tasks, whether those tasks are carried out by the same actor or different actors.
3. The interactions concerning the tasks of the different actors allow us to identify the application interfaces of the subsystems that must be located at each actor. In this way, interfaces can be designed to adequately handle the desired flow data.
4. From the defined data model, the subset of data (transactions) to be sent to the ledger must be identified. This analysis must be carried out by identifying the assets and, consequently, the significant data that need to be immutably kept to ensure the integrity of the data to which they refer.
5. There is no single point of trust such as a central Certificate Authority (CA). Instead, according to the decentralization needs of the healthcare domain described in Section 6.1, a decentralized CA should eliminate the need for a single "trusted third party".
6. Smart contracts should be designed to encode the logic of the transactions that manage the assets. In particular, they must be able to acquire the data produced by the activities of the processes and, if the specific codified preconditions are met, send them to the ledger. In this way, smart contracts can act as an interface between the application and network layers and automate the execution of processes, ensuring their correct implementation and integrity.

As a case study for implementing this approach, we consider the Patient Summary (PS) feeding process [47]. The PS is a relevant document generated and regularly updated by the General Practitioner (GP), as it contains the most relevant clinical information of the patient (such as possible allergies, interventions, problems, pharmaceutical therapies in progress, etc.). In general, the GP is not always immediately informed of the existence of new clinical information relating to a patient, as it is constantly produced during the patient's contact with the health system. However, the importance of this document to physicians is paramount, as it allows for an instant view of the patient's history. For this reason, tools are needed to facilitate communication with the GP regarding the new clinical information available. It is worth noting that the GP has the ultimate responsibility for deciding what information should or should not be included in a PS.

Figure 5 shows a BPMN model that describes a possible process. In this process, there are three actors: a medical specialist, a GP and the EHR. The scenario considered consists of a medical visit that a patient makes with the medical specialist, at the end of which he/she generates a clinical report, which is indexed in the EHR and notified to the GP. In this way, the GP can retrieve the document from the EHR and—if he/she deems it appropriate—update the patient's PS. This latter document must also be retrieved from the EHR. In addition, its updated version must be indexed in the EHR, replacing the old version.

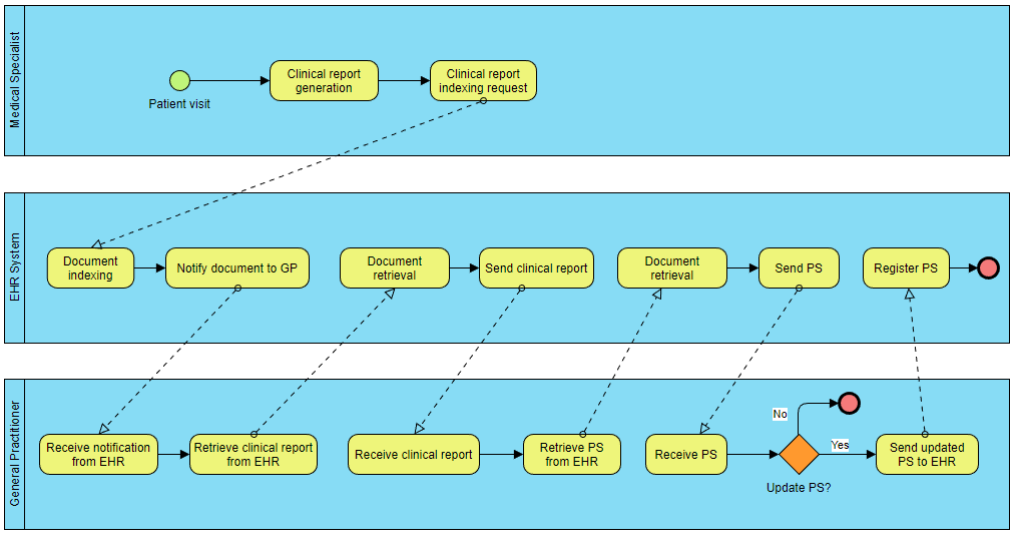

Figure 5. Example of a process represented in BPMN, where a clinical report produced after a patient's visit with a specialist doctor is recorded in the EHR, retrieved by the GP and used for a possible update of the PS.

All types of data produced in each activity of the entire process permit us to identify the data model of the system architecture. In particular, the task *"Notify document to GP"* in the figure has the purpose of sending the most relevant metadata of the clinical report to the GP, in order to inform him/her about the generation of the new document. In this scenario, these data consist of the document identifier, document type, patient identifier and process identifier. Therefore, it is necessary to design and develop an application interface used by the GP in able to manage this data.

As shown in the UML component diagram in Figure 6, a specific software component of the system (i.e., an interceptor) must be able to intercept this data and send it to the *"Register document"* and *"Retrieve document"* smart contracts, which have been designed and developed specifically for this purpose. These smart contracts are intended to (i) analyze the correctness of the data; (ii) identify the process; (iii) verify the user access rights; (iv) possibly transform the data into a specified transaction format; and (v) send/retrieve the transaction to/from the blockchain. Using the properties of blockchains is possible to build architectures based on decentralized CAs, as described in [48]. In this context, the transaction is composed of the same types of data received by the smart contract, even if the smart contract is in charge of replacing the patient identifier with a pseudonym. The verification of the user access rights is carried out by using specific access control models. In healthcare, the Attributed-Based Access Control Model (ABAC) is often used, as it easily permits us to design policies that grant or deny access to services and resources on the basis of a number of attributes declared by the user (represented in digitally signed authorization tokens), such as professional role (GP, specialist, nurse, etc.), purpose of use (emergency/urgent or ordinary access), patient consent and so on.

After receiving the notification, the GP can thus retrieve the clinical report from the EHR system, which verifies the integrity of the data and the process by interacting with the blockchain. Finally, the GP decides whether to add more information to the PS.

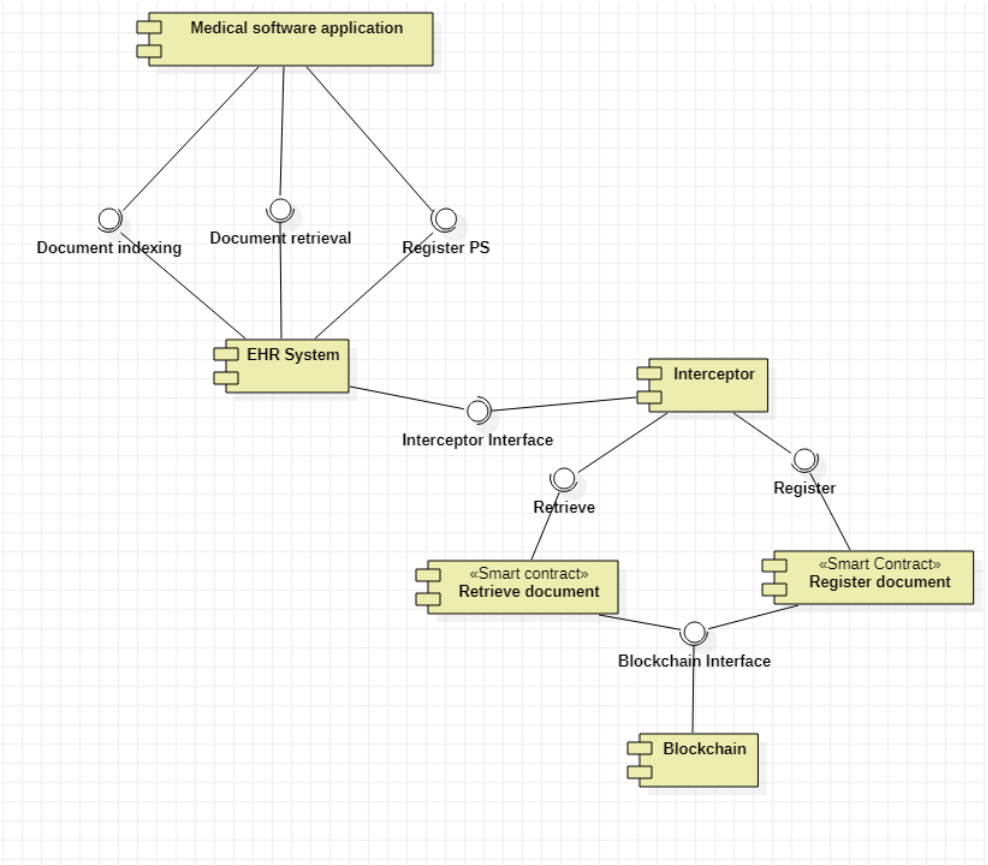

Figure 6. Component diagram of the system architecture.

7. Conclusions

Trying to address the general problem of process authentication, in this work, we identify some structural concepts and tools functional to introduce a framework for the authentication of the digital twin of a process using DL technology.

Through the description of three different case studies, we present the outcome of unrelated activities, sharing the difficulties of authenticating the process steps in a non-natively digital activity. Each case study exposes limitations and possible solutions to implement the process authentication such that it could result reasonably acceptable for the community of the application context.

We believe that through a methodological approach, the analysis of a process may expose the critical points requiring stringent authentication, often leading to a more profound awareness of inherent weaknesses. On the other hand, creating a digital twin of a process could lead to deceitful confidence in the chosen strategies. By implementing our case studies, we found that a digital twin of a process could introduce new issues not present in the physical context.

Therefore, we envision the creation of the digital twin of a process as a cyclic procedure involving: an initial mapping of the steps to the physical-world process, a fixing of the digital issues that could arise, and a final inverse checking of the mapping, repeating the procedure until necessary. A formalization of this idea will be the object of future study.

Moreover, human concurrence is focal when managing some of the steps or sub-steps in physical processes, and its respective integrity is also still the object of future research.

Author Contributions: Conceptualization, G.S.; methodology, M.C., D.R. and G.S.; writing—original draft preparation, M.C., D.R. and G.S.; writing—review and editing, M.C., D.R. and G.S.; supervision, G.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the European Commission, grant number 883273, AI4HEALTHSEC—A Dynamic and Self-Organized Artificial Swarm Intelligence Solution for Security and Privacy Threats in Healthcare ICT Infrastructures, and by the Italian Ministry of Economic Development, prog. n. F/190111/01-02/X44, PRODE—Processo di Dematerializzazione in qualità.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Tommaso Coppola, Maria Francesca De Tullio, Mauro Forte, Davide Lo Pilato, Rosario Scognamiglio, and Erica Vaccaro for their significant contribution to the volunteer working group on blockchain e-voting.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Natarajan, H.; Krause, S.; Gradstein, H. *Distributed Ledger Technology (DLT) and Blockchain*; World Bank Group: Washington, DC, USA, 2017.
2. Belchior, R.; Correia, M.; Vasconcelos, A. JusticeChain: Using Blockchain to Protect Justice Logs. In *On the Move to Meaningful Internet Systems: OTM 2019 Conferences*; Springer: Berlin, Heidelberg, 2019; pp. 318–325, https://doi.org/10.1007/978-3-030-33246-4_21.
3. Rosa, M.; Barraca, J.P.; Rocha, N.P. Blockchain structures to guarantee logging integrity of a digital platform to support community-dwelling older adults. *Clust. Comput.* **2020**, *3*, 1887–1898.
4. Maull, R.; Godsiff, P.; Mulligan, C.; Brown, A.; Kewell, B. Distributed ledger technology: Applications and implications. *Strateg. Chang.* **2017**, *26*, 481–489. https://doi.org/10.1002/jsc.2148.
5. Belotti, M.; Božić, N.; Pujolle, G.; Secci, S. A vademecum on blockchain technologies: When, which, and how. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 3796–3838.
6. Chowdhury, M.J.M.; Ferdous, M.S.; Biswas, K.; Chowdhury, N.; Kayes, A.S.M.; Alazab, M.; Watters, P. A comparative analysis of Distributed Ledger Technology platforms. *IEEE Access* **2019**, *7*, 167930–167943. https://doi.org/10.1109/ACCESS.2019.2953729.
7. Monrat, A.A.; Schelén, O.; Andersson, K. A survey of Blockchain from the perspectives of applications, challenges, and opportunities. *IEEE Access* **2019**, *7*, 117134–117151. https://doi.org/10.1109/ACCESS.2019.2936094.
8. Calandra, D.; Secinaro, S.; Massaro, M.; Dal Mas, F.; Bagnoli, C. The link between sustainable business models and blockchain: A multiple case study approach. *Bus. Strategy Environ.* **2022**. https://doi.org/10.1002/bse.3195.
9. Spanò, R.; Massaro, M.; Iacuzzi, S. Blockchain for value creation in the healthcare sector. *Technovation* **2021**, 102440. https://doi.org/10.1016/j.technovation.2021.102440.
10. Secinaro, S.; Dal Mas, F.; Brescia, V.; Calandra, D. Blockchain in the accounting, auditing and accountability fields: A bibliometric and coding analysis. *Account. Audit. Account. J.* **2021**, 102440. https://doi.org/10.1108/AAAJ-10-2020-4987.
11. Kawa, A.; Maryniak, A. *SMART Supply Network*; Springer: Berlin/Heidelberg, Germany, 2019.
12. Wang, Y.; Chen, C.H.; Zghari-Sales, A. Designing a blockchain enabled supply chain. *Int. J. Prod. Res.* **2021**, *59*, 1450–1475.
13. Vu, N.; Ghadge, A.; Bourlakis, M. Blockchain adoption in food supply chains: A review and implementation framework. *Prod. Plan. Control* **2021**, 1–18. https://doi.org/10.1080/09537287.2021.1939902.
14. Kurpjuweit, S.; Schmidt, C.G.; Klöckner, M.; Wagner, S.M. Blockchain in additive manufacturing and its impact on supply chains. *J. Bus. Logist.* **2021**, *42*, 46–70.
15. Wang, K.H.K.; Mondal, S.K.; Chan, K.C.; Xie, X. A Review of Contemporary E-voting: Requirements, Technology, Systems and Usability. *Data Sci. Pattern Recognit.* **2017**, *1*, 31.
16. Baudron, O.; Fouque, P.A.; Pointcheval, D.; Stern, J.; Poupard, G. Practical Multi-Candidate Election System. In Proceedings of the Twentieth Annual ACM Symposium on Principles of Distributed Computing, PODC '01, Newport, RI, USA, 26–29 August 2001; Association for Computing Machinery: New York, NY, USA, 2001; pp. 274–283. https://doi.org/10.1145/383962.384044.
17. Gibson, J.P.; Krimmer, R.; Teague, V.; Pomares, J. A review of E-voting: The past, present and future. *Ann. Telecommun.* **2016**, *71*, 279–286. https://doi.org/10.1007/s12243-016-0525-8.
18. Abuidris, Y.; Kumar, R.; Wenyong, W. A Survey of Blockchain Based on E-voting Systems. In Proceedings of the 2019 2nd International Conference on Blockchain Technology and Applications, ICBTA 2019, Xi'an, China, 9–11 December 2019; pp. 99–104. https://doi.org/10.1145/3376044.3376060.
19. Park, S.; Specter, M.; Narula, N.; Rivest, R.L. Going from bad to worse: From Internet voting to blockchain voting. *J. Cybersecur.* **2021**, *7*, tyaa025. https://doi.org/10.1093/cybsec/tyaa025.
20. Azaria, A.; Ekblaw, A.; Vieira, T.; Lippman, A. MedRec: Using Blockchain for Medical Data Access and Permission Management. In Proceedings of the 2016 2nd International Conference on Open and Big Data (OBD), Vienna, Austria, 22–24 August 2016; pp. 25–30. https://doi.org/10.1109/OBD.2016.11.
21. Zhang, P.; White, J.; Schmidt, D.C.; Lenz, G.; Rosenbloom, S.T. FHIRChain: Applying Blockchain to Securely and Scalably Share Clinical Data. *Comput. Struct. Biotechnol. J.* **2018**, *16*, 267–278. https://doi.org/10.1016/j.csbj.2018.07.004.

22. Liang, X.; Zhao, J.; Shetty, S.; Liu, J.; Li, D. Integrating blockchain for data sharing and collaboration in mobile healthcare applications. In Proceedings of the 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), Montreal, QC, Canada, 8–13 October 2017; pp. 1–5. https://doi.org/10.1109/PIMRC.2017.8292361.
23. Shahnaz, A.; Qamar, U.; Khalid, A. Using blockchain for electronic health records. *IEEE Access* **2019**, *7*, 147782–147795. https://doi.org/10.1109/ACCESS.2019.2946373.
24. Capece, G.; Lorenzi, F. Blockchain and healthcare: Opportunities and prospects for the EHR. *Sustainability* **2020**, *12*, 9693. https://doi.org/10.3390/su12229693.
25. Wang, H.; Song, Y. Secure cloud-based EHR system using attribute-based cryptosystem and blockchain. *J. Med. Syst.* **2018**, *42*, 1–9. https://doi.org/10.1007/s10916-018-0994-6.
26. Grieves, M. Digital twin: Manufacturing excellence through virtual factory replication. *White Pap.* **2014**, *1*, 1–7.
27. Cheng, Y.; Zhang, Y.; Ji, P.; Xu, W.; Zhou, Z.; Tao, F. Cyber-physical integration for moving digital factories forward towards smart manufacturing: A survey. *Int. J. Adv. Manuf. Technol.* **2018**, *97*, 1209–1221.
28. Kholopov, V.; Antonov, S.; Kurnasov, E.; Kashirskaya, E. Digital twins in manufacturing. *Russ. Eng. Res.* **2019**, *39*, 1014–1020.
29. Boschert, S.; Rosen, R., Digital Twin—The Simulation Aspect. In *Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and Their Designers*; Springer International Publishing: Cham, Switzerland, 2016; pp. 59–74. https://doi.org/10.1007/978-3-319-32156-1_5.
30. Mandolla, C.; Petruzzelli, A.M.; Percoco, G.; Urbinati, A. Building a digital twin for additive manufacturing through the exploitation of blockchain: A case analysis of the aircraft industry. *Comput. Ind.* **2019**, *109*, 134–152.
31. Szabo, N. Formalizing and Securing Relationships on Public Networks. 1997. Available online: http://journals.uic.edu/ojs/index.php/fm/article/view/548/469 (accessed on 11 August 2017).
32. Romano, D.; Schmid, G. Beyond Bitcoin: Recent Trends and Perspectives in Distributed Ledger Technology. *Cryptography* **2021**, *5*, 36. https://doi.org/10.3390/cryptography5040036.
33. Cooper, D.; Santesson, S.; Farrell, S.; Boeyen, S.; Housley, R.; Polk, W. Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. Standards Track 5280, Internet Engineering Task Force. 2008. Available online: www.rfc-editor.org/rfc/rfc5280.txt (accessed on 11 August 2017).
34. Soelman, M.; Andrikopoulos, V.; Pérez, J.A.; Theodosiadis, V.; Goense, K.; Rutjes, A. Hyperledger Fabric: Evaluating Endorsement Policy Strategies in Supply Chains. In Proceedings of the 2020 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS), Oxford, UK, 3–6 August 2020; pp. 145–152. https://doi.org/10.1109/DAPPS49028.2020.00019.
35. Wolford, B. What is GDPR, the EU's New Data Protection Law? 2020. Available online: https://gdpr.eu/what-is-gdpr/ (accessed on 5 September 2021).
36. DECRETO LEGISLATIVO 7 marzo 2005, n. 82. Codice dell'amministrazione Digitale. 2020. Available online: https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legislativo:2005-03-07;82 (accessed on 5 September 2021).
37. ISO. *Space Data and Information Transfer Systems—Open Archival Information System (OAIS)—Reference Model*; Iso Standard; International Organization for Standardization: Geneva, Switzerland, 2012.
38. ISO/IEC. *Information Technology—Security Techniques—Information Security Management Systems—Requirements*; ISO/IEC Standard; International Organization for Standardization/International Electrotechnical Commission: Geneva, Switzerland, 2013.
39. EU. *Regulation on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation)*; Eu Regulation; European Union: Brussels, Belgium, 2013.
40. EU NIS Cooperation Group. Compendium on Cyber Security of Election Technology. Technical Report. 2018. Available online: http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=53645 (accessed on 18 September 2021).
41. Almohri, H.M.; Yao, D.; Kafura, D. Process Authentication for High System Assurance. *IEEE Trans. Dependable Secur. Comput.* **2014**, *11*, 168–180. https://doi.org/10.1109/TDSC.2013.29.
42. Iakovidis, I. Towards personal health record: Current situation, obstacles and trends in implementation of electronic healthcare record in Europe1Disclaimer: The view developed in this paper is that of the author and does not necessarily reflect the position of the European Commission.1. *Int. J. Med. Inform.* **1998**, *52*, 105–115. https://doi.org/10.1016/S1386-5056(98)00129-4.
43. Shi, S.; He, D.; Li, L.; Kumar, N.; Khan, M.K.; Choo, K.K.R. Applications of blockchain in ensuring the security and privacy of electronic health record systems: A survey. *Comput. Secur.* **2020**, *97*, 101966. https://doi.org/10.1016/j.cose.2020.101966.
44. Ciampi, M.; Esposito, A.; Marangio, F.; Schmid, G.; Sicuranza, M. A blockchain architecture for the Italian EHR system. In Proceedings of the Fourth International Conference on Informatics and Assistive Technologies for Health-Care, Medical Support and Wellbeing, Valencia, Spain, 24–28 November 2019; Volume 35.
45. HL7 FHIR: Welcome to FHIR. Available online: https://hl7.org/fhir/ (accessed on 16 September 2022).
46. Ciampi, M.; Esposito, A.; Marangio, F.; Sicuranza, M.; Schmid, G. Modernizing healthcare by using blockchain. In *Applications of Blockchain in Healthcare*; Springer: Singapore, 2021; pp. 29–67. https://doi.org/10.1007/978-981-15-9547-9_2.
47. Ciampi, M.; Marangio, F.; Schmid, G.; Sicuranza, M. A blockchain-based smart contract system architecture for dependable health processes. In Proceedings of the Italian Conference on CyberSecurity, CEUR-WS, Online, 7–9 April 2021.
48. Xie, R.; Wang, Y.; Tan, M.; Zhu, W.; Yang, Z.; Wu, J.; Jeon, G. Ethereum-Blockchain-Based Technology of Decentralized Smart Contract Certificate System. *IEEE Internet Things Mag.* **2020**, *3*, 44–50. https://doi.org/10.1109/IOTM.0001.1900094.