Luka Hribar[1]
(Slovenia)

BLOCKCHAINS AND E-RECORDKEEPING SYSTEMS

ABSTRACT

Purpose: To describe the characteristics and the functionality of blockchains and investigate a possible integration of blockchains into e-recordkeeping systems.

Method/Approach: Analysis of scientific and professional literature including primary and secondary sources, and use of descriptive and comparative methods. Brief presentation of two pilot projects from the field of e-archiving. Identification of critical and unresolved issues.

Results/Discussion: The field of blockchains is developing rapidly and is not comprehensively addressed. The definition of a blockchain is not yet fully established. Excessive early expectations regarding the use of blockchains are already fading. There are three types of blockchain integration in the e-recordkeeping systems: storage of (cryptographic) metadata on the chain; storage of complete records on the chain; storage of records related to the virtualized representation of goods and services. The terminology is still evolving. The standardization process is at an early stage of development, the first international standards have already emerged. Archives and other stakeholders verify approaches with pilot projects.

Conclusions/Findings: Blockchains try to solve the problem of trust with the help of technology. Compared to traditional databases, which in some cases blockchains try to replace, they lack some functionality. Due to the design, which is usually based on multiple autonomous and distributed nodes, blockchain management presents new challenges. Developers are designing complex systems that combine the use of public and private blockchains and classic databases. As with most new technologies, the full extent of possible use and abuse is still unclear. To realize the potential of blockchains, issues of privacy, security, efficiency, scalability, and legal problems will need to be addressed. It is also necessary to check the compliance of solutions with the guidelines, recommendations, and standards for e-recordkeeping systems. Analysis of pilot projects shows that they are maybe not yet fully compliant. The first models appear to help designers decide if/when and what type of blockchain to use for a specific e-recordkeeping problem. The continuation of intensive work in this area would be beneficial. Research on the knowledge and acceptance of blockchain technology by the general and professional public has not yet been fully addressed, especially in relation to issues that go beyond the scope of cryptocurrencies.

Keywords: *e-recordkeeping, e-repository, blockchain, trust*

[1] Luka Hribar, PhD student of Archival Sciences at Alma Mater Europaea – ECM, Maribor, Slovenia, luka.hribar@gmail.com
Luka Hribar manages information and communication technology at the National Gallery of Slovenia. He has been involved in the digitalization of the gallery's art collection and documentation from the very beginning. His field of interests in connection to archival science are blockchain technology and artificial intelligence.

1 INTRODUCTION

Digital information technology poses significant risks that e-records can change in e-recordkeeping systems, both intentionally and unintentionally. Today, public confidence in the credibility of records is based primarily on institutional reputation. At a time when technologies for counterfeiting e-records are becoming increasingly widespread, trust in the authenticity of e-records will have to be increasingly based on organizational, security, and technological measures (Hajtnik, 2019). The essential properties of e-records are described using metadata, which is included in the system as additional information. Metadata is key to ensuring that e-records survive and remain available in the future (Hajtnik & Babič, 2018). Hajtnik (2019) states that the presumption of authenticity must be supported by evidence that the record is in its original form and its essential properties have not been altered or damaged.

Cryptographic methods, as one of the technological measures, best address the problem of integrity and partly solve the problem of authenticity. By implementing a one-way hash function, we can be sure of the integrity of the record. Asymmetric cryptography, which is used in digital signing, at the same time ensures integrity and, moreover, through the mechanisms of the public key infrastructure provides some essential methods of ensuring authenticity. However, most of the metadata generated as a result of these methods is still stored in a similar way to the original e-records. This means that cryptographic (and other) metadata (hash values or so-called fingerprints of documents, digital signatures and time stamps...) is stored within the information system (or added to documents) where e-records themselves are stored. At best, the storage of these metadata is entrusted to a third, authorized, independent person or organization, which by itself can present a weak link in the chain of trust. Researchers have long found that centralized trust is problematic (Barometer, 2017). Despite the use of verified cryptographic methods, altering an e-record (and thereby destroying its integrity and authenticity) is still possible if it is owned by only one person or organization.

Instead of trusting a third party, blockchain uses a mechanism of cryptographic evidence. Any transaction (the exchange of data) is protected by a digital signature and is transferred to *all* nodes. As a rule, there are as many copies of data as there are active nodes in the system, which also means big data redundancy. Figure 1 shows the fundamental difference between the centralized and the distributed system. Such a system also does not have a central weak link. If one (or more) nodes fail, the system will continue to operate. The connections between the nodes are significantly more numerous in a distributed system than in a centralized one, thus increasing the number of possible interactions between nodes. Due to the need to ensure that all copies of the data in each of the nodes are identical, the complexity of the distributed system is significantly greater than the centralized one.

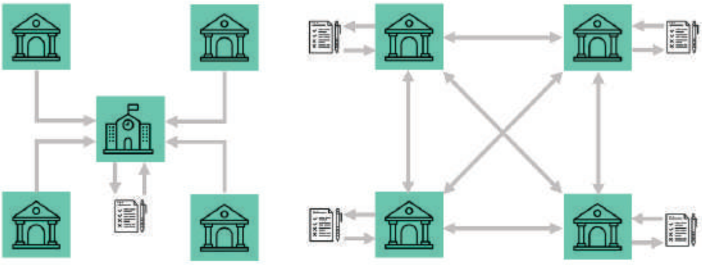

Figure 1: The difference between centralized and distributed system. (Source: Belin, 2020, modified)

1.1 Research limitations

We presume that the greatest limitation of this research is the fact that blockchain technology is rapidly evolving and current findings are quickly becoming obsolete. Although many scientific contributions are being made that focus on individual blockchain elements, comprehensive in-depth discussions on the subject are scarce. One of the problems we faced in preparing this paper is also the still-emerging and maturing terminology. Researchers also mostly focus on the oldest of the blockchains, which is the basis of the Bitcoin and/or Ether cryptocurrency, and its characteristics in discussions. However, development has already produced some new solutions that try to eliminate the perceived weaknesses of the first blockchains. Pilot projects, especially those that have chosen public and established blockchains as their core, are also negatively marked by the speculative nature of cryptocurrencies.

2 BLOCKCHAINS

2.1 Purpose and history of Blockchains

Blockchains' beginnings date back to the end of October 2008 when the author, known under the pseudonym Satoshi Nakamoto – there are also assumptions that it is a group of individuals – published a paper (Nakamoto, 2008) in which he detailed the innovative system of electronic money Bitcoin that operates on the principle of a peer-to-peer network and allows online payments without the need for a broker or central authority. The first block in the chain was created on 3 January 2009, the first transaction between two users took place on 12 January 2009, and the first purchase of goods using this cryptocurrency took place in May of the same year (Skaza, 2020). Many activities in the digital environment are related to the trust of central authority (communication systems, social networks, etc.) that our data is being processed by approved and regulated rules. The most revolutionary novelty of the blockchain is that it does not need a central authority to control or regulate fair cooperation between users. Through various technological mechanisms and incentives, it convinces users to follow the rules and play fair.

2.2 Blockchain definition

According to Lemieux (2016b), there is as of yet no generally agreed upon blockchain definition. Often it is described as a distributed ledger that keeps a growing list of accessible records, which are cryptographically protected against tampering. Walport (2016) states that blockchain is a type of database that combines records into blocks. Each block is chained to the next block using a cryptographic signature. The elements of the chain can be used as a record book, which is shared by everyone with granted rights. By adding new blocks, older blocks are more difficult to change, creating resistance to tampering. Blocks are replicated in copies within the nodes of the network and any disputes about the state of the system are resolved automatically by applying agreed-upon rules. For Vitalik Buterin (2015), author of the widely used Ethereum blockchain, in which in addition to transactions we can also store software code, blockchain is: "a magic computer that anyone can upload programs to and leave the programs to self-execute, where the current and all previous states of every program are always publicly visible, and which carries a very strong cryptoeconomically secured guarantee that programs running on the chain will continue to execute in exactly the same way that the blockchain protocol specifies."

2.3 Properties and operation of Blockchains

Blockchain is technologically similar to a distributed database. Its main purpose is to record digital transactions and belongs therefore to a group of technologies also called *Distributed Ledger Technology* (DLT). Data stored in a blockchain is generally distributed between nodes in complete copies. Nodes can be personal computers, tablets, mobile phones, or even devices of the Internet of Things. The functions of the nodes are to: **verify transactions**; participate in the **construction of new blocks** where transactions are recorded; keep a **copy of the blockchain data** and maintain a **consensus of the blockchain's state**. As a general rule, all nodes perform all of the above tasks. There are also so-called light nodes that do not carry out all tasks and are partly dependent on the full nodes.[2]

The main characteristics of the blockchain are:

- **Distribution.** The system consists of many equivalent nodes that are autonomous in their operation. In this way, the whole is not dependent on a single (crucial) node that could fail.
- **No need for central authority.** As a rule, no node is more important than the other. The state of transactions (block data) is agreed through consensus mechanisms/algorithms.
- **Immutability.** Each block contains hash value (fingerprint) of the contents of the previous block. Any attempt to modify data in previous blocks intentionally or unintentionally is immediately detected. Stored and linked hash values are those elements that *combine data blocks into a chain*.

Several transactions are recorded within each block. Blocks can only be added, editing is not possible or allowed. The header in each block contains hash value of the header of the previous block, a timestamp, some random (called *nonce*) value (if necessary for the purpose of the consent mechanism), and data from the root block, which is of course the only one that has no predecessor since it represents the beginning of the chain. Figure 2 shows a simplified blockchain scheme.

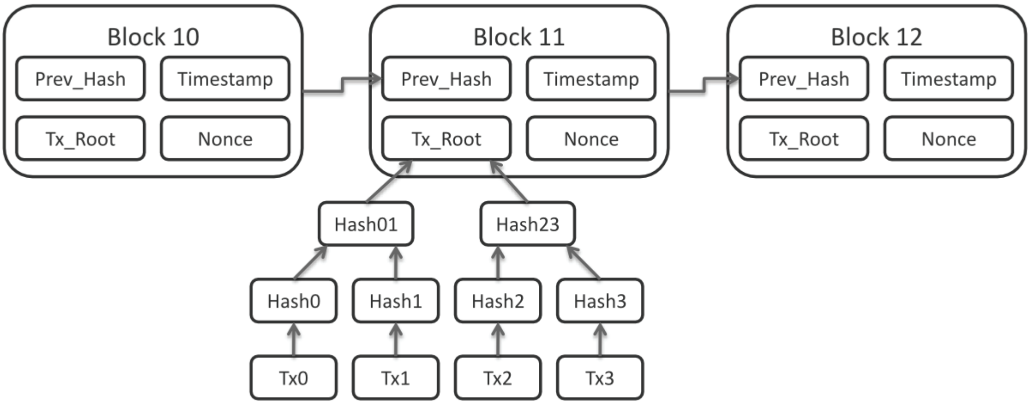

Figure 2: Simplified blockchain scheme. (Source: Bitcoin Block Data, 2020)

[2] Chapter is based on Walport, 2016; Kostanjšek, 2017; Yaga et al, 2018 and complemented by the author of this paper.

A consensus mechanism (there are some variations that differ in their suitability for specific purposes) assures that all transactions in the block are verified and that a new block can be created and linked to the previous block only after the current block has reached the required size or some other criteria (e.g. time interval) is met that is crucial in protecting the credibility and function of the chain. The nodes that perform these processes are sometimes called miners, mints, or publishing nodes, depending on the applicable consensus mechanism. Nodes are typically rewarded for their work, most often in the form of a cryptocurrency, tokens, or commission that the chain charges for transactions.

Blockchain developers and researchers soon realized that the blockchain could also be a code execution environment, not just a collection of records. The code that can be executed by blockchains is often referred to as *smart contracts*. Smart contracts are contracts whose terms are written in computer language instead of legal language (Walport, 2016). The blockchain is an impartial intermediary in a distributed system that executes such code. Smart contracts can be checked and implemented in the same way the digital transactions are being checked. As a rule, any action carried out within the contract will be carried out and checked by all nodes in the network. In this way, we achieve fair implementation of the contracts without the need for third-party trust (Kostanjšek, 2017).

2.4 Areas of application and maturity of technology

Although blockchains have been most frequently used in the field of cryptocurrencies, this is not the only area where attempts have been made to implement the technology. We can quickly find areas where pilot projects are being launched: banking and investment sector, insurance sector, legal services, creative industries (music, film, publishing – mainly in connection with copyright protection), energy and raw materials, transport and logistics, ICT services, distributed storage systems, anti-counterfeiting systems, systems to prove the existence of something at a given moment, etc. Developers focus primarily on industries where the need for trust in a central institution may potentially be a source of problems.

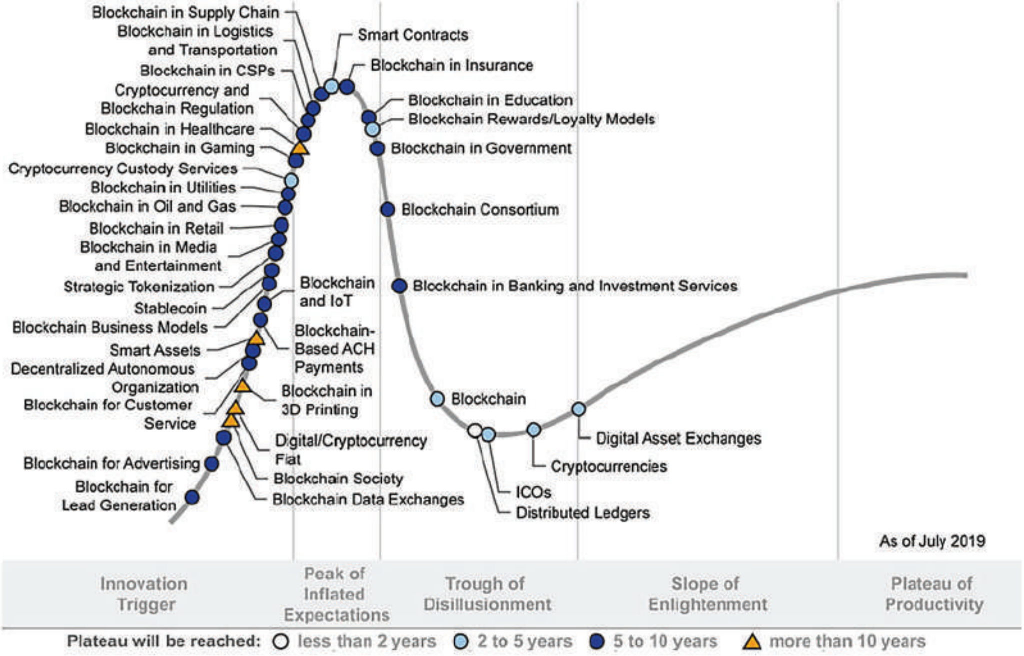

Figure 3: Blockchain hype-cycle. (Source: Gartner, 2019)

Gartner, one of the world's leading research and consulting company, predicts that blockchain technology will have a significant transformational impact on organizations' operations over the next five to ten years. In a survey conducted, 60% of the CIO said that they expected a certain level of acceptance of blockchain technologies over the next three years. At the same time, they are raising concern about the lack of clear principles in technology management and acceptance (Gartner, 2019). In Figure 3, which shows the blockchain hype-cycle, we can see that it is past the peak of inflated expectations and is currently located in the trough of disillusionment area, meaning that the cycle is slowly entering the phase of understanding. The question of whether or not the blockchain will reach the plateau of productivity and to what extent, remains open.

2.5 Blockchain typology

What most distinguishes blockchains from conventional databases is the built-in resistance to erasing or altering stored records (Galiev et al, 2018). Soon after 2009, when the first public blockchain with public permissionless access was created and served as the basis of the Bitcoin cryptocurrency, developers also began developing blockchains whose purpose would not solely be confined to cryptocurrencies. Depending on their characteristics, they can be classified in a few different ways.[3] Most frequently, the discussion around blockchains had been limited to blockchains that understand the **cryptocurrency** transaction as their underlying transaction, however, chains that carry out **token** transactions have soon begun emerging. Tokens represent the right to a service provided by the system in which the blockchain is used. From the perspective of addressing the very principles of the operation of blockchains, this difference is not essential (however, it is very important in connection to taxation if the chain deals with financial transactions) and will not be explained in more detail in this paper.

Classification according to **accessibility and need for identification** is also observed:

- In a **public** blockchain access to the network of chain nodes is available to all interested users.
- In a **private** blockchain access to the network of chain nodes is limited to certain participants.
- In a **permissionless** blockchain there are no restrictions on the identity of transaction participants. Users may be (pseudo) anonymous.
- In a **permissioned** blockchain transactions can only be carried out by identified users.

The chains also differ according to the consent-finding mechanism/algorithm. On a public permissionless blockchain, for example, there are generally many nodes competing for the publication of the next block in which transactions or data will be stored. A key aspect of blockchain is the ability to determine which node will publish the next valid block. In most public permissionless blockchains, nodes that publish a new block (space for new transactions) are rewarded in form of cryptocurrency, tokens or they collect commissions that the network can charge for each transaction. Transaction fees also solve the problem of spamming through unnecessary or uneconomic use and protect against network overloading attacks. Nodes that publish a new block are rewarded for a very clear reason: the possibility of winning rewards encourages nodes to be on-line, to be connected to the network and to verify and validate transactions on the blockchain. This initiative is the key reason public blockchains operate. In private or permissioned chains, this problem is not as pronounced as it is mostly in the interest of known stakeholders that the system

[3] Chapter is based on Okada et al, 2017; Galiev et al, 2018; Lemieux, 2016b; Lemieux et al, 2019 and complemented by the author of this paper.

works. Poorly involved members can be identified and appropriate action taken. The consensus mechanism must therefore solve two fundamental problems: to determine **which node has the task of creating a new block** and **resolve a dispute** that can happen if several **nodes justifiably wish to publish a new block at about the same time.**

Figure 4 in red shows blocks that are not valid. Transactions located in these blocks must be transferred or retransmitted into valid green blocks (longest chain).

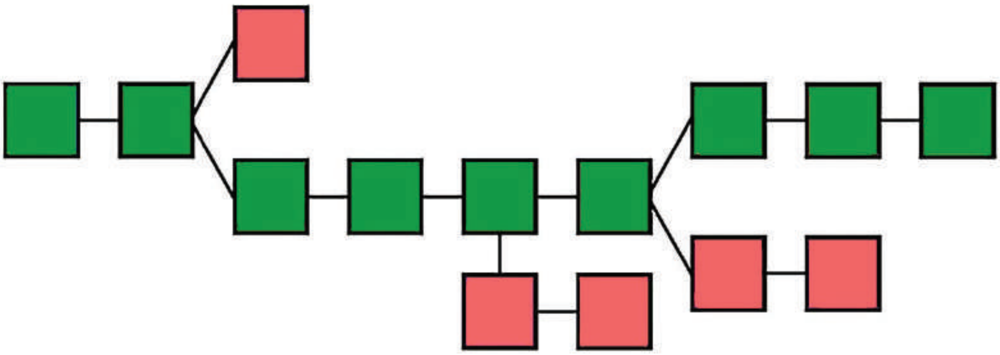

Figure 4: Possible temporary state of a blockchain. (Source: Singh, 2020, adapted)

A dispute that arises in the event of a possible (in each competition two or more participants may achieve exactly the same result without breaking the rules) simultaneous publication of a new block is usually resolved by choosing the chain that most nodes see as the longest. Information about the state of the system can only travel at a finite speed so there is a different perception of the state of the system. The winner in this case is chosen on the basis of geographical distribution or the delay of their network connections (of otherwise matched competitors). Nodes must also be able to re-submit orphan transactions to the next applicable block.

Solving the first of the problems, that is which node has the task of creating a new block, is a fundamentally harder problem. For the most part, all nodes wish to get chosen and would like to publish newer and newer blocks whether or not there is a need for them since they want to collect the prize. As a result, developers have devised ways to prove that the nodes are *worthy* of the task entrusted. The two most common algorithms are *Proof of Work* (PoW), used by the Bitcoin blockchain, and *Proof of Stake* (PoS). The latter is planned to be implemented on the second best-known blockchain, the Ethereum blockchain, as a replacement for PoW.

Proof of work is characterized by the fact that nodes solve a mathematical puzzle that relates to data in the transactions contained in the block (so transactions must be checked first – to put in the necessary work). The first node to solve the puzzle can publish a new empty block and collect the prize. The algorithm also adjusts the difficulty of the puzzle so that the addition of new blocks runs at regular intervals and at least roughly corresponds to the current network needs to store transactions. In this way, the excessive formation of blocks is effectively prevented. The PoW algorithm has proven to be very reliable, however, it also has a problem with energy efficiency. Solving a difficult mathematical puzzle requires a lot of energy. Operations on the Bitcoin blockchain are believed to consume as much energy as a medium-sized country (Vries, 2018). Energy waste is an extremely effective protection against forgery of blocks. A bad actor would have to use an enormous amount of energy to carry out a so-called 51% attack that theoretically could interfere with past transactions in the blocks.

Due to energy inefficiency, developers are looking for ways to replace PoW with an effective but less energy-consuming approach. **Proof of Stake** is based on the idea that the more the user invested in the system, the more likely he would want the system to succeed, and less likely he would want to undermine it. The stake is often in the form of a cryptocurrency or tokens, which the node sends to the system via a specific transaction to a specific address and thus locks the stake until certain conditions are met. The greater the stake the greater the probability that the node will be selected as the issuer of the next block. To keep the system from giving too much advantage to the richest, further approaches have been put in place: a random choice (with a probability weighted according to the stake); voting in several rounds; token ageing systems and delegate systems. Since the PoS model is less energy-consuming than the PoW model and as such spends fewer resources, some blockchains have decided that the reward for creating new blocks can be smaller – in the form of collected transaction fees only. PoS based systems are sometimes designed in such a way that all of the available cryptocurrency or tokens are already distributed among users at the start of the blockchain operation. This approach can also be a weakness since it can lead to allegations of an unjust initial division.

In addition to PoW and PoS models, developers also experiment with *Proof of Authority* (PoA), *Proof of Elapsed Time* (PoET), Round-robin method and others. In private or consortium blockchains none of these proofs have to be implemented. Stakeholders (node owners) can mutually agree on which nodes will take over the task of new block creation. For the most part, an impartial, fast, and computationally efficient model for determining the nodes that issue new blocks, especially in connection to consortium blockchains, tend to be implemented.

If we try to outline the development of blockchain technology chronologically, we can distinguish between three generations (Franks, 2020). The first generation dealt exclusively with financial transactions, whereas the second generation of blockchain technology has also become a runtime environment. For example, Ethereum has introduced the Solidity programming language in which smart contracts are written. The third generation seeks to address interoperability issues and also introduces Blockchain as a Service (BaaS).

All the giants of the IT industry already offer BaaS. Currently, IT systems developers are able to choose between three major platforms:

- Ethereum (public blockchain), which is the most generic platform governed by Ethereum developers;
- Hyperledger Fabric (consortium blockchains), which is modular and governed by the Linux Foundation;
- R3 Corda (consortium blockchains), specialized DLT platform for the financial industry, governed by the enterprise software firm R3.

2.6 Blockchain standards

The standardization process of blockchain technology is at a very early stage of development. The first proposals to initiate the process date back to 2016. First standards are already published. **The International Standards Organisation (ISO)** is an independent, non-governmental international organization that develops international standards. The main technical committee on the blockchain field (TC ISO 307) was established in 2016. It currently has 44 participating members and 13 observers. Two standards have been released so far (ISO/TR 23244: 2020 and ISO/TR 23455: 2019) and further eight are being developed. **The International Telecommunication Union (ITU)** is a specialized agency of the United Nations for Information and Communication Technologies. The agency's core group on the use of distributed ledger technologies (FG DLT, 2017) was established in May 2017. One of the priorities of the group is the creation of an evaluation framework to support efforts to understand the strengths and weaknesses of the DLT. **IEEE** (Institute of Electrical and Electronics Engineers), a technical expert organization for technology advancement, has also launched a number of ongoing initiatives related to the development of standards in connection to the use of blockchain technology (IEEE Blockchain, 2020).

National authorities often prepare their standards in compliance with global standards (ISO, ITU...). The US **National Institute of Standards and Technology (NIST)** so far issued an internal report *NISTR 8202* (Yaga, 2018). The report is a concise technical review, examines, and identifies a possible wider use for blockchain technologies other than cryptocurrency.

In 2017 widespread use of the Ethereum blockchain triggered a flood of *Initial Coin Offering* (ICO) to finance a wide range of projects. The **Ethereum community** quickly grasped the value of interoperability and introduced several of its own standards. The ERC-20 (Ethereum Request for Comment) documents mainly contain the rules for issuing tokens in the Ethereum ecosystem. The Ethereum Enterprise Alliance also operates within the Ethereum ecosystem and aims to develop standards-based open-source specifications that can be trusted and applied globally.

3 DISCUSSION

3.1 Trust as the foundation of society

It was said that trust is a 'social bond' and society could not function without it (Lemieux et al, 2019). Trust essentially means the ability to act without the full knowledge or information required to act – trust fills this gap (Duranti & Rogers 2014). If we ignore different views on the nature of trust, there is a growing global consensus that a crisis of trust exists today (Barometer, 2017). In addition, many people feel decreasingly trusting in centralized authorities in any form (MacNeil, 2011). Decentralization has also proved to be unreliable (Collomosse et al, 2018). Blockchain technology is offered as a solution to the global crisis of trust. However, blockchain technology does not eliminate the need for establishing trust. Instead, it offers a new way to compensate for the lack of information from other sources, in order to extend trust to something or someone and act accordingly. Many believe that consensus mechanisms are a key component of the disruptive potential of blockchain technology – trust is placed in algorithms and the impartiality of technology (Hofman et al, 2019).

3.2 Typology of blockchain application approaches in e-recordkeeping systems

It is important to determine whether the chain itself is solely a storage of records or whether it is part of a larger system (Okada et al, 2017; Lemieux, 2017; Lemieux et al, 2019). By analyzing case studies Lemieux (2017, 2019) identified three emerging typologies of blockchain solutions and characterized them as: mirror type, digital record type, and tokenized type.

Mirror type: With this type, documents are not created or stored on a chain. The blockchain is only in the function of storing cryptographic (and other) metadata of documents (digital fingerprints, digital signatures, etc.). The blockchain serves as a mechanism for confirming the integrity and partly also the authenticity of documents by verifying the equality of cryptographic data associated with the documents and copies of these metadata stored on the blockchain. It can be said that this approach mirrors current good practices to improve the credibility of records.

Digital record type: For this type complete documents are stored on a blockchain, not just metadata. The blockchain must be tailored to this approach. In particular, it must be able to store a much larger amount of data and be able to synchronize all that data between all nodes. With this approach, we should pay great attention to the issues of protecting sensitive data (if the blockchain network is publicly accessible) and the issue of ownership of e-records.

Tokenized type: This is the most innovative type that is characterized by the fact that we store records and tokens on the chain. Tokens often symbolize ownership of assets to which the records relate: e.g. land, real estate, property… With this type, we can also extend the usage of blockchains to products of the financial industry such as futures, derivatives, etc.

It should be noted that the oldest and by far the most widespread blockchain, which is the basis of the Bitcoin cryptocurrency, offers the possibility of including other types of data right from the beginning. This is possible by using the OP_RETURN field in the transaction where 40 bytes of data can be stored (Apodaca, 2017). Although this is not much, a number of projects are using it for storing fingerprints of documents. This is particularly important because it indicated to early developers the possible ways of expansion and development.

3.3 Pilot projects in the field of Archives

Projects that try to use blockchain technology in the private sector are plentiful. CoinGecko (2020), which maintains a database of public blockchains, lists over 7.500 different blockchains covering a wide range of applications. At the national level or in areas of public administrations, the attitude towards this technology is more restrained. But according to Lemieux, et al (2019), almost every country in the world is considering or already using blockchain technology to keep records.

ARCHANGEL is a project that explores the transition from institutional proof of trust to a demonstration of trust by using DLT to ensure the integrity and proof of origin of the digital records entrusted to archives. The project includes the British National Archives, the University of Surrey, and Tim Berners-Lee’s Open Data Institute.

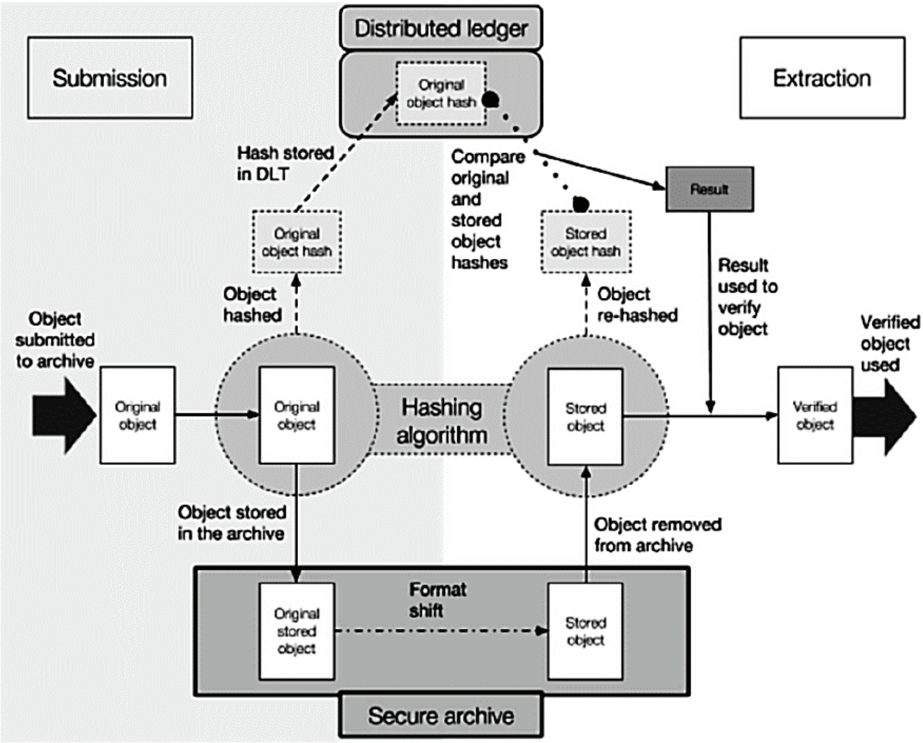

Figure 5: Schematic representation of the Archangel system design. (Source: Colomosse et al, 2018)

Archangel combines the techniques of computer vision and artificial intelligence to obtain fingerprints of documents entrusted to the National Archives. The prototype version uses the Ethereum platform and smart contracts to store the hash values of stored documents (Collomose et al, 2018; Lemieux et al, 2019). Figure 5 shows a schematic representation of the system design.

InterPARES Truster is a project led by researchers at the University of Zagreb that are addressing the issue of the long-term preservation of digital signatures. The problem with the digital signature is that over time the digital certificate used in the signature expires or the certificate issuing body ceases to exist (even time stamping does not completely solve this problem). When this happens, the signature can no longer be completely reliably confirmed. To solve this problem, the research team proposed the TrustChain system for long-term preservation of metadata of digitally signed documents using blockchain technology. Any interested individual or institution may request the addition of a record to the blockchain, but only authorized nodes can enter a new record in the chain (after confirming the validity of the document's digital signature). In addition to cryptographic metadata, document metadata is stored in the system to facilitate queries. The architecture of the system is shown in Figure 6. The TrustChain system cannot extend the lifetime of the digital certificate itself, however, it allows checks to determine whether the signature remained unchanged from the time of entry into the system. That indirectly and practically means that the signature can be trusted (because it was verified when entering the chain). Since the digital signature contains the name of the owner, it can also be used to confirm the creator/provider of the document (Bralić et al, 2017; Lemieux et al, 2019).

Authors of this system are developing an update – TrustChain 2.0 – where they hope to alleviate some of the limitations of TrustChain 1.0. The most obvious limitations of the 1.0 system are, firstly, that it can confirm the validity of digital signatures (or seals) only if they were valid and confirmed at the time they were ingested (problems with expired certificates are already apparent in our daily lives) and, secondly, that validation of digital signatures (or seals) can only be performed by the validation node (problematic with confidential documents) (Bralić et al, 2020).

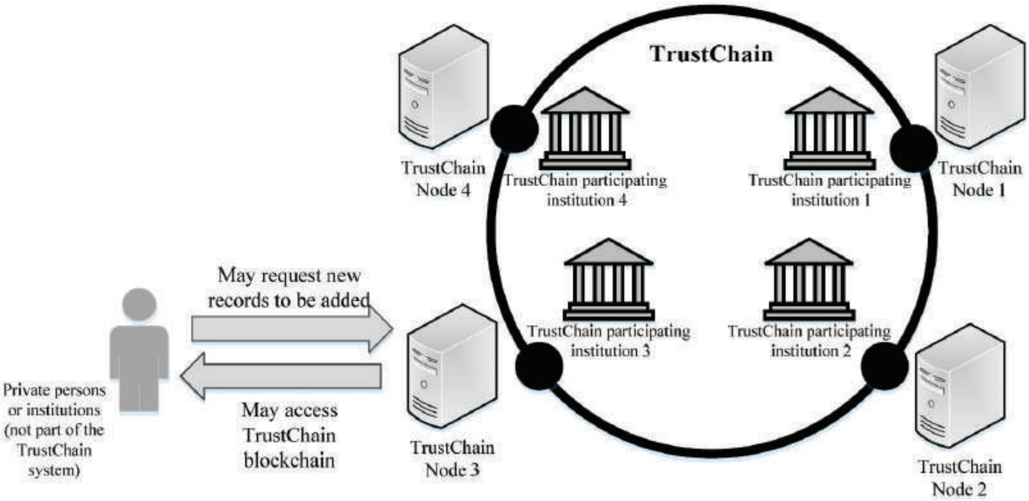
TrustChain
TrustChain Node 4
TrustChain participating institution 4
TrustChain participating institution 1
TrustChain Node 1
TrustChain participating institution 3
TrustChain participating institution 2
May request new records to be added
May access TrustChain blockchain
Private persons or institutions (not part of the TrustChain system)
TrustChain Node 3
TrustChain Node 2

Figure 6: The architecture of the TrustChain system. (Source: Bralić et al, 2017)

3.4 Unresolved Questions

Blockchain features, such as immutableness, distribution, and a lack of need for a central authority, can also be disadvantages when considering their use in relation to e-recordkeeping. In the following, we will summarize and complement some of the observations gathered by Okada et al (2017), Yaga, et al (2018), and Lemieux et al (2019). These considerations apply mainly to public blockchains. For private or consortium blockchains, many of the following questions and concerns are unnecessary. But private chains lose the essential characteristics they are defined by in the element of trust. Collusion is much more likely in private blockchains. Some think (Martinus, 2019) that private chains do not even make sense, because in such cases it is much better to use familiar and verified approaches using traditional databases. However, there are views that many national, regional, and academic organizations will likely choose to prefer private, or at least consortium blockchains where roles and responsibilities are easier to define and control (Bhatia et al, 2020).

The most obvious "weakness" of each blockchain is the inability to correct invalid data. There are always cases where this is justifiably necessary – users make mistakes. As existing transactions on the blockchain cannot be changed, the problem is being addressed by subsequent cancellation transactions. Such Errata needs to be actively monitored, which introduces a new very complex component into the system. A system that includes blockchains and the right to be forgotten (right to data erasure) needs to be carefully planned ahead (Hofman et al, 2019). Data erasure can for instance also be achieved by using smart contracts that can render required records cryptographically inaccessible (Bhatia et al, 2020).

With regard to immutability, it should also be made aware that in systems where entire documents (not just metadata) are stored on the blockchain, they can pose a major threat to users (Matzutt et al, 2018) if they also find unauthorized content; node owners – each with a full copy – could be criminally prosecuted. Theoretically, every miner on a permissionless public Blockchain could be a data controller within the scope of the EU General Data Protecting Regulation (Hofman et al, 2019).

In a distributed system of autonomous users it is very difficult to introduce changes to protocols and to introduce updates. The nodes are autonomous and have to agree to the changes. If there is no consensus, a chain fork (Yaga et al, 2018) occurs, when some nodes insist on old rules and some adopt new ones. If changes are backward compatible, we are talking about *soft fork*, if they are not, a *hard fork occurs*, which results in two functioning but mutually unrelated and incompatible blockchains.

Developers and researchers have not yet answered the question of what happens when all coins or tokens are mined or minted (incentive lost) or transaction commissions suddenly becomes prohibitively high (users can no longer afford transactions). The latter is a frequent occurrence on public permissionless blockchains. Deadcoins.com (2020) lists 1.928 orphan or dead blockchains. How do we reliably archive a blockchain? When blockchain is shut down, we cannot be completely sure of its state anymore.

In systems using blockchains to store hash values, a discrepancy may develop between the fingerprint of the document stored in the chain and the fingerprint of the document kept in the local e-storage system. This may happen because the document in the classical e-recordkeeping system was subsequently changed or amended (perhaps justifiably). What information will users trust? Thinking about this problem also leads us to the complex and unresolved issue of legal validity and ownership of records on the blockchains.

In the case of public blockchains it is often stated that there is no central authority. That statement is not completely accurate. Blockchain developers are connected in strong communities and can make significant changes through technical approaches. Not all technical changes are welcomed by all users, some can seriously impair projects that relies on certain features of a particular blockchain. Developers in a practical sense represent a concrete representation of central authority.

An important authoritarian role is also observed in economically strong node owners (Lemieux, 2016a), who can afford large investments in the form of energy or other resources to control a large portion of active nodes. This is observed when changes on public blockchains need to be implemented.

Another issue in connection to blockchains that is not yet well resolved is the processing rate of transactions. In the most established blockchain, which is the foundation of the Bitcoin cryptocurrency, new blocks are on average created every ten minutes (Median Confirmation Time, 2020). A transaction that is sent into a block, strictly cryptographically speaking, becomes valid only when the block is connected to the next one, so in this case after ten minutes. But many users wait even longer, for multiple blocks, to harden the cryptographic link. Improvements in later implementations of blockchains shortened this time. On the Ethereum blockchain the average time of a new block formation is around twelve seconds (Ethereum Average Block Time Chart, 2020). While it is possible to expect speed improvements in this area, blockchains seem slow compared to traditional databases where transactions are executed in a few milliseconds.

The projects created over the last few years have tried to overcome these weaknesses and limitations through a wide variety of techniques. Developers design systems that combine the use of public and private chains, and classic databases. Blockchains can be interconnected, leading to systems that include side-chains and sub-chains.

4 CONCLUSIONS

Information science experts must closely monitor the development of technologies used to create, manage, and store e-records. Over the past few decades, several such changes have been introduced. Technical innovation, such as blockchains, can trigger significant and long-term changes in business structures and, consequently, in the way in which the economy and society are organized and managed. Rules in the digital world, especially in the area of blockchains, are governed by technology *and* written rules that can be legally assessed. In the case of systems containing elements of distributed ledgers, careful consideration should be given to this complex entanglement. As with most new technologies, the full range of potential uses and abuses is still unclear. It should be made aware that when introducing new information technology we do not immediately perceive all the problems and changes that they create, leading to new professional doubts and ambiguities (Novak, 2009). Before the full potential of blockchains can be realized, issues of privacy, security, efficiency, and scalability will have to be resolved and legal problems addressed.

Any serious implementation of blockchains into the e-recordkeeping system will therefore require compliance with guidelines, recommendations and standards. Lemieux points out (2019) that the first analysis of the designs of different blockchain systems indicates that they do not meet archival standards. Researchers have already noted in 2016 that claims related to the use of blockchain technology to store e-records are in many cases exaggerated (Lemieux, 2016b). Lemieux also emphasizes that there is little awareness in blockchain development communities regarding archival requirements and standards.

A report by the National Institute of Standardization and Technology (Yaga et al, 2018) states that all too often organizations try to adapt the problem so that it could fit in the blockchain technological paradigm, rather than treating blockchains in the same way as any other technological solution available at the moment. Yaga et al (2018) further states that the introduction of blockchains is most meaningful in systems where: there are many participants who do not wish to trust central authority; the nature of the interactions between them is transactional with assets that are limited (money, securities, virtualized representations of physical goods or intellectual property…); an impartial and automated mechanism for resolving ownership disputes is required; there is a need for monitoring real-time transactions and transfer them to permanent storage.

Human society has changed dramatically in recent decades; socially, politically, and economically. These changes are also due to phenomena such as participatory culture, peer-to-peer networks, and trust through computing (Findlay, 2017). The emergence of a technological paradigm such as blockchain is of no surprise. Blockchains are entering information systems of many industries, sometimes complementing existing solutions, sometimes trying to replace them. The first models (Peck 2017; Wüst & Gervais, 2017; ACT-IAC, 2017; Chand, 2018; Hochstein, 2018; ACT-IAC, 2019; Franks, 2020) that help developers of IT solutions to decide when/if and what type of blockchain to use are also emerging. It would be advisable to continue intensive work in order to understand what is at stake in the transformation that is taking place. Moreover, research on the knowledge and acceptance of blockchain technology by the general and professional public has not yet been fully investigated, especially with regard to issues that go beyond cryptocurrency. Institutions that are considering to implement blockchain technology as an element of their e-recordkeeping systems should state their requirements towards DLT developers as early as possible (Bhatia et al, 2020).

Accumulation and dissemination of knowledge is one of the fundamental activities of archival science (Novak, 2010), which is a highly complex, interdisciplinary, and multidisciplinary field (Semlič Rajh et al, 2013). The study of blockchains touches on archival theory, practice, and techniques. In the paper on the study of archival science Klasinc (2011) notes that archivists will not be able to avoid intensive encounters with the theory and practice of information science.

REFERENCE LIST

ACT-IAC. 2017. Enabling Blockchain Innovation in the U.S. Federal Government. American Council for Technology – Industry Advisory Council. (https://www.actiac.org/act-iac-white-paper-enabling-blockchain-innovation-us-federal-government) (23. 5. 2020)

ACT-IAC. 2019. Blockchain Playbook for the U.S. Federal Government. American Council for Technology – Industry Advisory Council. (https://www.actiac.org/act-iac-white-paper-blockchain-playbook-us-federal-government) (23. 5. 2020)

Apodaca, R. L. (2017). OP_RETURN and the Future of Bitcoin. *Bitzuma*. (https://bitzuma.com/posts/op-return-and-the-future-of-bitcoin/) (28. 3. 2020)

Barometer (2017). *Edelman*. (https://www.edelman.com/research/2017-edelman-trust-barometer) (27. 4. 2020)

Bhatia, S., Douglas, E. K., & Most, M. (2020). Blockchain and records management: Disruptive force or new approach? *Records Management Journal*, ahead-of-print. https://doi.org/10.1108/RMJ-08-2019-0040 (1. 12. 2020)

Bralić, V., Kuleš, M. & Stančić, H. (2017). A Model for Long-term Preservation of Digital Signature Validity: TrustChain. *Conference: INFuture2017 –Integrating ICT in Society*, Zagreb. (https://bib.irb.hr/datoteka/906471.TrustChainV11-final.pdf) (26. 4. 2020)

Bralić, V., Stančić, H., & Stengård, M. (2020). A blockchain approach to digital archiving: Digital signature certification chain preservation. *Records Management Journal*, ahead-of-print. https://doi.org/10.1108/RMJ-08-2019-0043 (1. 12. 2020)

Belin, O. (2020). The Difference between Blockchain & Distributed Ledger Technology. *Tradeix*. (https://tradeix.com/distributed-ledger-technology/) (28. 4. 2020)

Bitcoin Block Data (2020). *Wikimedia*. (https://commons.wikimedia.org/wiki/File:Bitcoin_Block_Data.png) (18. 4. 2020)

Buterin, V. (2015). Visions, Part 1: The Value of Blockchain Technology. *Ethereum Blog*. (https://blog.ethereum.org/2015/04/13/visions-part-1-the-value-of-blockchain-technology/) (22. 4. 2020)

Chand, M. (2020). Do You Need A Blockchain. C# Corner, 28. 5. 2020. (https://www.c-sharpcorner.com/article/do-you-need-a-blockchain2) (1. 6. 2020)

Collomosse, J. P., Bui, T., Brown, A., Sheridan, J., Green, A., Bell, M., Fawcett, J., Higgins, J. & Thereaux, O. (2018). *ARCHANGEL: Trusted Archives of Digital Public Documents*. (https://arxiv.org/abs/1804.08342) (22. 4. 2020)

Cryptocurrency Prices & Market Capitalization (2020). *CoinGecko*. (https://www.coingecko.com/en) (3. 5. 2020)

Dead Coins (2020). (https://deadcoins.com/) (6. 5. 2020)

Duranti, L. & Rogers, C. (2014). Trust in Records and Data Online. *Integrity in Government through Records Management: Essays in Honour of Anne Thurston, 2nd ed.*: 203–214. (https://www.researchgate.net/publication/290042093_Trust_in_Online_Records_and_Data) (5. 5. 2020)

Ethereum Average Block Time Chart (2020). *Etherscan*. (https://etherscan.io/chart/blocktime) (25. 5. 2020)

Ethereum Improvement Proposals (EIPs) (2020). *Github*. (https://github.com/ethereum/EIPs) (3. 5. 2020)

FG DLT (2017). Focus Group on Application of Distributed Ledger Technology. *ITU*. (https://www.itu.int/en/ITU-T/focusgroups/dlt/Pages/default.aspx) (3. 5. 2020)

Findlay, C. (2017). Participatory cultures, trust technologies and decentralisation: Innovation opportunities for recordkeeping. *Archives and Manuscripts*, 45(3), 176–190. https://doi.org/10.1080/01576895.2017.1366864

Franks, P. C. (2020). Implications of blockchain distributed ledger technology for records management and information governance programs. *Records Management Journal*, ahead-of-print. https://doi.org/10.1108/RMJ-08-2019-0047

Galiev, A., Ishmukhametov, S., Latypov, R., Prokopyev, N., Stolov, E. & Vlasov, I. (2018). ARCHAIN: A Novel Blockchain Based Archival System. *Second World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*: 84-89. London. (https://arxiv.org/ftp/arxiv/papers/1901/1901.04225.pdf) (17. 4. 2020).

Gartner (2019). *Hype Cycle for Blockchain Business Shows Blockchain Will Have a Transformational Impact across Industries in Five to 10 Years*. (https://www.gartner.com/en/newsroom/press-releases/2019-09-12-gartner-2019-hype-cycle-for-blockchain-business-shows) (28. 4. 2020)

Hajtnik, T. & Škoro Babič, A. (2018). Ali nam lahko pri vrednotenju in odbiranju elektronskega gradiva pomaga tehnologija? *Moderna arhivistika. Časopis arhivske teorije in prakse št. 1*. 169–196. Maribor. (http://www.pokarh-mb.si/uploaded/datoteke/radenci_20181/1_2018_169-196_%C5%A0koro.pdf) (18. 12. 2019)

Hajtnik, T. (2019). E-repozitorij: kdaj bo zaupanja vreden sistem dolgoročne e-hrambe. Škoro Babić, A. (Ed.). *Zbornik prispevkov 7. znanstvene konference: Za človeka gre: Prihodnost zdaj* (53–66). Maribor. Alma Mater Europaea – ECM.

Hochstein, M. (2018). Don't Use a Blockchain Unless You Really Need One. CoinDesk, 16. 1. 2018. (https://www.coindesk.com/dont-use-blockchain-unless-really-need-one/) (23. 4. 2020)

Hofman, D., Lemieux, V. L., Joo, A., & Batista, D. A. (2019). "The margin between the edge of the world and infinite possibility": Blockchain, GDPR and information governance. *Records Management Journal*. (https://doi.org/10.1108/RMJ-12-2018-0045) (1. 12. 2020)

IEEE Blockchain (2020). *IEEE*. (https://blockchain.ieee.org/) (3. 5. 2020)

ISO/TR 23244:2020. *Blockchain and distributed ledger technologies – Privacy and personally identifiable information protection considerations*. (https://www.iso.org/standard/75061.html) (3.5. 2020)

ISO/TR 23455:2019. *Blockchain and distributed ledger technologies – Overview of and interactions between smart contracts in blockchain and distributed ledger technology systems*. (https://www.iso.org/standard/75624.html) (3. 5. 2020)

Klasinc, P. P. (2011). Študij arhivistike kot znanstvene vede. In Tovšak, S. (Ed.) *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja, 10. Zbornik referatov dopolnilnega izobraževanja s področij arhivistike, dokumentalistike in informatike v Radencih*. Pokrajinski arhiv Maribor (73–81). (http://www.pokarh-mb.si/uploaded/datoteke/Radenci/radenci2011/08_klasinc_2011.pdf) (18. 12. 2019)

Kostanjšek, B. (2017). *Uporaba verige blokov za zagotavljanje zaupnosti in integritete podatkov v obstoječih sistemih*. Univerza v Ljubljani, Fakulteta za računalništvo in informatiko. (https://repozitorij.uni-lj.si/Dokument.php?id=102434&lang=slv) (2. 4. 2020)

Lemieux, V. L. (2016a). Trusting records: Is blockchain technology the answer? *Records Management Journal*, 26(2), 110–139. (https://doi.org/10.1108/RMJ-12-2015-0042)

Lemieux, V. (2016b). *Blockchain for Recordkeeping; Help or Hype? Final Report.* (http://blogs.ubc.ca/recordsinthechain/files/2018/06/FinalReport_Volume2.pdf) (15. 4. 2020)

Lemieux, V. (2017). A typology of blockchain recordkeeping solutions and some reflections on their implications for the future of archival preservation. *IEEE International Conference on Big Data*: 2271-2278. Boston, MA. (https://ieeexplore.ieee.org/abstract/document/8258180) (2. 4. 2020)

Lemieux, V., Hofman, D., Batista, D. & Joo, A. (2019). *Blockchain Technology & Recordkeeping*. ARMA International Educational Foundation. (http://armaedfoundation.org/wp-content/uploads/2019/06/AIEF-Research-Paper-Blockchain-Technology-Recordkeeping.pdf) (16. 4. 2020)

Novak, M. (2009). Znanstveno informiranje v arhivistiki. In *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Zbornik referatov z dopolnilnega izobraževanja, Maribor 8/2009* (513–527). (http://www.pokarh-mb.si/uploaded/datoteke/Radenci/radenci2009/50_novak_2009.pdf) (17. 12. 2019)

Novak, M. (2010). Celostno strokovno izobraževanje v okviru slovenske arhivske službe. In *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja 9 (2010)* (483–498). (http://www.pokarh-mb.si/uploaded/datoteke/048novak2010.pdf) (17. 12. 2019)

MacNeil, H. (2011). Trust and professional identity: narratives, counter-narratives and lingering ambiguities. *Archival Science 11* (175–92). (https://doi.org/10.1007/s10502-011-9150-5) (15. 4. 2020).

Martinus, H. (2019). Here's Why You Don't Need Blockchain. *OfferZen*. (https://www.offerzen.com/blog/heres-why-you-dont-need-blockchain) (21. 5. 2020)

Matzutt, R., Hiller, J., Henze, M., Ziegeldorf, J., Müllmann, D., Hohlfeld O. & Wehrle, K. (2018). A Quantitative Analysis of the Impact of Arbitrary Blockchain Content on Bitcoin. *Conference: Financial Cryptography and Data Security*. Curaçao. (https://fc18.ifca.ai/preproceedings/6.pdf) (27. 4. 2020)

Median Confirmation Time (2020). Blockchain.com. (https://www.blockchain.com/charts/median-confirmation-time) (25. 5. 2020)

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. *Bitcoin.org*. (https://bitcoin.org/bitcoin.pdf) (2. 4. 2020)

Okada, H., Yamasaki, S., & Bracamonte, V. (2017). Proposed classification of blockchains based on authority and incentive dimensions. *19th International Conference on Advanced Communication Technology* (593–597). PyeongChang. (https://ieeexplore.ieee.org/document/7890159) (12. 4. 2020)

Peck, M. E. (2017). Do You Need a Blockchain? *IEEE Spectrum: Technology, Engineering, and Science News, IEEE Spectrum*, 29. 9. 2017. (https://spectrum.ieee.org/computing/networks/do-you-need-a-blockchain) (25. 5. 2020)

Semlič Rajh, Z., Šabotić, I. & Šauperl, A. (2013). Znanstveno raziskovalno delo v arhivistiki: Značilnosti uporabe dveh metod. In Fras, I. (Ed.) *Tehnični in vsebinski problemi klasičnega in elektronskega arhiviranja. Zbornik mednarodne konference, Radenci* (125–144). (http://www.pokarh-mb.si/uploaded/datoteke/Radenci/Radenci2013/11_Semlic_Sabotic_Sauperl_2013.pdf) (17. 12. 2019)

Singh, G. (2020). SinghBlockchain Technology: All Set To Revamp the Future of Transactions. *Appinventiv*. (https://appinventiv.com/blog/blockchain-revamping-transactions/) (10. 5. 2020)

Skaza, G. (2019). *Primerjalna analiza perspektivnih blockchain platform*. Univerza v Mariboru, Ekonomsko-poslovna fakulteta. (https://dk.um.si/IzpisGradiva.php?lang=slv&id=73885) (2. 4. 2020)

Vries de, A. (2018). Bitcoin's Growing Energy Problem. *Joule*. Volume 2, Issue 5 (801–805). (https://www.sciencedirect.com/science/article/pii/S2542435118301776) (28. 4. 2020)

Walport, M. (2016). *Distributed Ledger Technology: beyond block chain*. London. Government Office for Science. (https://www.gov.uk/government/publications/distributed-ledger-technology-blackett-review) (18. 4. 2020)

Wüst, K. & Gervais, A. (2017). Do You Need a Blockchain? *IACR ePrint Archive* (https://eprint.iacr.org/2017/375.pdf) (25. 5. 2020)

Yaga, D., Mell, P., Roby, N. & Scarfone, K. (2018). *NISTIR 8202. Blockchain Technology Overview.* U.S. Department of Commerce. National Institute of Standards and Technology. (https://doi.org/10.6028/NIST.IR.8202) (11. 4. 2020)