Records Management Journal

Trusting records: is Blockchain technology the answer?

Victoria Louise Lemieux,

Article information:

To cite this document:
Victoria Louise Lemieux, (2016) "Trusting records: is Blockchain technology the answer?", Records Management Journal, Vol. 26 Issue: 2, pp.110-139, https://doi.org/10.1108/RMJ-12-2015-0042
Permanent link to this document:
https://doi.org/10.1108/RMJ-12-2015-0042

Downloaded on: 25 January 2018, At: 23:40 (PT)
References: this document contains references to 49 other documents.
To copy this document: permissions@emeraldinsight.com
The fulltext of this document has been downloaded 8676 times since 2016*

Trusting records: is Blockchain technology the answer?

Victoria Louise Lemieux
*School of Library, Archival and Information Studies,
The University of British Columbia, Vancouver, Canada*

Received 4 December 2015
Revised 8 February 2016
16 April 2016
Accepted 16 April 2016

Abstract

Purpose – The purpose of this paper is to explore the value of Blockchain technology as a solution to creating and preserving trustworthy digital records, presenting some of the limitations, risks and opportunities of the approach.

Design/methodology/approach – The methodological approach involves using the requirements embedded in records management and digital preservation standards, specifically ISO 15,489, ARMA's Generally Accepted Recordkeeping Principles, ISO 14,721 and ISO 16,363, as a general evaluative framework for a risk-based assessment of a specific proposed implementation of Blockchain technology for a land registry system in a developing country.

Findings – The results of the analysis suggest that Blockchain technology can be used to address issues associated with information integrity in the present and near term, assuming proper security architecture and infrastructure management controls. It does not, however, guarantee reliability of information in the first place, and would have several limitations as a long-term solution for maintaining trustworthy digital records.

Originality/value – This paper contributes an original analysis of the application of Blockchain technology for recordkeeping.

Keywords Reliability, Authenticity, Risk, Digital preservation, Blockchain, Trusted digital repository

Paper type Conceptual paper

1. Introduction

Ensuring trustworthiness of records is a necessary requirement in a range of different contexts where systems of record provide critical underlying infrastructure necessary to achieve development objectives. This is not only a problem for traditional archives, but also for many organizations that may never have thought of themselves as performing an archival function. This includes organizations responsible for civil registries of births, deaths and marriages, land registries and repositories of financial transactions, to offer but a few examples. In each of these cases, if digital records are insecure or lack integrity, development or organizational objectives may be thwarted. For example, untrustworthy civil registration entries may mean that citizens are unable to prove their identities as a necessary precondition of accessing social protection benefits, or that opportunities for identity fraud emerge that undermine a country's immigration policies and national security. Insecure land registries may create opportunities for corrupt politicians to acquire properties that they are not entitled to by fraudulently entering title transfers. Additionally, such records are often required for long periods that may extend well beyond the life span of a single database system or server. Loss or irretrievability of the records may prevent citizens from making future claims to

Records Management Journal
Vol. 26 No. 2, 2016
pp. 110-139
© Emerald Group Publishing Limited
0956-5698
DOI 10.1108/RMJ-12-2015-0042

citizenship, land, social protection or other entitlements. In such cases, the inability to secure long-term trust in records can lead to a more generalized breakdown in trust in government and throughout society.

One technology that is increasingly being discussed as a solution to system of record problems, such as the need for trusted digital records, is Blockchain technology. Blockchain uses "cryptographic signatures and public keys […] chain-linked to form an unforgeable record of transactions for, say, digital cash (or any ledger record for that matter). Crypto proof replaces the notary" (Levy, 2014). Enthusiasm for the potential of this new technology has been spreading fast, even among professional recordkeepers. In a recent blog post, for example, Cassie Findlay of the Recordkeeping Roundtable writes:

> A decentralised archive utilising the blockchain as a storage mechanism could offer an uncontested space from which records could be accessed. Documents and other sets of data can be validated by the blockchain – even if an application you used to get it there is not working. It is decentralized proof which can't be erased or modified by anyone; competitors, third parties, governments. This is what distinguishes using the blockchain from other forms of data timestamping and authentication […] The technology potentially offers a means for society – or at least groups within society – to keep their own records with some assurance about inviolability and longevity that was not possible before (Findlay, 2015).

Given the spread of Blockchain-based solutions and growing interest in using Blockchain technology as a solution to recordkeeping issues, there is an urgent need for records professionals to gain an understanding of the implications of relying on this technology for the long-term management and preservation of trusted digital records. This paper therefore examines whether Blockchain technology truly is capable of meeting this objective.

2. Methodology and approach

The paper first offers a brief general discussion of the conditions under which one might trust records. This is followed by an overview of the high-level requirements for preservation of trustworthy digital records, drawing upon internationally recognized general records management and digital preservation standards (e.g. ISO 15,489, ISO 14,721 and ISO 16,363). The paper then offers background information about Bitcoin Blockchain technology, followed by consideration of a particular proposed implementation of Blockchain technology for the land registry system of a developing country (Honduras), drawing upon an analysis of publicly available sources of information about the architecture and operation of the proposed solution. The Honduras example has been chosen to give the reader a concrete example of how Blockchain-based recordkeeping solutions might operate in practice and how risks associated with such systems might materialize. The Honduran example is based on a solution that, at the time of writing, is only a proposal, which may or may not be implemented by the Honduran government. Even so, it provides the most detailed publically available information about a Blockchain-based recordkeeping solution to date and thus offers the best current illustration of the issues discussed in this paper. The paper concludes by reflecting upon some of the limitations, risks and opportunities presented by the application of Blockchain technology in recordkeeping, and Appendix 2 links this discussion back to the high-level requirements for trusting records.

3. Trust in records

The Merriam Webster dictionary defines "Trust" as:

> […] assured reliance on the character, ability, strength, or truth of someone or something […] one in which confidence is placed […] a charge or duty imposed in faith or confidence or as a condition of some relationship […] something committed or entrusted to one to be used or cared for in the interest of another (Merriam Webster, 2015).

What it means to trust records, and the conditions needed to achieve such trust, is still an open research question, however. The 2002 RLG-OCLC Report on Trusted Digital Repositories noted:

> Computer scientists worldwide have grappled with definitions and performance measures of trusted military systems for nearly 20 years […] Likewise, the airlines industry has required trustworthy, responsible, and authentic systems. In the last decade, ground-breaking work by archivists in Australia, North America, and Europe has resulted in fundamentally new approaches and tools that specify the nature and performance of accountable record-keeping systems. And in the past few years digital library experts have contributed their experience to a growing body of literature and applications pertaining to the construction and maintenance of secure systems accommodating large quantities of digital resources. But what is meant by the phrases "trusted archives" or "trusted repository"? (RLG-OCLC, 2002, p. 8f).

Much of the discussion about trusted records or systems boils down to two interlinking concepts: reliability and authenticity, and closely related concepts such as identity, integrity and provenance (see, for a summary discussion, Mak, 2014). Reliability, as it relates to records, is defined as the trustworthiness of a record as a statement of fact, based on the competence of its author, its completeness and the controls on its creation (Duranti and Rogers, 2012, p. 525). In diplomatic theory, this is sometimes also referred to as historical truth (Duranti, 1990), and in other contexts it is referred to as validity (Merriam Webster, 2015). This concept is more closely linked to the way in which records are originated, and who originates them than the way in which the records are subsequently maintained. Trust in the truth of the facts in a record stands in contrast to documentary truth. Documentary truth, also referred to as juridical truth (Duranti, 1990), is about the trustworthiness of the record as a record; in other words, its authenticity or the quality of a record in relation to what it purports to be (relating to identity of the record) and that it is free from tampering or corruption (related to integrity of the record) (InterPARES, 2015). This concept is more closely linked to the way in which records are maintained over time (i.e. their chain of custody or preservation).

Based on what is known about the reliability and authenticity of a record, an inference is made about how much to trust the record in question. From this standpoint, determining trust is a matter of making a reasoned risk assessment: if the risk is low enough, it is possible to trust the object or artifact concerned (Yeo, 2013, p. 380). Inferences can be made by humans (consciously or unconsciously), or computationally. Regardless, such inferences generally are said to depend upon assessment or computation of four types of knowledge that derives from information about the provenance of records: reputation, which results from an evaluation of the trustee's (which may be the original creator of the record or a subsequent custodian's) past actions and conduct; performance, which is the relationship between the trustee's present actions and the conduct required to fulfill his or her current responsibilities as specified by the truster; competence, which consists of having the knowledge, skills, talents and traits required to be able to perform a task to any given standard; and confidence, which is an "assurance of expectation" of action and conduct the truster has in the trustee (Duranti and Rogers, 2012, p. 522).

4. Standards for creation and preservation of trustworthy digital records

As noted, trustworthiness of records derives, in large part, from establishing reliability and authenticity.

4.1 Reliability

Reliability of records as a statement of fact begins with the process of record creation: who creates the record and how they go about creating it. Record reliability is mandated in standards for current records management such as ISO 15,489 (ISO, 2001) and ARMA's Generally Accepted Recordkeeping Principles (ARMA International 2013), two of the most widely accepted general international recordkeeping standards for management of current records. For example, ISO 15489 (2001) states:

> A reliable record is one whose contents can be trusted as a full and accurate representation of the transactions, activities or facts to which they attest and can be depended upon in the course of subsequent transactions or activities (ISO, 2001, p. 13) (7.2.3 Reliability).

The standard recognizes that records are likely to be more reliable if they are created in the usual and ordinary course of business; that is, at the time of the transaction or incident to which they relate, or soon afterward, by individuals who have direct knowledge of the facts or by instruments routinely used within the business to conduct the transaction (ISO, 2001, p. 13). The standard further notes the importance of retaining a history of the context of the record to enable future users to judge the reliability of records, even in cases where the record systems in which they are retained have been closed or have undergone significant changes (ISO, 2001, p. 18).

To produce reliable records, records systems must routinely capture all records within the scope of their business activities; organize the records in a way that reflects the business processes of the organization; protect the records from unauthorized alteration or disposition; routinely function as the primary source of information about actions that are documented in the records; and provide ready access to all records and metadata (ISO, 2001, p. 14).

4.2 Authenticity

Authenticity, on the other hand, is reliant upon establishing and preserving the identity and the integrity of a record from its point of creation and thereafter (Rogers, 2015). Like reliability, record authenticity is mandated in standards for current records management such as ISO 15,489 (2001) and ARMA's Generally Accepted Recordkeeping Principles (ARMA International 2013). When digital records are created they are often maintained for a period of time in the systems that have generated them (i.e. "originating systems"). This period may vary according to the purpose of the record, such as in environments that have a well-established records management program with retention schedules, or may simply be linked to the decommissioning of the originating system, especially in organizations where standard record controls are not in place. In the context of originating systems, establishing the identity of records includes processes of registration, wherein records are entered into a register and assigned unique identifiers, and classification wherein records are logically linked to other records related to the same function according to a classification scheme. These identifiers may be embedded within, or logically linked to, the record. Assuring integrity in such systems includes measures such as access control, user verification, audits trails, as well as documentation that demonstrates the normal functioning, regular maintenance and frequency of upgrades of records systems (ISO, 2001; DLM Forum, 2010). These activities are also closely linked to, and specified in, standard IT Security controls (ISO, 2013), indicating that maintaining the security of a system will help ensure the integrity of the data within it.

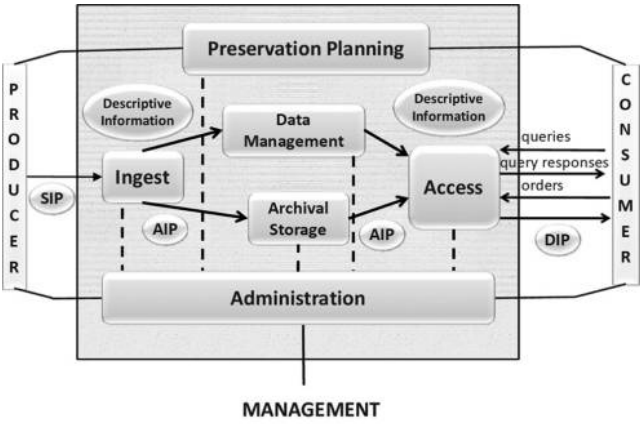
Authenticity also is mandated in frameworks for long-term digital preservation such as OAIS – Open Archival Information Systems (ISO, 2012a) and standards relating to the establishment and certification of Trusted Digital Repositories (ISO, 2012b). The authenticity requirements embedded in these standards are discussed in the section below.

4.3 Long-term digital preservation

If records are of the type that will have continued value to society or are of historical significance, steps need to be taken to ensure their long-term preservation even from the outset of their creation. Standards for long-term preservation of records are relevant to systems of registering land titles because of the long-term value of these records to society. By long-term, it is meant “long enough to be concerned with the impacts of changing technologies, including support for new media and data formats, or with a changing user community. Long Term may extend indefinitely” (ISO, 2012a, p. 18). Long-term preservation of information in digital form requires that technical dangers to the longevity of authentic information be addressed. These include rapid changes to software, hardware, network links to related information and failure to capture or loss of semantic information. That said, the problem of long-term preservation is not just a technical issue. There are also organizational, legal, industrial, scientific and cultural issues to be considered in protecting records over the long-term (ISO, 2012a).

Developed in 2002 by the Consultative Committee for Space Data Systems, the OAIS Reference Model is an approved ISO standard (ISO, 2012a) and has become the *de facto* benchmark for building digital preservation systems. The standard addresses all aspects of long-term preservation of digital information: ingest, archival storage, data management, access, dissemination and migration to new media and forms (see Figure 1).

The OAIS reference model does not dictate the means of implementation, but prescribes requirements to ensure that an OAIS-compliant repository is “an organization of people and systems that has accepted the responsibility to preserve information and make it available for a Designated Community” (ISO, 2012a, p. 30). An OAIS Archive, therefore, is situated in the context of a single, generalized “Designated Community” (e.g. every citizen of a country), or several distinct designated communities with highly specialized needs, each requiring different functionality or support from the repository (ISO, 2012b; Rogers, 2015). The reference model describes the external environment, the functional components or internal mechanisms, which collectively fulfill preservation responsibilities. At a high level, these include:



Source: ISO (2012a)

Figure 1.
OAIS functional entities

- negotiating for and accepting appropriate information from information producers;
- obtaining sufficient control of the information to the level to ensure its long-term preservation;
- determining which is its designated community; and
- therefore to whom the information must be independently understandable and interpretable.

The importance of this function deserves to be underscored: it is not enough to preserve the integrity of bits, or file formats, the semantic meaning for a particular community must be preserved as well. To illustrate, it may be possible to preserve a bit stream (e.g. a sequence of 1's and 0's), and even to preserve the information that renders the bit stream interpretable as a series of, for example, ASCII characters; however, the ability to understand the significance and meaning of the bits will be dependent upon preservation of information that also puts the information in a context interpretable by the knowledge base of a designated community (i.e. enables a subsequent user of the information to understand and interpret the ASCII characters as a file reference number, for example) (ISO, 2012a). OAIS-compliant repositories also must follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, including the demise of the archive and *ad hoc* deletions, and they must make the preserved information available to the designated community and enable the information to be disseminated as copies of, or as traceable to, the original submitted data objects with evidence supporting its authenticity (ISO, 2012a, p. 45).

In performing these functions, an OAIS Archive ultimately is concerned with the preservation of an information object (see Figure 2), which the standard defines as a data object that, if digital, is composed of a set of bit sequences together with representation information (ISO, 2012a, p. 29). Representation information is that information needed to map the bits to higher-level concepts to render the information interpretable and meaningful to a designated community. For example:

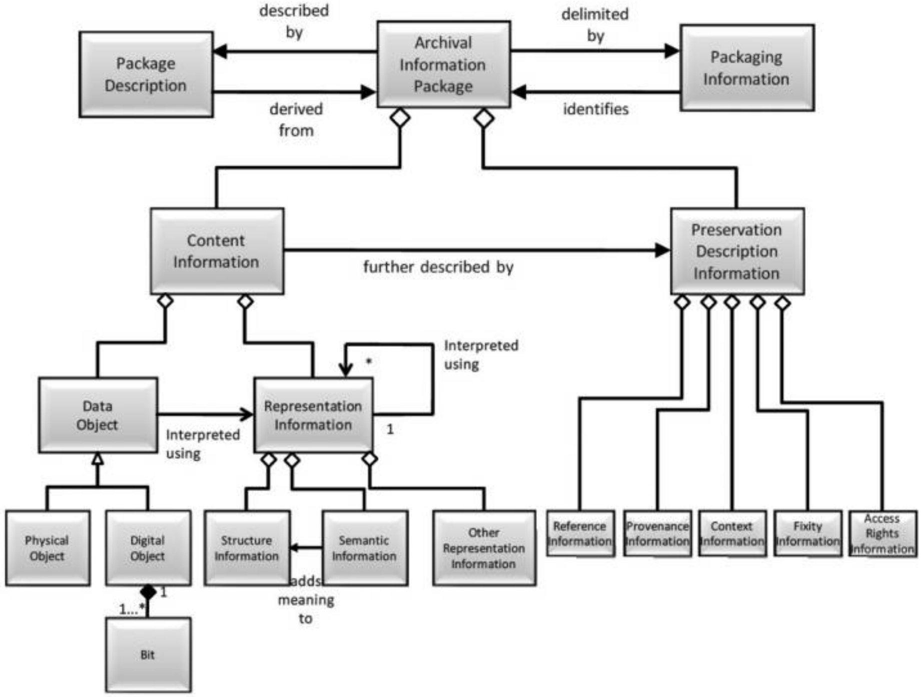


Figure 2.
Composition of information packages

Source: ISO (2012a)

> JPEG software which is used to render a JPEG file; rendering the JPEG file as bits is not very meaningful to humans but the software, which embodies an understanding of the JPEG standard, maps the bits into pixels which can then be rendered as an image for human viewing (ISO, 2012a, p. 31).

In order for an information object to be successfully preserved, it is critical for an OAIS to identify and understand clearly the data object and its associated representation information. For digital information, this means the OAIS must identify the bits and the representation information that applies to those bits (ISO, 2012a, p. 37).

OAIS archives receive information objects as transmissions, called information packages. An information package is a conceptual container of two types of information called content information and preservation description information (PDI) (see Figure 2). The content information and PDI are viewed as being encapsulated and identifiable by the packaging information. The resulting package is viewed as being discoverable by virtue of the descriptive information. The PDI is divided into five types of preserving information called provenance, context, reference, fixity and access rights. These are defined in the OAIS framework as follows:

(1) *Provenance* describes the source of the content information, who has had custody of it since its origination, and its history (including processing history). Some of this information may derive from an originating system, such as a

transaction processing system, as well as from subsequent systems involved in the further processing or storage and retrieval of the content information.

(2) *Context* describes how the content information relates to other information outside the information package. For example, it would describe why the content information was produced, and it may include a description of how it relates to another content information object that is available. This is important for preserving digital records, since records derive their meaning, in part, from their archival bond (referring to the relationship that each archival record has with the other records produced as part of the same transaction or activity and located within the same grouping) and the context of their creation.

(3) *Reference* provides one or more identifiers, or systems of identifiers, by which the Content Information may be uniquely identified. Examples include an ISBN for a book, or a set of attributes that distinguish one instance of content information from another.

(4) *Fixity* provides a wrapper, or protective shield, that protects the content information from undocumented alteration. For example, it may involve a checksum over the content information of a digital information package.

(5) *Access Rights* provide the terms of access, including preservation, distribution and usage of content information. For example, it would contain the statements to grant the OAIS permissions for preservation operations, licensing offers (for distribution), specifications for rights enforcement measures, as well as access control specifications (ISO, 2012a, pp. 33-34).

The components are wrapped together with packaging information. The packaging information is that information which, either actually or logically, binds, identifies and relates the content information and PDI. For example, packaging information could be a file system, such as Documentum™, holding the file content bits, the directory structure holding the pointers to the content bits, the information that is used to distinguish the content bits from the PDI and an encapsulating data structure that identifies the files and other data structures as the components of the information package. The associated archival storage mapping infrastructure might then be implemented as a database that relates the information package ID to the location of the encapsulating data structure (ISO, 2012a, p. 109).

Authenticity of information preserved in an OAIS archive, and "the degree to which a person (or system) regards an object as what it is purported to be", is judged on the basis of evidence (Giarretta *et al.*, 2009, p. 69). Part of the necessary evidence is provided by provenance information, which tells the origin of the source of the content information, documents changes to it and the chain of custody since creation. Authenticity, a stated objective of long-term preservation, is the responsibility of the repository to protect (CCSDS, 2012, pp. 1.9-1.14). The challenge of ensuring records integrity – that is, among other actions, protection from tampering and unauthorized removal or destruction – is achieved in a variety of ways, often including numbered entries in registers, listing file contents and page numbers and restricted permissions on digital systems enforced according to a user's login details. It has always been the case that such methods were not foolproof, and record tampering and removal could take place where there was a strong enough incentive coupled with sufficient technical capability.

ISO 16,363 (2012b) is a standard describing the metrics of an OAIS-compliant digital repository which developed from work done by the OCLC/RLG Programs and the National Archives and Records Administration task force initiative (Giarretta, 2011). This standard sets out the characteristics and functional features required of trusted digital repositories along three dimensions: organizational structure, digital object management and infrastructure and security management. Each of these major dimensions is composed of requirements that a repository must meet to be certified as OAIS-compliant. In terms of organizational structure, the main requirements relate to governance and organizational viability; a suitable organizational structure and staff; a procedural, accountability and preservation framework; financial stability; and appropriate contracts, licenses and liability protections. For digital object management, the requirements are concerned with modes of acquiring or ingesting content; the creation of archival information packages (AIP), consisting of the content information and the associated PDI; preservation planning; AIP preservation; information management; and access management. Finally, infrastructure and security management lay out requirements for technical infrastructure risk management and security risk management.

Together, ISO 14,721 (2012a) and ISO 16,363 (2012b) establish the benchmark for long-term preservation of authentic records within a trusted digital repository.

5. About Bitcoin Blockchain technology

Blockchain technology is a distributed transaction database in which different computers – called nodes – cooperate as a system to store sequences of bits that are encrypted as a single unit or block and then chained together (hence the name Blockchain). The initial and now infamous application of the Blockchain technology is Bitcoin, a form of digital cryptocurrency. The origins of Bitcoin are shadowy. It was first proposed in a 2008 posting attributed to Satoshi Nakamoto. The name turned out to be a pseudonym, and no one has yet been able to uncover its real inventor (Coindesk, 2015). Since its invention, Bitcoin has been linked to criminal activity, experienced wild fluctuations in valuation and been associated with the spectacular failure of Mt. Gox, a Tokyo-based Bitcoin exchange (The Economist, 2015) – hardly the origins of what one would think of as the basis of trustworthy recordkeeping!

Despite its rather shadowy reputation, the technology underlying Bitcoin (i.e. Blockchain) is increasingly seen as a solution to recordkeeping problems where there is a need for a trustworthy public ledger, such as ledgers of financial transactions (as in Bitcoin), land registries (Szabo, 2005), civil registration and other types of public registration systems (Wild *et al.*, 2015), especially in combination with “smart contracts” which automate the rules associated with certain types of contractual transactions (e.g. property purchases) (Johnston, 2014; Cohen 2015). Almgren and Stengard (2015) discuss use of the technology without digital signatures for recordkeeping, and Andrew Miller and colleagues suggest a method for repurposing how Bitcoin operates, in particular its mining function, to achieve distributed storage of archival data (Miller *et al.*, 2014). In the US State of Vermont, legislation has been passed calling for a study on creating a presumption of validity for electronic facts and records that use Blockchain technology (State of Vermont, 2015)[1].

How does the technology work? Bitcoin Blockchain technology essentially establishes a distributed public ledger that contains the payment history of every

Bitcoin in circulation, providing proof of who owns what at any given juncture. This distributed ledger is replicated on thousands of computers – Bitcoin's nodes – around the world and is publicly available (The Economist, 2015, p. 3). For purposes of more easily comparing the operation of native Blockchain to the case study implementation using Factom's proposed solution for the Honduran land registry system, discussion of the specifics of how the technology works will be divided into three parts:

(1) recording transactions;

(2) validating transactions; and

(3) updating a public ledger and authenticating transactions.

An overview of the entire process using the Bitcoin Blockchain is provided below:

- Bitcoin wallet A proposes the transfer of Bitcoin to another wallet B.
- The Bitcoin distributed "mesh" network checks the public ledger that sufficient Bitcoin exists in wallet A.
- If there is sufficient Bitcoin, specialized nodes called miners will bundle the proposal with other reputable transactions to create a new block for the Blockchain.
- 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. See Appendix A.
- This header then becomes the basis for the "proof of work" performed by the miner nodes on the Bitcoin 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.
- Wallet A's payment to Wallet B's transaction, and all the other transactions the block contains, are confirmed.

Box 1. summarizes three critical FAQs for records professionals to understand about the Bitcoin Blockchain.

5.1 Recording transactions

Blockchain technology works by using the Blockchain – made up of an electronic chain of hashes of digital signatures (see Figure 3). Digital signatures are a form of asymmetric cryptography (i.e. they use one private key and one public key) for demonstrating the authenticity of a digital message or documents. A valid digital signature gives a recipient reason to believe that the message was created by a known sender (i.e. it is authentic), that the sender cannot deny having sent the message (i.e. it is non-repudiable) and that the message was not altered in transit (i.e. that it has integrity). Digital

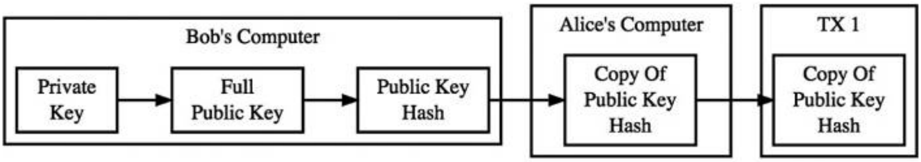
Box 1. Three key FAQs about the Bitcoin Blockchain

(1) Does the Bitcoin Blockchain function as a decentralized archive, storing original records from which records can be accessed? No. Original records are not stored on the Bitcoin Blockchain, only hashes of original records.

(2) Is it possible to reproduce an original record from the hash of the record stored on the Bitcoin Blockchain? No. It is not possible to reverse engineer a hash to reproduce a record.

(3) Does using the Bitcoin Blockchain ensure the trustworthiness of the records? No. Trustworthiness is only guaranteed if the records are both reliable and authentic. Blockchain solutions do not address the reliability of records, and there are many features of the Bitcoin Blockchain that may negatively affect the authenticity of information as well.

Source: Author's own analysis

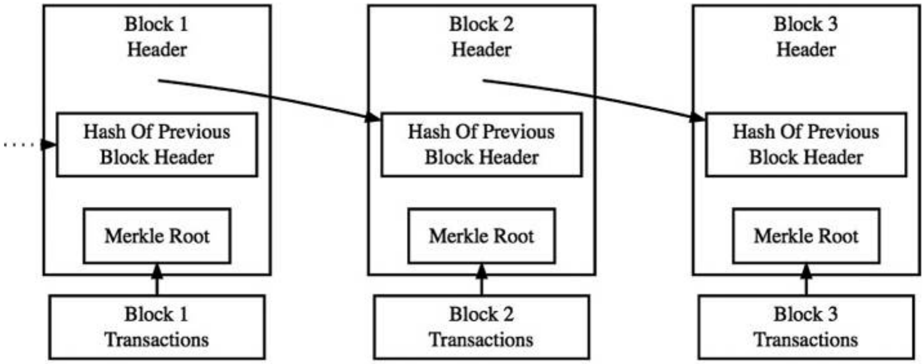


Source: Nakamoto (2009)

Figure 3. Bitcoin digital signature generation

signatures are commonly used for software distribution, financial transactions and in other cases where it is important to detect forgery or tampering.

Each party completes a transaction, for example, sale of a Bitcoin or other asset, by digitally signing (with their private key) a hash of the previous transaction and the public key of the next owner and adding these to the end of the hash chain. The receiving party (e.g. a payee) can verify the signatures to verify the chain of ownership (Nakamoto, 2009; Bitcoin.org, 2015). See Figure 4. To complete this process, Bitcoin uses the Elliptic



Source: Bitcoin.org (2015)

Figure 4. Simplified blockchain

Curve Digital Signature Algorithm with the secp256k1 curve. Secp256k1 private keys are 256 bits of random data. A copy of that private key data is computationally transformed into a secp256k1 public key, which avoids the need for a central authority (called a Certificate Authority) to generate and hold the public keys as is typical of Public Key Infrastructure cryptography (Bitcoin.org, 2015).

5.2 Validating transactions

To avoid a situation wherein a party could transfer an asset twice (the problem of "double-spending" in Bitcoin terms), the transactions are broadcast out to a distributed network of nodes to agree and approve the order of the transactions. Nodes in the network collect the broadcasts of the transactions into blocks, which are then hashed, and receive a timestamp. As explained:

> The solution we propose begins with a timestamp server. A timestamp server works by taking a hash of a block of items to be timestamped and widely publishing the hash, such as in a newspaper or UseNet post […] The timestamp proves that the data must have existed at the time, obviously, in order to get into the hash. Each timestamp includes the previous timestamp in its hash, forming a chain, with each additional timestamp reinforcing the ones before it (Nakamoto, 2009).

To achieve this, the system uses the Hashcash proof of work function; the Hashcash algorithm requires the following parameters: a service string, a nonce (a random or pseudo-random number issued in an authentication protocol to ensure that old communications cannot be reused in replay attacks) and a counter. In Bitcoin the service string is encoded in the block header data structure, and includes a version field, the hash of the previous block, the root hash of the Merkle tree[2] of all transactions in the block, the current time and the difficulty (Bitcoinwiki, 2015a, 2015b). Proof of work, also called mining in Bitcoin parlance, occurs when a computer in the network scans for a value that when hashed begins with a required number of zero bits.

5.3 Updating a public ledger

When a computer finds the proof, it broadcasts the block to all nodes (see Appendix 1 for an example of the content of a Bitcoin block). Nodes accept the block only if all transactions in it are valid. Nodes express their acceptance of the block by working on creating the next block in the chain, using the hash of the accepted block as the previous hash. Nodes work on a consensus system; that is, together with little coordination. Their behavior is such that they do not need to be identified, can leave and rejoin the network at will, accept the proof-of-work chain as proof of what happened while they were gone and express their acceptance of valid blocks by working on extending them and can reject invalid blocks by refusing to work on them (Nakamoto, 2009). This process ultimately establishes a single, but distributed, agreed history for each transaction (a trusted chain of transactions, or Blockchain) and creates a way for the receiver of an asset to know that the previous owners did not sign any earlier transactions (or double spend) (Nakamoto, 2009). Advocates argue that trust is increased among the parties because there is no possibility for abuse by a node in a dominant position, as there can be when a system relies upon a single trusted third party that may be breached or turned rogue (Wild *et al.*, 2015). According to Bitcoin's inventor, the system is secure as long as honest nodes collectively control more CPU power than any cooperating group of attacker nodes because:

> […] in order to modify a past block, an attacker would have to redo the proof-of-work of the block *and all blocks after it* [emphasis added] and then catch up with and surpass the work of the honest nodes. (Nakamoto, 2009, p. 3)

6. A proposed implementation of Blockchain technology for the Honduras land registry system as an illustrative example of the use of Blockchain technology in recordkeeping

Honduras is a middle- to low-income country facing significant challenges, with more than two thirds of the population living in poverty and five of ten citizens suffering from extreme poverty as of 2012. In rural areas, six of ten live in extreme poverty (World Bank, 2015). Security of land tenure in Honduras is essential for the economy to grow and to reduce poverty levels. The fragile situation of property rights has discouraged investment in the country and created a highly unequal distribution of land, contributed to social instability (violent invasions and disputes) and irrational land use (USAID, 2010). A recent Economist article described an anecdote that captures the difficulties facing Hondurans because of weaknesses in the country's system of land registration:

> When the Honduran police came to evict her in 2009 Mariana Catalina Izaguirre had lived in her lowly house for three decades. Unlike many of her neighbours in Tegucigalpa, the country's capital, she even had an official title to the land on which it stood. But the records at the country's Property Institute showed another person registered as its owner, too – and that person convinced a judge to sign an eviction order. By the time the legal confusion was finally sorted out, Ms Izaguirre's house had been demolished. It is the sort of thing that happens every day in places where land registries are badly kept, mismanaged and/or corrupt – which is to say across much of the world. This lack of secure property rights is an endemic source of insecurity and injustice. It also makes it harder to use a house or a piece of land as collateral, stymying investment and job creation. (The Economist, 2015, p. 1).

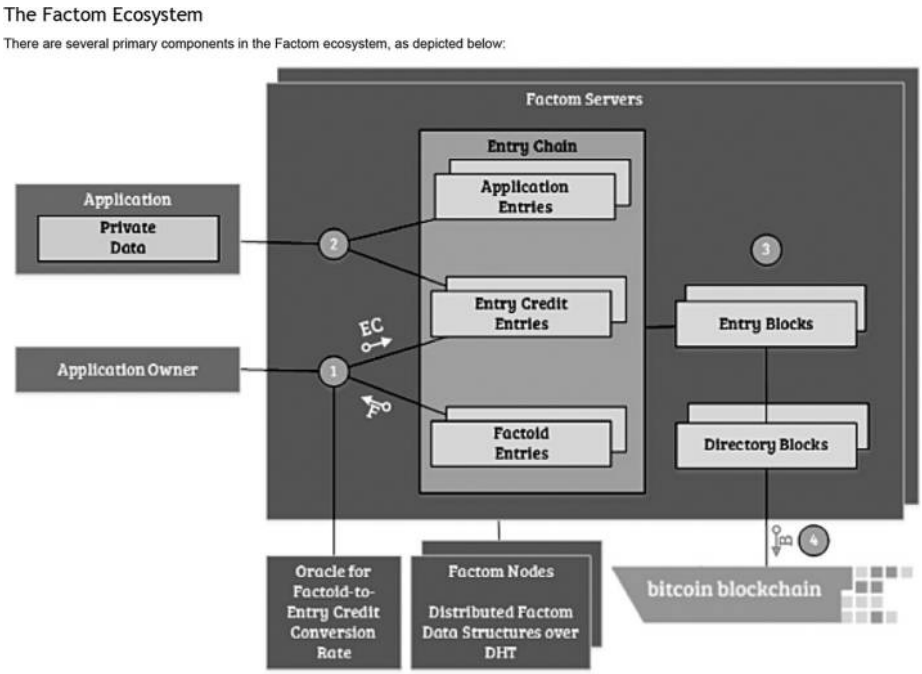
In 2004, Honduras enacted a new law on property, which created a comprehensive legal framework and strengthened land administration in the country (Government of Honduras, 2004). In addition to legal reforms, various institutional reforms were made including establishment of a new national system for property management (SINAP – Sistema Nacional de Administracion de la Propriedad), which provides the technology platform modules for the country's unified registry system (SURE – Sistema Unificado de Registros) (www.sinap.hn). Unfortunately, these systems have been vulnerable to manipulation involving land title fraud. According to one source, "The country's database was basically hacked. So bureaucrats could get in there and they could get themselves beachfront properties" (Chavez-Dreyfuss, 2015). Thus, not only were there concerns about the reliability of the information in the Honduran land registry system, but also about ongoing authenticity of Honduran land registrations.

In November 2015, the *Financial Times* reported that the government of Honduras had approached a company called Factom, a Texas-based start-up, about the use of Blockchain technology to address the recordkeeping problems associated with the Honduran land registration system (Wild *et al.*, 2015). Given the proposal to use Blockchain technology to address recordkeeping issues in Honduras, it is worthwhile – even urgent – to consider how well the Factom Blockchain solution proposed for Honduras might work as a solution to the kinds of recordkeeping problems faced by the country, which are also not atypical of other countries facing similar challenges.

Prefatory to any assessment, however, it is necessary to understand what Factom is and how the Factom implementation is structured and intended to function.

Factom is essentially an open-source Blockchain-based solution layer that sits on top of the Bitcoin Blockchain, and is designed to maintain a permanent, timestamped record in the Bitcoin Blockchain. This is intended to establish a record's – e.g. a record of a land transfer's – proof of existence, proof of process and proof of audit. It has also been designed with the stated objective of addressing three key issues with the way in which the Bitcoin Blockchain typically operates: slow speed to confirmation because of the Bitcoin Blockchain's proof-of-work consensus method of validation over a mesh network of servers, cost of transactions and bloat in Bitcoin Blockchain size (Snow *et al.*, 2014). Factom works with the Bitcoin Blockchain by inserting an "anchor", comprising a hash of a "directory block", into the Bitcoin Blockchain, usually every 10 minutes (see Figure 5). Servers use Bitcoin public keys to post hashes into the Bitcoin Blockchain. However, clients need to independently scan the Bitcoin Blockchain and check relevant transactions for validity.

To advance the discussion, this paper will now examine three aspects of the Factom system that, at a high level, follow the way in which the Bitcoin Blockchain operates and, ultimately, also rely upon the Bitcoin Blockchain:



The Factom Ecosystem

There are several primary components in the Factom ecosystem, as depicted below:

Source: Snow *et al.* (2014)

Figure 5. The Factom ecosystem

(1) recording transactions;

(2) validating transactions; and

(3) updating a public ledger and authenticating transactions.

6.1 Recording transactions

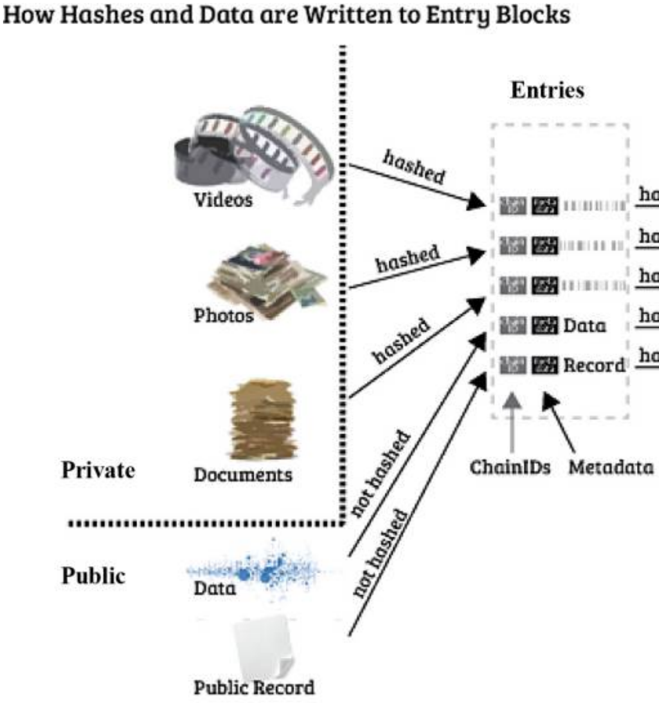
Like the Bitcoin Blockchain, the Factom process of recording transactions, such as land registrations, relies upon a mesh network of federated nodes fulfilling various roles, such as acceptance of entry information from the network of full nodes. These nodes rotate responsibility for different aspects of the system. No node is ever fully in charge, and their decisions and behavior are always in view and change frequently, arguably making them less vulnerable to manipulation and attack (Snow *et al.*, 2014, p. 7).

The process also relies upon a Factom-specific data structure that includes these layers:

- *Entries*: Contains an application's raw data or a hash of its private data.
- *Chains*: Grouping of entries specific to an application.
- *Entry Block Layer*: Organizes references to entries.
- *Directory Layer*: Organizes the Merkle Roots of entry blocks. A directory block is a hash of a "microchain" of "entry ChainIDs" that have been updated during a specified period, usually every minute.

The precise process of recording transactions using this architecture is outlined below and shown in Figure 6:

- All nodes reset their process lists to empty. Process lists are lists of chains that the federated nodes construct from the minute they are responsible for processing. The process list is important for broadcasting decisions made by a node to the rest of the network.
- The user submits an entry payment using a public key associated with entry credits (called "Factoids", which are Factom's own Bitcoin-like brand of cryptocurrency). Users have public keys to sign Factoid transactions, entry commitments, and chain commitments. If they wish to have a vote in the system, they will create an identity and associate public keys holding entry credits with the identity. Factom nodes must have an identity including known Bitcoin keys to post anchors into the Bitcoin Blockchain (Snow et al 2015).
- Based on the public key used to pay for the entry, one of the nodes accepts the payment.
- That node broadcasts the acceptance of the payment.
- The user sees the acceptance and submits the entry.
- Factom organizes entries into chains, linking them via a ChainID. ChainIDs are computed from chain names, which may be derived from natural language with semantic meaning relevant to the recordkeeping context (i.e., a land registration number). This enables faster searching than if all entries were combined together in a single ledger. Actual entries and their ChainIDs are stored separately. There is no dependency of chains with different ChainIDs on one another at the protocol level, meaning that an application recording a chain can ignore other chains, but



Source: Snow *et al.* (2014)

Figure 6.
Factom transaction processing flow

also that Factom does not validate integrity to the same level as the Bitcoin Blockchain since the most recent entry does not embed all previous entries in the digital signature, only those previous entries relating to the specific chain in question. Based on the ChainID of the entry, one of the nodes adds the entry onto its process list, and adds the entry to the appropriate entry block for that ChainID.

- The node broadcasts an entry confirmation, containing the process list index of the entry, the hash of the entry (linked to the payment) and the serial hash so far of the node's process list.
- All the other nodes update their view of the node's process list, validate the list, and update their view of the entry block for that ChainID.
- As long as the user can validate the relevant process list for their entry, then they have a fair level of assurance it will be successfully entered into Factom.
- At the end of the minute, all nodes confirm the process list height, reveal a deterministic private key and the serial hash of the process block (that will match the last item in the process list).
- The directory block for that minute is constructed from all the entry blocks defined by all the nodes. So, each node has all entry blocks, all directory blocks and all entries.

- The collection of deterministic private keys are is combined to create a seed to reallocate ChainIDs among the nodes for the next round. (Snow *et al.*, 2014, p. 8).
- Wallet A's payment to Wallet B"s transaction, and all the other transactions the block contains, are confirmed.

6.2 Validating transactions

Factom is very clear, stating in no less than three places in its original Whitepaper, that it does not validate entries. Such validation must be done client-side by users and applications (Snow *et al.*, 2014, pp. 4, 6, 7). A Factom Whitepaper notes that:

> If Factom were used, for example, to record a deed transfer for real estate, Factom would be used simply to record a process occurred. The rules for real estate transfer are very complex. For example, a local jurisdiction may have special requirements for property if the buyer is a foreigner, farmer or part-time resident. A property might also fall into a number of categories based on location, price or architecture. Each category could have its own rules reflecting the validation process for smart contracts. In this example, a cryptographic signature alone is insufficient to fully verify the validity of a transfer of ownership. Factom then is used to record the process occurred rather than validate transfers (Snow *et al.*, 2014, p. 4).

Thus reliability in the sense of ISO 15,489 is *ultra vires* the Factom solution. The solution is only focused on ensuring documentary reliability; that is, authenticity downstream of the user and land registration application validation of the transaction.

This downstream validation is achieved via the operation of two different types of nodes: Factom nodes and Auditing nodes. Factom nodes accept entries, assemble them into blocks and fix their order. After 10 minutes, the entry order is made irreversible through insertion as an anchor into the Bitcoin Blockchain. The auditing of entries is a separate process that can be done in two different ways according to Factom. In a "with trust" approach, a "thin-client" (a node that does not store a complete copy of every block in the entire block chain) chooses an Audit node it trusts. Audit nodes submit their own cryptographically signed entry, and in so doing validate that the entry passed all checks and conforms to standard. "Trustless auditing", on the other hand, operates in a similar manner to the Bitcoin Blockchain (i.e. computationally via proof-of-work and consensus-based confirmation of entries).

6.3 Updating a public ledger and authenticating transactions.

> Factom creates a consensus system that ensures Entries are quickly recorded, in the order provided, without centralized control, and without requiring public identities of participants […] Factom uses a set of cryptographic protections to ensure the protocol is executed properly. The protocol is also backed up by a user voting system to ensure bad actors are removed from authority where misbehavior cannot be detected in real time from just data within the Factom protocol (Snow *et al.*, 2015, p. 1).

As long as more than half of the participants are honest, and run the protocol as defined, Factom will perform as needed. Performance slows with increased dishonesty among system actors, however. But, as actors fail to perform, they eventually are replaced with new actors, according to Factom.

Records are preserved in two places: the Federated and Audit nodes, which need to maintain this data to make correct decisions about adding new entries. Since they have this data, they can provide it as a service, as part of being a full node. There will also be partial nodes, which share only part of the Factom data set. The partial nodes can share only the data that is relevant to their specific application. The following is an overview of how Factom updates the Bitcoin distributed public ledger:

- Once there are all the entries, and the entry blocks, with each minute separated within the entry blocks for a 10- minute window.
- Build the directory block header, containing the Merkle Root and serial hash of the previous directory block.
- Continue building the directory block with the ChainIDs of all the chains that had entries over the last 10 minutes (sorted by ChainID).
- Compute the serial hash of the directory block, and have the majority of Federated nodes sign it. Append the signatures to the directory block.
- Create a Merkle Root of the directory block and signatures.
- Node #1 for the last minute is delegated to record the Merkle Root into the Bitcoin Blockchain with one of their Bitcoin addresses.
- Clear all the process lists.

Presenting the document at a later time allows one to create its hash, and compare it to the hash recorded in the past. An important point to emphasize is that this process only works if the original record and all the related systems architecture needed to replicate exactly the hash stored on the Blockchain (Factom and Bitcoin) remains available, a point to which this paper will return in the following section. This is because a Blockchain hash cannot be reverse engineered; that is, it is impossible to reproduce the original record by undoing the hash. Instead, to authenticate a document, it must be processed using the same protocols and procedures originally used to process the record stored in the Bitcoin Blockchain and then the two hashes must be compared to ensure that they match.

7. Assessing the limitations, risks and opportunities of the approach

So, is Blockchain technology the answer to trusting records in the case of the Honduran land registry and in general?[3] The answer is mixed. This section provides an overview of some of the limitations, risks and opportunities. The analysis of limitations and risks is summarized in Appendix 2 as well, which also links back to the requirements of the international records management and digital preservation standards discussed in Section 4.

7.1 Who controls the Bitcoin Blockchain?

One of the key questions to be answered about the Bitcoin Blockchain relates to who controls the nodes that are collectively essential to the operation of this technology as a distributed public ledger. Recent reports discuss secret Bitcoin mines operating in China with one mine alone generating 1.5M in Bitcoin per month, and estimated to control 3 per cent of the total Bitcoin distributed network (Franco, 2015). Given this, it is crucial to ask how truly decentralized the Bitcoin public registry really is, and whether concentration of Bitcoin nodes with their combined computing power could allow collusion among nodes and erode the basis of trust upon which Bitcoin is built. In addition, for countries relying on storing elements of their public records on the Bitcoin Blockchain, or any Blockchain not operating within its sovereign jurisdiction, questions arise about where the country's public records really reside, who has access to the country's data and how the distribution of the country's data affects its ability to trust its public records.

7.2 Reliability

As noted above, the Factom solution proposed for Honduras does not address the issue of records reliability at all. This remains completely outside of what the Factom solution offers. Thus, it is difficult to see how the recordkeeping issues faced by Hondurans would be entirely solved by the Factom solution as currently engineered, despite the company's claims. Individual Hondurans must still register their land ownership with the Property Institute, which is responsible – as a trusted third party – to record the title to land in the SURE System. This alone requires a behavioral "nudge" that is well outside the technical architecture of a solution like Factom. It is also still quite possible, even when registrations are recorded, for erroneous and unauthorized entries to be entered into the upstream land registration system well before any records are sent as Entries to a Factum server. Thus, assuring records reliability is one of the major limitations of the proposed solution and, indeed, of any Blockchain-based solution.

7.3 Authenticity

Maintaining authenticity, especially record integrity, is at the heart of what the Factom solution and Blockchain technology has to offer. This is, in essence, the major opportunity that this technology promises to deliver. Yet, even in this regard, the technology's ability to maintain the authenticity of records is highly dependent upon how vulnerable the system is to faults and security breaches. While a full information assurance and security risk assessment is not possible without access to information about the specific system architecture for the Honduran land registry, including risk mitigation measures, and would, in any case, generate several additional papers in its own right, it is possible to highlight in this paper a few key areas of possible system vulnerability based on what is known at a high level about how the system is intended to operate.

7.3.1 Man-in-the-middle attacks. Whenever one system passes information to another system, there exists a possibility for a Man-in-the-Middle Attack (MitMA). MitMA occurs when an attacker secretly intercepts and possibly alters the communication between two parties who believe they are directly communicating with each other. There are two points, in particular, where the proposed Honduran land registry solution may be vulnerable to a MitMA. The first is at the point at which a new land registration entry (an entry) in the SURE system enters the Factom solution, particularly if the transmission is unencrypted. The second is at the point at which the Factom solution anchors the Merkle Root of the signed Factom directory block in the Bitcoin Blockchain. Since Bitcoin miners do not audit these transactions for validity, it is possible to insert invalid transactions designed to look like valid transactions into the Blockchain. The probability of this type of attack is extremely likely in an environment such as exists in Honduras, where system hacking is already occurring, and where the data may pass between systems in unencrypted form.

7.3.2 SYN Flood attacks. A SYN Flood attack is a form of Denial-of-Service attack in which an attacker sends repeated, rapid SYN requests to a target's system in an attempt to consume enough server resources to make the system unresponsive to legitimate traffic. A SYN request is made when a server requests a connection to communicate with another server by sending a SYN (synchronize) message to the server. This is followed by a "handshake" procedure in which the two servers acknowledge one another. In a SYN Flood attack the server receiving the request is unable to complete the handshake procedure before a new request comes in, which ultimately floods the server's resources with requests and causes it to become unresponsive. Although the Bitcoin Blockchain has implemented several measures to prevent denial-of-service attacks, such as SYN Flood attacks (Bitcoinwiki, 2015b), it is still difficult to rule out such attacks, especially in a technology solution that relies heavily on broadcast of communications over a public network.

7.3.3 Sybil attack. A Sybil attack occurs when an attacker fills a Blockchain mesh network with nodes controlled by him, which increases the probability of connecting only to attacker nodes. This type of attack can allow an attacker to refuse to relay blocks and transactions, even disconnecting an entry registration communication from the network. It can also allow an attacker to relay only blocks that he creates (Bitcoinwiki, 2015b). The probability of this type of attack is likely increasing with growing concentration of Bitcoin miners.

7.3.4 Timing errors and attacks. In the Bitcoin Blockchain, each individual block contains a list of transactions and a timestamp representing the approximate time the block was created, among other additional information. The block timestamps allow the system to regulate the production of Bitcoins and generate proof of the chronological order of the transactions as a guard against the double-spending problem. In the context of a land registry system, timestamping enables 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 (Culubas, 2011). 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 to prevent timestamp errors. 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 (Culubas, 2011). Similar to Sybil attacks, growing concentration of Bitcoin miners will increase the probability of this type of attack.

7.3.5 Key management. Key management is essential in a system that relies upon cryptography, such as Blockchain. This includes the generation, exchange, storage, use and replacement of keys, which is difficult to achieve in practice because users must ensure that potentially millions of keys are simultaneously accessible, resistant to digital theft and resilient to loss. The features of effective key management, including system policy, user training, organizational and departmental interactions, and coordination between all of these elements, remain an unresolved problem (Eskandari *et al.*, 2015). The complexity of key management leaves private keys, such as those created to support operation of the Factom and Bitcoin Blockchain systems, vulnerable to loss or open to theft. For example, Bitcoin software manages several private keys by storing them on a node's local storage in a file or database in a pre-configured file system path. A file containing private keys can be read by any application with access to the user's application folder. Attackers may exploit this to gain immediate access to the transaction records. Users must be careful to not inadvertently share their Bitcoin application folder (e.g. through peer-to-peer file sharing networks, off-site backups or on a shared network drive), and must also be cautious about the possibility of physical theft when using portable computers or smartphones. Another threat is loss of keys as a result of general equipment failure due to natural disasters and electrical failures, acts of war or mistaken erasure (e.g. formatting the wrong drive or deleting the wrong folder) (Eskandari *et al.*, 2015).

7.3.6 Audit node attacks in Factom. Factom proposes two methods of transaction verification using Audit nodes. As described above, one method involves using a trusted thin-client to validate entries. Ironically, this approach re-introduces into a Blockchain-based system the very vulnerability that the Bitcoin Blockchain originally was designed to mitigate – reliance upon a trusted third party. The Factom trusted audit approach has the weakness inherent in any system that relies on a single trusted third party, which is that that party may be vulnerable to attack. This could potentially leave the Factom Audit node open to hijacking and manipulation of entries.

7.4 Long-term digital preservation

This paper earlier noted that digital preservation is necessary to consider whenever records have long-term societal or historical value, as is the case with the Honduran land registrations. On the question of long-term preservation and access, Factom has said:

> Nothing is forever. That said […] Factom will hold the data as long as the protocol and Bitcoin are running. Even if Factom went away, the data can be validated as long as the user kept a copy of their data, and has access to the Bitcoin blockchain (Factom.org 2015).

This raises three critical questions:

Q1. How long will the Bitcoin Blockchain survive?

Q2. Who and by what means will they ensure that a copy of the user's data will be preserved?

Q3. Assuming a copy of the user's data is preserved, can it be verified using the data stored on the Bitcoin Blockchain?

None of these questions has a clear answer in the Honduran context. Though it is tempting to think of digital preservation as a legacy issue and thus something that can be dealt with at a later point of time, there is now widespread consensus that, like IT security, digital preservation must be considered early on. This lesson has already been learned the hard way in the USA in relation to use of the Mortgage Electronic Registration System (MERS). Before the financial crisis, MERS was put forward as a solution to a slow paper-based mortgage registration process; however, post-financial crisis, it was found to have contributed to widespread land title resolution issues (Kreiger, 2013).

The remainder of this section explores a number of limitations of the Factom solution and the Bitcoin Blockchain in relation to long-term access and authentication.

7.4.1 Preservation and availability of original records. For long-term validation of entries, the Factom solution relies upon preservation of a copy of the original data entered into the Blockchain, as Factom itself has stated above. This is because Factom, and many other Blockchain-based solutions, work by comparing a digitally signed hash of the original with a digitally signed hash stored on the Bitcoin Blockchain. As the digitally signed version on the Bitcoin Blockchain cannot be reverse engineered to produce a copy of the original record, originals must always be preserved so that they can be re-hashed and digitally signed for purposes of comparison. As discussed in section 4, the level of organization and investment needed to preserve originals in not inconsiderable, involving, according to ISO 14,721 (2012a), the establishment of a Trusted Digital Repository and such additional elements as technical, policy and institutional capacity for ingest of records, archival storage, data management, access, dissemination and migration to new media and forms. All of these functions and investments are beyond the scope of the Factom solution, leaving the question of how the preservation of originals, so necessary for comparison with entries on the Bitcoin Blockchain, is to be handled. It is, moreover, unclear as to what components of the Factom and Bitcoin Blockchain technical architecture must be preserved alongside the original records to render a digitally signed hash capable of comparison with a Blockchain entry for purposes of ascertaining the authenticity of the original. One small change to any of the bits of an original digital record (quite common with bit rot or even as a result of digital preservation processes) or some alteration in the protocols in the recording process, or with their implementation using various components, could make it impossible to authenticate a record at a point in the future. The additional challenge here is that management and preservation of these protocols and components – in part across the Bitcoin Blockchain's extensive distributed network in nodes held in and under the control of other jurisdictions – may be completely outside of the control of the Honduran government, making the proposition of digital preservation a truly challenging one to realize in practice.

7.4.2 Digital signatures. Both the Factom solution and Bitcoin Blockchain technology rely upon the use of digital signatures, which combine a hash message digest with encryption. As encryption relies upon key management, storing private keys for a long period, as required by Factom and the Bitcoin Blockchain, requires that users must also preserve a specification of the file format to ensure the keys can continue to be read (Eskandari *et al.*, 2015). An additional challenge is that, over time, the initially small probability that two different entries will produce the same message digest (called a collision) naturally increases. According to Rosenthal *et al.* (2005):

> Digest algorithms are inherently subject to collisions, in which two different inputs generate the same digest. Digest algorithms are designed to make collisions unlikely, but some of the assumptions underlying these designs do not hold in digital preservation applications. For example, the analysis of the algorithm normally assumes that the input is a random string of bits, which for digital preservation is unlikely.

Finally, and very importantly where the impetus behind the use of Blockchain technology is the prevention of fraud, over time the encryption algorithms used to generate the message digests used in digital signatures can become vulnerable. In the past, widely used algorithms have been broken and obsolescence in the cryptography underlying digital signatures is also an issue (e.g. Estonia, a country that introduced digital signatures as part of its drive toward digital governance, now potentially faces having to make an enormous additional investment to replace outdated and breakable digital signatures; Aas 2014). A digital preservation system that audits against previous message digests must pre-emptively replace its digest algorithm with a new one before the current one is broken or obsolete. To do so, it would need to audit against the current digest to confirm that the item is still good then compute a digest using the replacement algorithm. This could be appended to the stored list of digests for the item (Rosenthal *et al.*, 2005). However, if this process is to be followed, it would necessitate having a trusted third party to carry it out (which defeats the purpose of using a Blockchain approach) and, moreover, raises the likelihood of challenges to the authenticity and integrity of "recomputed" digests as the basis of validating original records. This latter issue would likely have to be resolved in a court of law, which, in turn, requires a legal system capable of making such determinations.

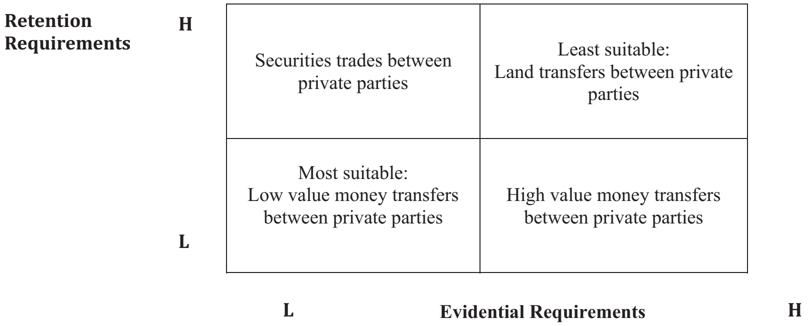
7.4.3 Preservation of link with ChainIDs in Factom. ChainIDs are computed from chain names, which may be derived from natural language with semantic meaning relevant to the recordkeeping context (i.e. a land registration number). Factom does not embed all previous entries in the digital signature, as in the case of the Bitcoin Blockchain, but only those previous entries relating to the specific chain in question. Based on the ChainID of the entry, one of the nodes adds the entry to its process list, and adds the entry to the appropriate entry block for that ChainID. To later establish the authenticity of an original record, it is necessary to link the original record to the appropriate ChainID, which in turn links it to the digitally signed Merkle Root of the directory block that is ultimately anchored in the Bitcoin Blockchain. Unless that link is maintained, either by explicitly annotating the original record with the relevant Factom ChainID or by retention of a key that links the natural language meanings used to construct the name that generates the ChainID, users may have difficulties retrieving the appropriate Bitcoin Blockchain entry against which to test the authenticity of an original entry.

7.5 Monetizing recordkeeping with cryptocurrencies

As noted, operation of the proposed Factom solution for the Honduran land registry relies upon two different crytocurrencies: Factoids and Bitcoin. Original land registrations can only be published to the Factom Blockchain solution by spending Factoids to purchase entry credits, which are then used to publish entries. This is all well and good if the Factom nodes are controlled by the Honduran government and Factoids serve as a type of "monopoly money" used only within the system to facilitate its operation; however, the implications are different if the nodes are controlled by an outside party with the potential to manipulate the Factoid currency-to-entry credit exchange rate. Even without the risk of currency conversion manipulation, instability of currency conversion rates as trading volumes and demand fluctuates creates a currency exposure that the Honduran government may find it difficult to manage and that could threaten the long-term financial sustainability of the solution. A related risk arises from the way in which the proposed Factom solution relies upon anchoring blocks into the Bitcoin Blockchain. To do this, Bitcoin must be purchased and used in the anchoring process. Given wild fluctuations in Bitcoin currency valuation, the potential for currency risk is a real possibility. Since less-well-resourced countries like Honduras already have fragile economies that can be susceptible to external economic shocks, the risk of exposing these economies to higher risks of such shocks through use of cryptocurrencies for recordkeeping is an area that requires additional research and careful consideration.

Despite the above-noted limitations and risks, there is general agreement that Blockchain-based public ledgers – Bitcoin being one example – do work well as mechanisms to validate transactions, such as money transfers or securities trades, at least in the short-term. This is because, as Bitcoin Embassy has argued, Bitcoin avoids the inefficiencies that result from using financial intermediaries to transfer or store assets because any individual is able to transfer Bitcoin to others at low cost, instantaneously and without the need for the usual documentation (Government of Canada, 2015, p. 32). This could be advantageous to individuals in developing countries, for example, where a significant portion of the population has a mobile phone, including individuals with lower incomes, but a large portion of the population remains without access to financial services (Government of Canada, 2015). The recordkeeping requirements associated with money transfers of this nature, for the most part, will be short-term, in contrast to the long-term retention requirements for land titles. In addition, if transactions turn out to be invalid, the loss is relatively minimal and would accrue only to an individual; whereas, in the case of land transfers, the losses could be relatively high and, given the public good nature of land, may have broader societal consequences. Figure 7 presents a heuristic for thinking about where different Blockchain technology use cases may fall along two important dimensions: record retention requirements and evidential requirements (for which the bar is higher when the loss may be borne by the public as well as by a single individual). Use cases wherein the retention requirements are short and the evidential requirements are low may be most suited to use of Blockchain-based solutions, while those with longer retention requirements and higher evidential requirements may be least well-suited. This does not take into consideration that new designs may mitigate some of the risks identified, making even use cases that now appear less well-suited more viable in future.

Plans to roll out Blockchain-based financial transaction processing are moving ahead quickly. In Australia, for example, at time of writing, the Australian Stock Exchange has struck a deal with the US Firm Digital Asset Holdings to use Blockchain technology for the settlement of securities trades (Australian Stock Exchange, 2016). The design of the system to be used will differ from the public Blockchain (i.e. Bitcoin) solution discussed in the Honduran land registry example in that it will rely on a private network in which all parties that participate will be permissioned to do so (Australian Stock Exchange, 2016). This design may point the way to future configurations of Blockchain-based recordkeeping solutions that mitigate some of the risks noted above (e.g. concerns about who controls the Blockchain). The Australian initiative, the Honduran proposal and studies in the US State of Vermont and by the Government of Canada underscore how quickly use of this technology may spread and the urgency with which records professionals need to bring themselves up to speed on how to respond to proposals to use Blockchain technology.



Retention Requirements	L	H
H	Securities trades between private parties	Least suitable: Land transfers between private parties
L	Most suitable: Low value money transfers between private parties	High value money transfers between private parties
	Evidential Requirements	

Source: Authors own analysis

Figure 7. Heuristic for thinking about the suitability of Blockchain solutions for recordkeeping

8. Conclusion

Recently, as this paper has discussed, there has been much interest in – even exuberance about – the potential of Blockchain as a recordkeeping technology. The purpose of this paper has been to try to separate the hype from the reality in a systematic fashion by using international recordkeeping and digital preservation standards as a frame of reference for an assessment of the limitations, risks and opportunities presented by this new technology. The paper examines only how the technology is said to currently function in the context of the Factom solution and the Bitcoin Blockchain, setting aside consideration of novel new functionality such as that proposed by Miller *et al.* (2014), or that may be possible with the introduction other solution layers or risk mitigating features. Without reference to a specific implementation and relying only on publicly available documentation about how the Bitcoin Blockchain and the Factom solution function, the assessment provided in this paper is necessarily neither comprehensive nor definitive. Rather, it serves to concretely illustrate a few potential key areas of concern that require further investigation as a signpost for end users who may need guidance as to whether Blockchain technology is appropriate in their circumstances, researchers interested in further exploring the potential of this new technology as a recordkeeping solution and designers and developers wishing to build recordkeeping solutions using Blockchain technology. It is early days, and the potential of this technology is, as yet, not fully explored. The hope is that this paper takes a small, first step in the direction of building greater understanding.

Far from wishing to discourage use of Factom or any other Blockchain-based solution by discussing possible limitations and risks, this paper aims to stimulate the development of new releases of current offerings that more fully realize the opportunities that this technology may offer to the world of recordkeeping and, more broadly, the many important societal functions that depend upon long-term preservation of trustworthy records, as well as identifying areas of digital recordkeeping and preservation, such as the need for Trusted Digital Repositories for preservation of original records, that Blockchain does not address and that therefore may need to be designed and provided for as enhancements to existing solutions or as completely separate offerings for some use cases. The specific design of such solutions is outside of the scope of this paper, as it requires further research and experimentation. Overall, however, the message is one of caution about the role of Blockchain technology as a comprehensive public recordkeeping and digital preservation solution, even while acknowledging its apparent advantages as a low-cost transaction validation mechanism.

Notes

1. At time of publication, the study had been completed and is available at: http://legislature.vermont.gov/assets/Legislative-Reports/blockchain-technology-report-final.pdf
2. An important feature of the system in terms of saving disk space is that once the latest transaction has enough blocks, the spent transactions before it can be discarded to save disk space. To facilitate this without breaking the block's hash, transactions are hashed in a Merkle Tree with only the root included in the block's hash. Old blocks can then be compacted by stubbing off branches of the tree. The interior hashes do not need to be stored.

3. At time of writing, the government has not signed a contract with Factom and the technology has yet to be implemented. As a result, an assessment of an operational system is not possible. The following assessment therefore is based on facts about the solution derived from technical documentation that the company has made publicly available and other public company communications.

References

Aas, K. (2014), “Case study of digital record keeping in Estonia,”, World Bank Transparency and Information Management Open Discussion Forum, *4* December, available at: http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPUBLICSECTORANDGOVERNANCE/0,contentMDK:23585462~pagePK:148956~piPK:216618~theSitePK:286305,00.html (accessed 21 November 2015).

Almgren, H. and Stengard, M. (2015), “How to maintain authenticity and integrity of electronic information without utilizing electronic certificates?”, INFuture2015: e-Institutions – Openness, Accessibility, and Preservation, Zagreb, *11-13* November.

ARMA International (2013), *Generally-Accepted Recordkeeping Principles*, available at: www.arma.org/docs/bookstore/theprinciplesmaturitymodel.pdf?sfvrsn=2 (accessed 6 February 2016).

Australian Stock Exchange (2016), *ASX Selects Digital Assets Holdings*, available at: www.asx.com.au/asxpdf/20160122/pdf/434j6hyq911404.pdf (accessed 7 February 2016).

Bitcoin.org (2015), “Developer guide”, available at: https://bitcoin.org/en/developer-guide (accessed 21 November 2015).

Bitcoinwiki (2015a), “Block hashing algorithm”, available at: https://en.bitcoin.it/wiki/Block_hashing_algorithm (accessed 21 November 2015).

Bitcoinwiki (2015b), “Weaknesses”, available at: https://en.bitcoin.it/wiki/Weaknesses (accessed 15 November 2015).

CCSDS (2012), “Reference Model for an Open Archival Information System (OAIS): Recommended Practice Issue 2”, Consultative Committee for Space Data Systems, available at: http://public.ccsds.org/publications/archive/650x0m2.pdf (accessed 21 November 2015).

Chavez-Dreyfuss, G. (2015), “Honduras to build land title registry using bitcoin technology”, *Reuters*, available at: http://in.reuters.com/article/2015/05/15/usa-honduras-technology-idINKBN0O01V720150515 (accessed 21 November 2015).

Cohen, B. (2015), “Vermont considering blockchain tech for state records, smart contracts”, *The Coin Telegraph*, available at: http://cointelegraph.com/news/115064/vermont-considering-blockchain-tech-for-state-records-smart-contracts (accessed 21 November 2015).

Coindesk (2015), *Who is Satoshi Nakamot?*, Coindesk, available at: www.coindesk.com/information/who-is-satoshi-nakamoto/ (accessed 3 December 2015).

Culubas (2011), *Timejacking & Bitcoin*, available at: http://culubas.blogspot.com/2011/05/timejacking-bitcoin_802.html (accessed 21 November 2015).

DLM Forum Foundation (2010), *MoReq2010, Modular Requirements for Records Systems*, Publications Office of the European Union, Brussels.

Duranti, L. (1990), “ Diplomatics: new uses for an old science (Part 111)”, *Archivaria*, Vol. 30, pp. 4-20.

Duranti, L. (1997), “”The archival bond”, *Archives and Museum Informatics*, Vol. 11, pp. 213-218.

Duranti, L. (2005), *The Long-Term Preservation of Authentic Electronic Records: Findings of the InterPARES Project*, Archilab, San Miniato.

Duranti, L. and Rogers, C. (2012), "Trust in digital records: An increasingly cloudy legal area", *Computer Law & Security Report* Vol. 28 No. 5, pp. 522-531.

Eskandari, S., Barrera, D., Stobert, E. and Clark, J. (2015), "A first look at the usability of bitcoin key management", USEC 2015, San Diego, CA, available at: www.internetsociety.org/sites/default/files/05_3_3.pdf (accessed 21 November, 2015).

Findlay, C. (2015), "Decentralised and inviolate: the blockchain and its uses for digital archives", *Recordkeeping Roundtable*, available at: http://rkroundtable.org/2015/01/23/decentralised-and-inviolate-the-blockchain-and-its-uses-for-digital-archives/ (accessed on 18 November, 2015).

Franco, E. (2015), "Inside the Chinese Bitcoin Mine That's Grossing $1.5M a Month", *Motherboard*, available at: http://motherboard.vice.com/read/chinas-biggest-secret-bitcoin-mine (accessed 3 December 2015).

Giarretta, D. (2011), *Advanced Digital Preservation*, Springer-Verlag, Berlin-Heidelberg.

Giarretta, D., Matthews, B., Bicarregui, L. and Guercio, M. (2009), "Significant properties, authenticity, provenance, representation information and OAIS", iPres 2009: The Sixth International Conference on Preservation of Digital Objects, San Francisco, CA, pp. 67-73, available at: http://escholarship.org/uc/cdl_ipres09 (accessed 21 November 2015).

Government of Canada Standing Senate Committee on Banking, Trade and Finance (2015), "Digital currency: you can't flip this coin!", available at: www.parl.gc.ca/Content/SEN/Committee/412/banc/rep/rep12jun15-e.pdf (accessed 7 February 2016).

Government of Honduras (2004), "Ley de Propiedad (Property Law)", available at: www.ccit.hn/wp-content/uploads/2013/12/LEY-DE-PROPIEDAD.pdf (accessed 6 February 2016).

International Council on Archives (2007), *International Standard for Describing Functions*, 1st ed., Committee on Best Practices and Standards, Dresden, *2-4* May, International Council on Archives, Paris.

InterPARES (2015), "Terminology database", available at: http://arstweb.clayton.edu/interlex/ (accessed 18 November 2015).

ISO/IEC (2001), *ISO 15489-1:2001 – Information and Documentation – Records Management –Part I: General*, ISO, Geneva.

ISO/IEC (2012a), *ISO 14721: 2012– Space Data and Information Transfer Systems – Open Archival Information System (OAIS) – Reference Model*, ISO, Geneva.

ISO/IEC (2012b), *ISO 16363: 2012– Space Data and Information Transfer Systems –Audit and Certification of Trustworthy Digital Repositories*, ISO, Geneva.

ISO/IEC (2013), *ISO 27001: 2013 – Information Security Management*, ISO, Geneva.

Johnston, T. (2014), "Ethereum contracts as legal contracts", Silicon Valley Ethereum Meeting, available at: http://computationallegalstudies.com/2014/06/ethereum-contracts-legal-contracts/ (accessed 21 November, 2015).

Kreiger, D. (2013), "Wall Street's Mortgage Fraud Scandal 'You can have a house that is fully paid for and still end up in foreclosure'", Deadly Clear, available at: https://deadlyclear.wordpress.com/2013/08/20/wall-streets-mortgage-fraud-scandal-you-can-have-a-house-that-is-fully-paid-for-and-still-end-up-in-foreclosure/ (accessed 22 November 2015).

Levy, J. (2014), "'I love the Blockchain, just not bitcoin'", *CoinDesk*, available at: www.coindesk.com/love-blockchain-just-bitcoin/ (accessed 18 November 2015).

Mak, B. (2014), "Authenticity", in Duranti, L. and Franks, P. (Eds), *Encyclopedia of Archival Science*, Rowman & Littlefield, New York, NY.

Merriam Webster (2015), "Trust", *Merriam Webster Online Dictionary*, available at: www.m-w.com (accessed 18 November, 2015).

Miller, A., Juels, A., Shi, E., Parno, B. and Katz, J. (2014), "Permacoin: repurposing bitcoin work for data preservation", *IEEE Symposium on Security and Privacy*, IEEE Press, New York, NY.

Nakamoto, S. (2009), *Bitcoin: A Peer-to-Peer Electronic Cash System*, available at: https://bitcoin.org/bitcoin.pdf (accessed 21 November 2015).

OCLC/RLG Working Group on Preservation Metadata (2002), *A Metadata Framework to Support the Preservation of Digital Objects*, available at: www.oclc.org/content/dam/research/activities/pmwg/pm_framework.pdf (accessed 18 November 2018).

Rogers, C. (2015), "Virtual authenticity: authenticity of digital records from theory to practice", unpublished PhD dissertation, University of British Columbia, available at: https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/, 1.0166169 (accessed 18 November 2015).

Rosenthal, D.S.H., Robertson, T., Lipkis, T., Reich, V. and Morabito, S. (2005), "Requirements for digital preservation systems", *D-Lib Magazine*, Vol. 11 No. 11, available at: www.dlib.org/dlib/november05/rosenthal/11rosenthal.html (accessed 21 November 2015).

Snow, P., Deery, B., Lu, J., Johnston, D. and Kirby, P. (2014), "Factom: business processes secured by immutable audit trails on the blockchain", available at: www.factom.org (accessed 15 November 2015).

Snow, P., Deery, B., Kirby, P. and Johnston, D. (2015), "Factom ledger by consensus", available at: www.factom.org (accessed 15 November 2015).

State of Vermont, United States Act 51 (2015), *An Act Relating to Promoting Economic Development*, State of Vermont, United States Act 51.

Szabo, N. (2005), "Secure Property Titles with Owner Authority", available at: http://szabo.best.vwh.net/securetitle.html (accessed 21 November, 2015).

The Economist (2015), "The Great Chain of being sure about things; Blockchains", 31 October, available at: http://ezproxy.library.ubc.ca/login?url=http://search.proquest.com/docview/1728728735?accountid=14656 (accessed 18 November 2015).

USAID (2010), *USAID Country Profile: Property Rights and Resource Governance – Honduras*, available at: www.usaidlandtenure.net/sites/default/files/country-profiles/full-reports/USAID_Land_Tenure_Honduras_Profile_0.pdf (accessed 6 February, 2016).

Wild, J., Arnold, M. and Stafford, P. (2015), "Technology: banks seek the key to blockchain", *Financial Times*, available at: http://on.ft.com/1NiyWWs (accessed 2 November, 2015).

World Bank (2015), *Honduras – Overview*, available at: www.worldbank.org/en/country/honduras/overview (accessed 21 November 2015).

Yeo, G. (2013), "Trust and context in cyberspace", *Archives and Records* Vol. 34 No. 2, pp. 214-234.

Appendix 1

Sample block in the Bitcoin Blockchain

```
{"hash":"00000000000042ac033ac4b56e7783d28aa04c14fef6c09a1ea4a3fc5eb823f3","confirmations":257903,"size":6826,"height":126003,"version":1,"merkleroot":"183c22d234724b46d60b1e589b015acd32f64a74823d2100a02b54aa7355e729","tx":["a2993f7df48195a65877a62531dacd541d717918a435cd88e8f981b5c5a48853","35f3e201d368bfb7a58bc80c2c85880bf3b6de128810f3d17974d87591efc04c","783fbaf06cac02e1ae4f56324debcfddf643e003a227b9c1e921352f912baf7c","39dc162c6ae4d3b7f86109256cd4344c00a7e41578564fb7b05a2e777687c81e","a2244b86e621937c4951c96df8c2d8fe376a68fd19eda2cda872f56a1f2e0a73","60be8e16ba36029c1ce48a149f7a5bdda9ba696feec14f8a623bea81c73a15ae","1509c78bb6d84ebe97792dbea630d6b4687d6a367f028ad6063610eec9723f5d","9b1d7e4cb13cc1e2164feb22d9eecc4aa19207d11ddb401b086b31b5c265c292","9069f3d55c1ed8bae0158834430efefffe73b644eba776f20d98bab587fa3b2c","6ed8a55683adb1ba2ecef9150865b8200f7b1c743d6335d9f63d6386b830ef44","daae5daca5a16b0eddc604970d0a6e6b0d0cd71f81f83dec03dfe024f0dd7d99","00eb66b44935f1a329eeb71ebf57686961f92eff47d1897418ede35f951fd4c0","345721dcb9bff3b8db28725ca11fefbd7219582d6015a308ddda622365b09a05","710577e8f318e7d208b07c5bdf5166695a8226377710ef158484d2b0779df560","66e3937eed4797527b0ad3ff88b33804b33f32875c4bc73ae8267748c6f10316","c7a50aef819da69db7a764717478f5c617b9e4275b5229f3f395b88ec90943f4","5cc9c8cb31cdcc1bd3e919775f277031232d1a10a7dff789dbbd08e5f4f40773","019af03f78625672c1ca70384b1c4f17e07ce5daf6de1d3c9183f3d526219368","419cd0ecaf068daea4970e59dbbf4371e3981c4cb3297f19b46492047a20c1bc","866535a1fd93f78ebac9496e579e8543ba47a9f803e34b22864971699f3ae53f","4eee6122a678e024867f908de1a62bdbd5e125d7b67b74e13a465dbe1ccf19ea"],"time":1306148819,"nonce":3984488284,"bits":"1a44b9f2","difficulty":244112.48777434,"chainwork":"00000000000000000000000000000000000000000000000071383ff047b88614","previousblockhash":"00000000000013697134582ff34f9205bd311dd0c7a92be92e1fa74e79da7e5e","nextblockhash":"0000000000001a132111b6b516c790a124052a056f0f60eeb710397a2a569461","reward":50,"isMainChain":true}
```

Appendix 2

Threat/Vulnerability	Impact	Probability[a]	Risk
Control of Blockchain	Loss of control of nodes that validate transactions, leading to inauthentic records being anchored in the Blockchain	Low-medium, but will increase with concentration of control	Authenticity (ISO, 2001; ARMA International, 2013)
Control of record creation outside of Blockchain solution	Unreliable records may be anchored in the Blockchain	Medium-high	Reliability (ISO, 2001, s. 7.2.3; ARMA International, 2013).
Man-in-the-middle attack	Inauthentic records may be anchored in the Blockchain	High	Authenticity (ISO, 2001; ARMA International, 2013)
SYN Flood attack	Inauthentic records may be anchored in the Blockchain	Medium	Authenticity (ISO, 2001; ARMA International, 2013)
Sybil attack	Inauthentic records may be anchored in the Blockchain	Medium	Authenticity (ISO, 2001; ARMA International, 2013)
Timing errors	Inauthentic records may be anchored in the Blockchain	Medium	Authenticity (ISO, 2001; ARMA International, 2013)
Key management	Inauthentic records may be anchored in the Blockchain	Medium-high	Authenticity (ISO, 2001; ARMA International, 2013; ISO 2012a; ISO, 2012b)
Audit server attack (Factom-specific)	Inauthentic records may be anchored in the Blockchain	Medium	Authenticity (ISO, 2001; ARMA International, 2013)
Preservation of original records (needed to validate all transactions)	Inability to validate transactions because of inability to compare hash of original with hash on Blockchain	High	Authenticity (ISO, 2012a; ISO, 2012b)
Bit rot	Inability to compare hash of original with hash on Blockchain	High	Authenticity (ISO, 2012a)
Change to encryption algorithm or manner in which hash is generated	Inability to compare hash of original with hash on Blockchain	High	Authenticity (ISO, 2012a; ISO, 2012b)
Collision of hashes	Inability to validate transaction based on unique identity of original	Low-medium	Authenticity (ISO, 2012a)
Breaking of encryption code	Inability to ensure the integrity of Blockchain	Low-medium	Authenticity (ISO, 2012a)

Notes: [a] Assessment of probability is based on the illustrative Honduran example and will vary depending on the context of the Blockchain system and upon what risk mitigating measures, if any, have been implemented

Table AI. Summary of limitations and risks

Corresponding author

Victoria Louise Lemieux can be contacted at: vlemieux@mail.ubc.ca