Blockchain and Distributed Ledgers as Trusted Recordkeeping Systems: An Archival Theoretic Evaluation Framework

Conference Paper · November 2017

CITATIONS
45

READS
6,605

1 author:

Victoria Lemieux
University of British Columbia - Vancouver
103 PUBLICATIONS **646** CITATIONS

SEE PROFILE

Blockchain and Distributed Ledgers as Trusted Recordkeeping Systems:

An Archival Theoretic Evaluation Framework

Victoria L. Lemieux
School of Library, Archival and Information Science
The University of British Columbia
Vancouver, Canada
vlemieux@mail.ubc.ca

Abstract— Blockchains and distributed ledger technology promises trusted and immutable records in a wide variety of use cases involving recordkeeping, including real estate and healthcare. This paper presents a novel framework for evaluating the capability of innovative blockchain-based systems to deliver trustworthy recordkeeping based on archival science-an ancient science aimed at the long-term preservation of authentic records.

Keywords—archival science; authenticity; blockchain; distributed ledger; integrity; reliability; trust

I. INTRODUCTION

Blockchain and distributed ledgers have burst onto the scene in the past few years as an important future technology. Though there is, as yet, no single, internationally agreed upon definition of blockchain or distributed ledgers, they are often described as "an open-source technology that supports trusted, immutable records of transactions stored in publicly accessible, decentralized, distributed, automated ledgers" [1]. In a relatively short time span these technologies have become the innovation to watch according to just about every technology research and advisory firm, global consultancy, and international think tank. The technology is even said to have reached the top of the Gartner hype curve [2]. Blockhchain and distributed ledgers represent more than just hype, however.

Governments and organizations around the world are beginning to look seriously at the application of this technology, and some have already implemented it. The main drawing card of this innovative technology is the production of immutable trustworthy records without need of a trusted third party.

A case in point is the sale of land. In traditional land transfers, the process often begins with the listing of the property on a real estate market, the exchange of contracts during negotiations over price, and the completion of the sale by registration with a state-run land titles registration authority. The problem with the traditional approach to land transfers is that, at least in some jurisdictions, the process is slow and cumbersome, being often reliant on manual recording of transactions by land registration authorities, and open to fraud and corruption. A number of jurisdictions are experimenting with the application of blockchain technology to address these issues.

The government of Georgia, for example, piloted the registration of land titles using a private blockchain in 2016 and has plans to expand the service to sales and purchases of land titles, mortgages, rentals, new land title registration, property demolition and notary services [3]. The Swedish land registry authority, Lantmäteriet, has been testing a way to record property transactions on a blockchain and is estimated to have saved $106 million annually by reducing paper work, eliminating fraud, and speeding up transactions [4]. A pilot of the application of blockchain technology to land transfer registration in the municipality of Pelotas in Brazil has recently been launched by the

local real estate registration authority [5].

Recording of land transactions is by no means the only recordkeeping use case. Many jurisdictions are looking at blockchain to securely keep health records and a host of other types of government records as well [6].

Will blockchain technology deliver on the promise of producing and keeping trustworthy records? Critical reflections on the core capabilities of blockhains and distributed ledgers encompass a variety of perspectives and issues, including security, privacy, scalability, interoperability, legal uncertainty, and monetary concerns. So far, however, assessments of the technology have not incorporated an archival science perspective. This is an odd oversight given that many of the use cases involve recordkeeping, which is based on archival science. It is a gap that this paper aims to begin to fill by presenting an archival science-based theoretical framework for evaluating blockchain technology as systems for the production and keeping of trustworthy records and presenting an initial assessment of blockchain-based recordkeeping against this framework.

II. Archival Science – An Overview

Archival science is concerned with the long-term preservation of authentic records, which is commonly misidentified as being associated only with repositories of ancient, often very dusty, tomes. The ancient theories and principles of archival science are still relevant today, and they apply as readily to digital records and recordkeeping as to the dusty old volumes of yesteryear. These theories and principles began to be systematized in the middle ages with the first university course in the precursor to contemporary archival science, a course in notarial arts, offered in 1158 at the University of Bologna [7]. In the seventeenth century, the formal study of records (called Diplomatics), grew out of a need to establish the authenticity of medieval documents at a time when there was an increasing number of forgeries related to European legal conflicts [8]. As historical documents of questionable authenticity were often presented as evidence of rights, "the need for alternative ways of establishing authenticity increased, and techniques of documentary criticism began to be developed and formalized." [8] These techniques were first systematized in 1681 by Dom Jean Mabillon in *De re Diplomatica Libre VI,* which included instruction on the organization and operation of records offices, including personnel, regulations and the process of records creation, routing, storage, and preservation [7]. The teaching of Diplomatics spread to faculties of law across Europe, and in 1821 led to the founding of the Ecole des Chartes in Paris. At this time, archival education expanded to include the study of records in support of both legal and historical research, thus laying the foundation for contemporary archival science [7]. Just as the form of records has evolved through time – from cuneiform on clay tablets, to papyrus, to wax cylinders, paper and now digital systems, so the focus of archival science has shifted from these ancient physical forms of records to newer digital forms. Archival science now has a prospective aspect in its focus on the design of systems (broadly defined as comprising human and technical infrastructures and processes) that result in the long-term preservation of authentic records [9], a concurrent aspect in the active preservation of the authenticity of records throughout their life cycle, often in archival institutions, and a retrospective aspect in the assessment of the genuineness or authenticity of records, which is often referred to as digital records forensics [10].

III. An Archival Theoretic Framework for Evaluating Immutable Trust

In archival science, a record is said to be trustworthy if it is assessed as being accurate, reliable and authentic. These main attributes can be further decomposed into additional attributes, as shown in Fig. 1. This assessment, usually done by an expert human assessor, is probabilistic in nature. That is, it is an assessment that often has to rely on imperfect information, given the uncertain origins of many archival documents, and therefore involves inferential reasoning about the main characteristics of records upon which assertions about their trustworthiness can be based.

In order to be considered trustworthy, records first

must be seen to be accurate. The InterPARES terminology database defines accuracy as "The degree to which data, information, documents or records are precise, correct, truthful, free of error or distortion, or pertinent to the matter" [11]. The Society of American Archivists' glossary [12] defines it as: "The degree of precision to which something is correct, truthful, and free of error or distortion, whether by omission or commission." Accuracy thus refers directly to the truth-value of the content (facts) of the record. The archival definitions of accuracy align with common understandings of the term.

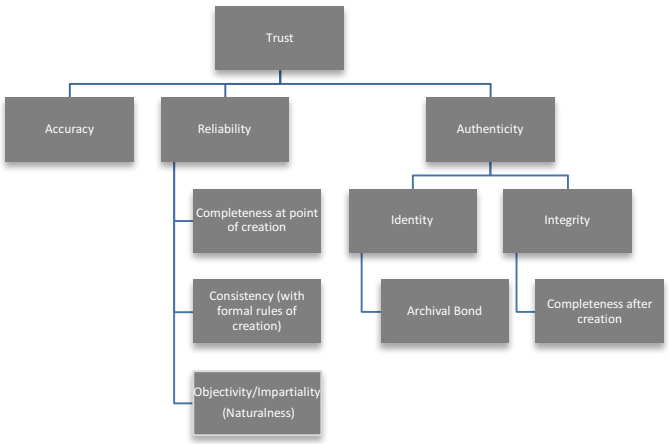

Fig. 1. A taxonomy of key archival concepts and their relationship to trust (author's own rendering)

Reliability is another of the main dimensions of trustworthiness from an archival science perspective. In archival science, the term reliability refers to "the trustworthiness of a record as a statement of fact; that is, to its ability to stand for the facts it is about" [13, 525]. Records are created to effect some real-world act, and to memorialize such acts. As such, a reliable record will serve as a mirror of the facts about the acts it enacts and be as such a "good" representation of these acts and the facts pertaining to them. A record can stand in for the act itself. Thus, an original copy of a land title registration stands for the transfer of title to a piece of land into the hands of a new landholder.

To achieve reliability, records must have three characteristics: completeness at the point of creation; consistency with formal rules of creation; and "naturalness". In archival terms, completeness is linked to the transactional nature of records and refers to the presence of all the elements required by the creator and a legal-administrative system for the record to be capable of generating consequences [14]. This typically includes signatures and dates of creation [7]. To illustrate, a contract for sale of land that does not possess a signature and date would not be considered complete. Completeness as an archival science concept is thus intrinsic to the record and associated with its formal characteristics. A trustworthy record is also one that possesses physical and formal elements which are consistent with authentic records of similar provenance (e.g., whether the ink used to write a document is contemporaneous with the document's purported date, or whether the style and language of the document is consistent with other related documents that are accepted as authentic) [7]. Finally, trustworthy records will possess naturalness. This refers to the fact that, typically, records are generated in the course of business or daily life, and are thus not usually designed purposefully to disseminate knowledge or opinion, like, for example, books or other publications. As such, they have traditionally been thought to possess qualities of unselfconsciousness that underpin their reliability as records [7]. This notion underpins the legal "business records exception to hearsay" rule, which accepts a record as standing for the facts referred to in it by virtue of the naturalness of its creation [15].

Beyond accuracy and reliability, records must also be judged to be authentic. Archival authenticity is defined as "the trustworthiness of a record as a record; i.e., the quality of a record that establishes that it is what it purports to be and that it is free from tampering or corruption" [16]. It also encompasses the idea that the records are entitled to acceptance, that they are authoritative or duly authorized, and that their origin or authorship is genuine. For a record to be considered authentic, it must have been created by the individual represented as the creator. The presence of a signature, whether it be physical or digital, serves as a test for authenticity; the signature identifies the

creator and establishes the relationship between the creator and the record.

Note that, in archival authenticity, genuineness of the creator of the record does not imply or provide a basis for inferences about the truth-value of the facts in the record; it merely establishes that the purported creator of the record is genuine and that the creator possesses the authority to make the record [7] [15]. Evidence scholars and historians traditionally distinguish between two kinds of trustworthiness: the reliability of a record, which refers to its truth-value as a statement of facts, assessed in relation to the proximity of the observer and recorder to the facts recorded, and authenticity of a record, which refers to its truth value as a representation of the facts it recorded, assessed in relation to the document's continued likeness to its original instantiation. To illustrate the distinction in reference to the issue of fake news: postings from mainstream and alternative media outlets may be treated as authentic even if the veracity of the reporting and writing is in question [17].

There are two necessary preconditions for authenticity: identity and integrity of the record. It is impossible to establish the genuineness of a record unless the identity of the record to be authenticated is clear and, in the case of detection of forgeries, distinct from the identity of the record to which it will be compared. The unique identity of a record as a record is established by the instantiation and maintenance of the archival bond. A record is an "intellectual object" that is "made or received in the course of an activity as an instrument or a byproduct of such activity, and set aside for action or reference" [18] Thus, "a record has a determinate relationship to the activity of which it is a record, to the actor who kept it as a record and to other records of the same activity. This relationship, called the 'archival bond,' not only relates a record to a specific context of creation and use but also defines the Archival Aggregate in which it belongs" [18]. Without reference to the archival bond, it is impossible to tell if a record is genuine or a forgery. For example, if handed a manuscript entitled "Ulysses" which discusses the hero's journey and purports to be written by James Joyce, it may be impossible to tell if that was a manuscript written by the famous author or a university professor's lecture notes on Greek mythology without reference to the archival bond. Understanding the archival bond is often done via the analysis of a record's provenance. That is, through an examination of the "relationships between records and the organizations or individuals that created, accumulated and/or maintained and used them in the conduct of personal or corporate activity" [18].

Further, if the integrity of a record is compromised, it is impossible to establish a record's genuineness with any degree of certainty. The concept of integrity, along with the concept of identity, forms the basis of establishing and assessing the authenticity of records over the long term. In order to remain authentic, records must remain free from tampering, corruption, or alteration over time [19]. In the pre-digital era, integrity controls included numbered entries in registers, listing file contents, and numbering individual documents in file folders. In the digital era, the concept of integrity has expanded to include reliable operation of information systems in which records are created and the infrastructures on which they are maintained. Assuring integrity in such systems consists of a broad range of measures such as access controls, user authentication and verification, audit trails, as well as documentation that demonstrates the normal functioning, regular maintenance, and frequency of upgrades of records systems [20]. Preservation of the integrity of records over the long term falls within the domain of digital preservation.

Within the digital preservation community, it is recognized that preserving the integrity of the bit structure of data is not a sufficient form of preservation because semantic loss may prevent later interpretability and accessibility. To illustrate, it may be possible to preserve a bit stream of a digital version of a land title, and even to preserve the software that renders the bit stream interpretable, but the ability to understand the significance and meaning of the bits depends upon preservation of information about the context of their creation in order to render them interpretable

and also so that the record does not lose its real world effect, such as conferring an entitlement [21]. It is possible to have some degree of bit loss without a detrimental impact upon "renderability", interpretability, or effect; however, even with perfect preservation of bits, interpretability and effect may be compromised. This understanding characterizes the archival notion of completeness after creation. Digital records preservation therefore involves preservation of the integrity of the identity of records, through preservation of the archival bond, in addition to preservation of the integrity of the general semantic context, content, and form of data.

IV. Analysis of the Trustworthiness of Blockchain Recordkeeping Systems

Having set out an archival theoretic framework for analysis of trust in records and recordkeeping systems, it is now possible to analyze whether blockchain recordkeeping systems deliver on their promise of producing trustworthy immutable records. Naturally, such analysis is limited without reference to specific implementations of the blockchain technology; however, it is still possible to make some initial observations in relation to generic reference architectures representative of current blockchain recordkeeping applications, as set out in Fig. 2.

A. *Reference Architecture and Operation of Blockchain-Based Recordkeeping Systems*

In a number of blockchain applications for recordkeeping, such as land titles registration, health recordkeeping, or tax recordkeeping, to name just a few use cases, distributed applications (or DApps) run as a user-facing web-based application layer that reads from or writes to the other layers of the blockchain technology stack [22] [23]. DApp web-forms embed domain-specific business and data logic and rules and typically enable structured data entry, presentation and processing. To function, DApps may access both blockchains/distributed ledgers and off-chain services, such as storage or operational transactional databases, through application interfaces [22] [23].

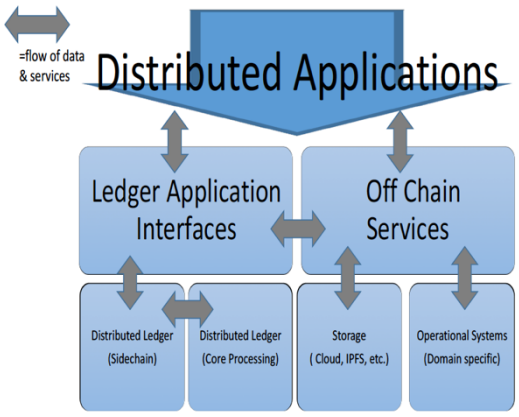

Fig. 2. Generic Blockchain Recordkeeping Reference Architecture (author's own rendering)

As a general rule, data representing transactional records is created off-chain in operational systems and usually stored off chain as well in, for example, a database, a cloud-based repository or the Inter-Planetary File System (IPFS). However, it is increasingly likely that records will be created and stored on chain using smart contracts [22] [23], which are user created business rules, such as for the transfer of ownership of a property, implemented in an executable software module which performs functions defined by the module [1]. The smart contract code, which is stored in a distributed ledger, determines what transactions are recorded into the blockchain, under what conditions, and what information they contain. Smart contracts also usually have the ability to read and write through program interfaces to data stores which are separate from the blockchain itself and can be updated when transactions occur [23]. The business logic contained in a smart contract creates or operates on business data that is contained in such external data stores [23].

It is often the case that the processing of transactions is handled in specialized, permissioned and private sidechains - a mechanism that allows tokens from one blockchain to be securely used within a completely separate blockchain but still moved back to the original chain if necessary- before being "anchored" into a public or main

blockchain [1]. This architecture allows for more efficient processing of transactions as well as for the flexibility to customize consensus mechanisms, contract capabilities, data capabilities, and other aspects of the operation of the blockchain-based recordkeeping service within the sidechain [24].

At the core blockchain processing layer, processing may proceed as follows: blockchain address A proposes the transfer of a token to another address B. Next the distributed "mesh" network checks the public ledger that sufficient tokens exist in the wallet at address A. If there is sufficient value, specialized nodes called miners will bundle the proposed transfer with other transactions to create a new block for the blockchain. Here it is important to note that the bundling of reputable transactions into blocks is completely agnostic as to the nature of those representations of transactions (i.e., they can relate to any transaction of any type from any source in a public blockchain like bitcoin). The blocks are cryptographically "hashed"; that is, they are used as input to an algorithm that converts them into a fixed-size alphanumeric string, which is called the hash value (sometimes also called a message digest, a digital fingerprint, a digest or a checksum).

That hash is put, along with some other data (e.g., a nonce), into the header of the proposed block. This header then becomes the basis for the "proof of work" performed by the miner nodes on the network. When a miner node arrives at a solution to the proof of work, other nodes check it and then each node that confirms the solution updates the blockchain with the hash of the header of the proposed block. This becomes the new block's identifying string, now part of the distributed ledger in the blockchain. Address A's payment to address B, and all the other transactions the block contains, are confirmed [39].

B. *Evaluating the Reference Architecture using the Archival Theoretic Framework*

The archival theoretic framework for evaluating immutable trust can be used to assess the reference architecture and operating model described in the previous section. Such an evaluation can expose gaps that must be filled in order to achieve the production and preservation of trusted and immutable records from an archival perspective. Fig. 3 shows those aspects of the archival theoretic framework for trusted immutable records that, based on preliminary analysis, are often in scope of current blockchain recordkeeping solutions (colored green). These include integrity of records, which blockchain-based systems are generally designed to protect. Those that are often out of scope (colored red) include accuracy and reliability, most often because they are instantiated in off-chain systems, such as databases, but also because such features are not explicitly designed into blockchain-based recordkeeping solutions, as in the cause of the archival bond. Persistence through time is also not usually explicitly addressed. The following sections discuss these issues in greater detail.

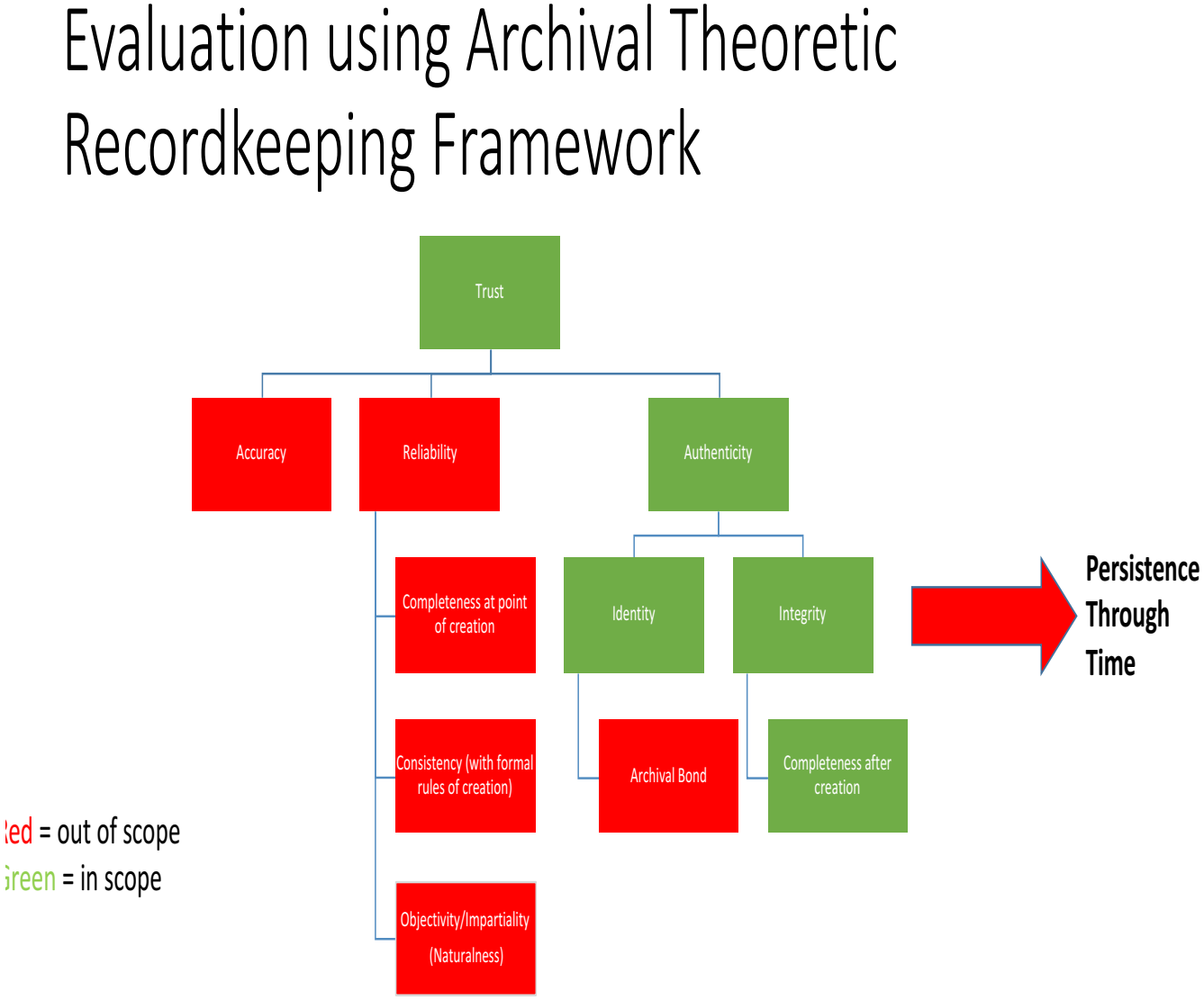

Fig. 3. Preliminary high-level evaluation of aspects of trustworthy recordkeeping addressed in blockchain recordkeeping solutions based on a generic reference architecture and operating model.

Accuracy and Reliability. Although blockchain-based recordkeeping solutions are often advanced as solutions to protect recordkeeping systems from tampering, corruption and fraud, thereby purporting to improve the accuracy and reliability of such systems. In theory, they should make no material improvement upon these dimensions of the trustworthiness of records. In systems where transactional records continue to be generated from data and processes in operational transactional systems, and only a hash of such transactional records is recorded into a blockchain, accuracy and reliability still are determined by the operation of the transactional system. There is nothing inherent in the blockchain architecture or mode of operation that influences the procedures and processes of

records creation-the main determinant of whether records will be accurate and reliable. Thus, it is quite possible, even when records are recorded as hashes on chain, for erroneous and unauthorized entries that have been entered into an upstream operational system, such as a land registration system, to be entered into the blockchain.

The answer is different, however, if the upstream creation of the content to be recorded on chain is generated by way of a smart contract. In this case, there is a relationship between the data being passed among the web-front end, an off-chain data store, and the blockchain. Given the uncertainty that remains around how reliably smart contract code represents the intent of smart contract creators, as illustrated by the DAO exploit [25], as well as the uncertainty associated with interoperability among different decentralized system components, there is currently a higher probability that accuracy and reliability will be affected negatively than positively.

From a technical standpoint, whether records are generated off- or on-chain, inconsistencies between nodes within a single blockchain, different blockchains in a system using more than one blockchain, or various components of a decentralized blockchain system, can lead to errors or inconsistencies that affect accuracy. Within a single blockchain, each individual block contains a list of transactions and a timestamp representing the approximate time the block was created, among other additional information. In some systems, the block timestamps allow the system to regulate the production of tokens (e.g., underlying cryptocurrencies) used to generate proof of the chronological order of the transactions. In the context of a land registry system, timestamping is required by the creator and a legal-administrative system for the record to be capable of generating consequences and will support determination of an authentic land registry record versus an inauthentic one. Nodes usually calculate the timestamp based on the median time of a node's peers, which is sent in the version message as nodes connect [26]. Given the reliance of Blockchain technology upon timestamps, it is extremely important that the counters of all the nodes that keep track of the network time be working properly in order to prevent timestamp errors. If this is not the case, the timestamp will be inaccurate. In addition, even when the counters are working properly, it is possible for an attacker to slow down or speed up a node's network time counter by connecting as multiple peer nodes and reporting inaccurate timestamps [26]. In certain types of blockchains (e.g., private, permissioned), it is more likely that a single party may gain control over a large number or majority of nodes, which increases the probability of such an attack. In blockchain-based recordkeeping systems operating multiple blockchains (e.g., a private side-chain and a public blockchain) there may arise inconsistencies, even if temporarily, by virtue of differential rates of consensus formation [27]. Finally, given the decentralized nature of such systems, failures in communication between different, distributed components of the system also could result in inconsistencies in transactional records.

Inaccuracies introduced into records in blockchain recordkeeping systems are not easy to correct. Given the design of such systems, aimed at preserving the immutability of records through time-ordered, cryptographically verified entries, it is not trivial to make changes to recorded transactions. One approach is to introduce “editability” to the blockchain, but this is generally seen as defeating one of the core rationales for using blockchain recordkeeping in the first place [28]. Another approach is to introduce a new transaction that corrects the previous transaction by way of entering a new, updated transactional record. For example, if the name of the new owner of a property was incorrectly recorded, the new transaction could register title by transferring the property from person A with the name misspelled to person A with the name correctly spelled. This is not a perfect solution, however, as subsequent transactions may link back to the incorrect record and, upon correction of the record, could become invalidated [29]. An example might be the subsequent purchase of a new car that uses the property as collateral, referencing the hash of the initial land title registration. By pointing to the

incorrect hash (i.e., the hash of previous, uncorrected document), this subsequent record could be invalidated.

Authenticity. The above-noted problem might be resolved by the instantiation of an archival bond that links records relating to the same transaction/procedural context throughout their lifecycle to one another. At present, however, blockchain-based recordkeeping systems generally lack such functionality. This also impedes the establishment of a unique identification of blockchain based records, since the data content of transactions alone is insufficient to distinguish the purpose of the data and the real world effects that it has been created to generate. This can only be ascertained from an understanding of the source and procedural context of the records.

There are several options currently in use to address the need to link hashed records on the blockchain back to their procedural context. Some blockchain-based recordkeeping systems employ sidechain solutions that link hash records via a unique ID (e.g., [24]). This approach relies upon the continued existence and operation of the sidechain. If the sidechain goes, so does the archival bond between the entries on the distributed ledger.

Some blockchain-based recordkeeping solutions hash all the documents that are part of the logical transaction (i.e., the same action) and place all the hashes into a metadocument, which is then hashed again [30] [31]. The latter hash is then the item that is placed into the blockchain. Depending upon how the document is constructed, it may fail to preserve the unique identity of each transaction record comprising the metadocument. While it is true that the archival bond between the documents is established and preserved in this approach, the hashing of the metadocument transforms the escapsulated hashes into a new document which destroys (since the hash cannot be reverse engineered) the individual identities of the documents within the metadocument that have contributed to the formation of the new hash. As a result, subsequent determination of the identity (and authenticity) of all those documents that contributed to the formation of the metadocument may become impossible. In addition, it is inefficient to have to wait to bundle all logically related transactions together into the metadocument before hashing and anchoring in the blockchain. In real-world recordkeeping, actions often take place in time ordered sequences that can span a considerable amount of time. For example, it can take some time for the sale of property to complete, with key steps in the transaction taking place over months or, in some cases, years. To instantiate and retain the archival bond using the above method may take too much time.

An alternative approach might be to use transaction metadata (e.g. the OP_RETURN field in Bitcoin) to establish an archival bond between transactions in a blockchain [32]. To illustrate how this approach could work on a Bitcoin blockchain, in a manner similar to the addition of a descriptor to a wire transfer, OP_RETURN script opcode could be used to mark a transaction with procedural metadata (e.g., a classificatory code). The primary difficulty with this approach is that the OP-RETURN data does not form part of the Bitcoin transaction per se and thus is not validated in the same way.

Another approach is to introduce a semantic layer within blockchain-based recordkeeping systems that, using ontologies, which provides a mechanism to establish the archival bond, since the entry can be linked by the ontology to the procedural action of which it forms a part in order to establish the record's identity. The reference to the ontology, or ontologies, including version numbers, is combined with the transactional data to generate the hash that is recorded on chain [33].

When technical innovations are introduced, there is also a danger that long-established legal and administrative procedural controls over the creation of records become obsolete or break down. This can introduce uncertainty surrounding the integrity of records, as previously happened with the US Mortgage Electronic Registration Systems (MERS) [34]. In addition, jurisdictional laws may not accept records generated by means of smart contracts or recorded on chain as standing in for the facts they

are about, as would usually be the case with reliable records generated and maintained in more traditional forms of recordkeeping systems. Thus, legal admissibility may be in question. Administrative procedures and legal regulations therefore must align with new blockchain based recordkeeping approaches to ensure that such records are accepted as being trustworthy. For example, in the state of Vermont, a new law establishes blockchain-based records as legally recognized facts, giving them an enforceable authority above and beyond "code as law" [35].

Protection of the integrity of records, or at least determination of whether integrity has been affected, is one of the strengths of blockchain-based recordkeeping solutions. The time-ordered generation of the blocks, together with cryptographic validation of transactions and a decentralized architecture offers increased assurances that the records have remained free from corruption. Yet, even in this regard, blockchain technology is highly dependent upon how vulnerable the system is to faults and security breaches. Issues such as Man-in-the Middle attacks, Sybil attacks, SYN floods, coding errors, timing errors and attacks, and cryptographic key loss are possible sources of blockchain system vulnerability [36].

Governance of blockchain-based systems has been noted as another possible source of weakness that could affect integrity negatively. Researchers have observed a systemic tendency towards centralization, at least in the case of the Bitcoin blockchain miners, and private or permissioned ledgers are controlled by an organization or a consortium of organizations, public or private [37]. Given this, it is crucial to ask how truly decentralized some blockchains really are, and whether concentration of nodes with their combined computing power could allow collusion among nodes, eroding the basis of trust (i.e., decentralization) upon which these networks are built, and allowing manipulation of blockchain entries.

Persistence through time. Long-term preservation of the authenticity of records within blockchain-based ecosystems is also uncertain. Preservation in such systems is premised upon redundancy through decentralization of nodes. Blockchain solutions are volatile, however, and the persistence of entire blockchain networks is not guaranteed. If a blockchain community were to shut down, or if miners moved on to a new fork or system, the specific records preserved on the obsolete fork or system ("orphaned chains") may no longer be preserved and, moreover, there may be no backup archive proving the existence (or execution) of these records. Even where records are preserved, the larger question may be: which version is considered legitimate and authoritative according to specific administrative or legal contexts in which these systems operate?

In the case of solutions that anchor only hashes of original records on the blockchain, the originally hashed records must be archived separately in a form that is unchanged and inviolate to later determine authenticity. The level of organization and investment needed to preserve originals in not inconsiderable, involving the establishment of trusted digital repositories and such additional elements as technical, policy and institutional capacity to ingest records and for archival storage, data management, access, dissemination and migration to new media and forms [38]. All of these functions and investments are beyond the scope of most blockchain solutions, but are, at the same time, critical to the effectiveness of any recordkeeping solution.

V. Conclusion

Understanding the theory and principles underlying trustworthy recordkeeping as articulated in archival science, helps provide a useful framework for the evaluation of blockchain-based recordkeeping systems that purport to provide trusted, immutable records. Using an archival theoretic evaluation framework, it is possible to identify gaps in, or threats to, the accuracy, reliability, and long-term authenticity of such systems. Understanding these weaknesses can point the way to design improvements that address gaps in this innovative new suite of technologies. If not addressed, such

gaps could prevent the successful adoption of blockchain-based recordkeeping solutions. Research and development to identity recordkeeping design options and trade-offs will lead, in the long run, to better technical, and downstream, social outcomes.

REFERENCES

[1] R. Pearce-Moses, ed. 2017. InterPARES Trust Terminology Database. http://arstweb.clayton.edu/interlex/expandedSearch.php?term=authenticity (accessed April 6, 2017).
[2] J. Redman, 2016, September 1, "We've Hit Peak Blockchain Hype, Says New Report," https://news.bitcoin.com/blockchain-hype-peak-newreport/.
[3] L. Coleman. 2017, February 2. "Georgia Expands Project to Secure Land Titles on the Bitcoin Blockchain" Cryptocoin News, https://www.cryptocoinsnews.com/republic-of-georgia-expands-projectto-secure-land-titles-on-the-bitcoin-blockchain/.
[4] Kairos Future. 2017. "The Land Registry in the Blockchain – a testbed," https://chromaway.com/papers/Blockchain_Landregistry_Report_2017.pdf.
[5] G. Keirns. 2017, April 5. "Blockchain Land Registry Tech gets Test in Brazil," CoinDesk, http://www.coindesk.com/blockchain-land-registrytech-gets-test-brazil/.
[6] S. Joyce and G. Roberts. 2016, May 4. "Delaware Introduces Blockchain Initiative," Bloomberg, https://www.bna.com/delaware-introduces-Blockchain-n57982070625/.
[7] L. Duranti. 1995. Reliability and authenticity: the concepts and their implications. Archivaria 39 (1995), 5-10.
[8] C. Rogers. 2015. Diplomatics. In Encyclopedia of Archival Science (Rowman & Littlefield, 2015), 176-179.
[9] L. Duranti, A. Jansen, G. Michetti, C. Mumma, D. Prescott, C. Rogers, and T. Kenneth. 2016. Preservation as a Service for Trust. In Security in the Private Cloud (CRC Press, 2016), 47-72.
[10] L. Duranti and B. Endicott-Popovsky. 2010. Digital records forensics: A new science and academic program for forensic readiness.
In Proceedings of the Conference on Digital Forensics, Security and Law (Association of Digital Forensics, Security and Law, 2010), 109.
[11] R. Pearce-Moses, ed. 2017. Accuracy. In InterPARES Trust Terminology Database.
http://arstweb.clayton.edu/interlex/expandedSearch.php?term=accuracy (accessed April 6, 2017).
[12] R. Pearce-Moses and L-A. Baty. 2005. A glossary of archival and records terminology. Society of American Archivists, 2005).
[13] L. Duranti and C. Rogers. 2012. Trust in digital records: An increasingly cloudy legal area." Computer Law & Security Review 28.5 (2012): 522-531.
[14] R. Pearce-Moses, ed. 2017. Completeness. In InterPARES Trust Terminology Database.
http://arstweb.clayton.edu/interlex/expandedSearch.php?term=completeness (accessed April 6, 2017).
[15] H. MacNeil. 2011. Trust and professional identity: narratives, counternarratives
and lingering ambiguities. Archival Science 11, 3-4 (2011), 175-192.
[16] R. Pearce-Moses, ed. 2017. Authenticity. In InterPARES Trust Terminology Database.
http://arstweb.clayton.edu/interlex/expandedSearch.php?term=authenticity (accessed April 6, 2017).
[17] E. Chabrow. 2013. iSMG Security Report. Http://
http://www.bankinfosecurity.com/interviews/data-integrity-in-era-fakenews-i-3489 (accessed April 6, 2017).
[18] (ICA) International Council on Archives. 2004. ISAAR (CPF). International Standard Archival Authority Record for Corporate Bodies, Persons and Families 2nd. Ed. (ICA, 2004).
[19] InterPARES 2. N.d. Glossary.
Http://interpares.org/ip2/ip2_terminology_db.cfm.
[20] International Organization for Standardization (ISO). 2016. TC 46/SC 11. ISO 15489-1:2016. Information and documentation. Records management. Part 1: General. (2st ed.) (International Organization for Standardization, 2016).
[21] (ICA) International Council on Archives. 2012. Committee on Best Practices and Standards: Progress report for revising and harmonising ICA descriptive standards (ICA, 2012).
[22] M. Jacobs. 2017. "A Proposed Blockchain Reference Architecture," LinkedIn, https://www.linkedin.com/pulse/proposed-blockchainreference-architecture-mike-jacobs
[23] ISO TC307-N38. 2017. United States NB-Contribution-SG-Vocabulary-RA. Unpublished document.
[24] P. Snow., B. Deery, J. Lu, D. Johnston, P. Kirby. 2014. "Factom: Business Processes Secured by Immutable Audit Trails on the Blockchain," http://www.factom.org (accessed 15 November, 2015).
[25] D. Seigel, (2016, June 25). "Understanding the DAO Attack," Coindesk, http://www.coindesk.com/understanding-dao-hack-journalists/.
[26] Culubas (2011), "Timejacking & Bitcoin" accessible at http://culubas.blogspot.com/2011/05/timejacking-bitcoin_802.html (accessed 21 November, 2015).
[27] T. Koeppl and J. Kronnick. 2017. "Blockchain Technology – What's in Store for Canadian Markets?" C.D. Howe Institute, https://www.cdhowe.org/sites/default/files/attachments/research_papers/mixed/Commentary_468_0.pdf.
[28] B. Kelly. 2016, September 30. "The Case Against Editable Blockchains," CoinDesk, http://www.coindesk.com/sorry-accenturebitcoins-un-editable-blockchain-feature-not-flaw/.
[29] Author's own empirical observation.
[30] W. Vaughn, J. Bukowski, R. Shea, C. Allen, P. Storz, J. Nelson. 2016, "Chainpoint – A Scalable Protocol for Anchoring Data in the Blockchain and Generating Blockchain Receipts,"
https://tierion.com/chainpoint.
[31] A. Sanchez de Pedro Crespo and L.I. Cuende Garcia. 2016. "Stampery Blockchain Timestamping Architecture (BTA)," https://s3.amazonaws.com/stampery-cdn/docs/Stampery-BTA-v5-whitepaper.pdf.
[32] Coin Sciences Ltd. 2014. Metadata in the Blockchain: The OP_RETURN Explosion. URL:
https://www.slideshare.net/coinspark/bitcoin-2-and-opreturns-theblockchain-as-tcpip.
[33] V. Lemieux and M. Sporny. 2017. "Preserving the Archival Bond in Distributed Ledgers: A Data Model and Syntax," WWW 2017 Companion , April 3-7, 2017, Perth, Australia.
[34] J.P. Hunt, R. Stanton and N. Wallace. 2011.. The End of Mortgage Securitization? Electronic Registration as a Threat to Bankruptcy Remoteness," unpublished paper, http://works.bepress.com/john_hunt/7/
[35] State of Vermont, United States (2015), Act 51, "An Act Relating to Promoting Economic Development," Chapter A.3.
[36] G. Karame and F. Androulaki. 2016.. Bitcoin and Blockchain Security. Norwood, MA: Artech House.
[37] A. Walch. 2015.. The Bitcoin Blockchain As Financial Market Infrastructure: A Consideration of Operational Risk. N.Y.U. Journal of Legislation & Public Policy, 18(4): 837–893.
[38] ISO/IEC (2012), ISO 16363: 2012– Space data and information transfer systems --Audit and certification of trustworthy digital repositories, ISO, Geneva.
[39] Satoshi Nakamoto, 2008. Bitcoin: A peer-to-peer electronic cash system. URL:
http://www.cryptovest.co.uk/resources/Bitcoin%20paper%20Original.pdf.