ArchContract: *using smart contracts for disposition*

Danielle Alves Batista
iSchool
University of British Columbia
Vancouver BC, Canada
danilusasm@gmail.com

Tim Weingaertner
School of Information Technology
Lucerne University of Applied Sciences and Arts
Rotkreuz, Switzerland
tim.weingaertner@hslu.ch

Abstract*— Disposition is one of the consequences of the appraisal archival function. Although appraisal writ large is a function that no technology could execute, disposition has been already supported by different tools. In this paper, we propose a blockchain-based application for disposition, a smart contract called ArchContract, using two different repositories. We discuss appraisal and disposition on blockchain systems, the use of smart contracts as a disposition tool and present the model of ArchContract. We conclude that blockchain and smart contracts have the potential to support some of the records management functions such as disposition.

***Keywords*—Disposition, appraisal, smart contracts, blockchain systems**

I. Introduction

Emerging technologies have long been a concern for records professionals and blockchain systems are no different. The promise of immutability and trustworthiness has inspired some archival science researchers to investigate blockchain's suitability as a recordkeeping system and analyze the records produced and kept in those environments [1], [2], [3]. Technological tools can provide records professional with valuable resources to improve the efficiency and effectiveness of records management practices; what remains to be seen is how best to utilize blockchain as such as a tool.

One of the most important – and challenging – duties of archivists is appraising records, selecting relevant records to be considered evidentiary sources to be preserved for legal rights, culture and history and ensuring the destruction of records with no value to the archivist's organization. Properly done, appraisal, including disposition, improves organizational efficiency. In this paper, we propose a solution using blockchain systems and smart contracts to make the disposition process more secure and efficient. We recognize that the appraisal is a delicate, intellectual procedure that no technology could accomplish. After retention and disposition decisions are made, however, it is perfectly possible to use technology to facilitate records management practices. This paper is composed of the following sections: *Appraisal and disposition on blockchain systems* presents the possibility of using blockchain systems only for disposition but not for appraisal; *Using smart contracts for disposition procedures* presents how a smart contract could execute the disposition procedures and discusses some of the benefits for reducing human intervention; and finally the *ArchContract model* section introduces the technical aspects of the smart contract for disposition, followed by the conclusions.

II. Appraisal and Disposition on Blockchain Systems

The need to select records for destruction became evident in the middle of the 20th century, as organizational models changed, and paper records production exploded. At the beginning of the digital era, the volume of digital records – logical objects stored in a digital medium – was sufficiently limited that the idea of keeping everything in magnetic tapes was not so irrational. As time went by it became clear that, as with traditional records, preserving every digital record produced was not only expensive but also challenging, given technological obsolescence and access difficulties. Maintaining records indiscriminately "leads to a host of information discovery issues, increase[s] cost, litigation and employee dissatisfaction" reinforcing the need for "better appraisal practices, as well as for scalable approaches for dealing with an increasing volume of digital information" [4].

Effective information management practices require both the selection and preservation of relevant records and the destruction of records with no ongoing value to an organization. The decisions over what to destroy and what to keep in the context of records management are made according to laws, regulations and international standards, with a focus on protecting one's organization from litigation. According to [5], all archival work depends on archivists' primary obligation, appraisal, considered one of the main responsibilities of records professionals. The appraisal process consists of "making a judgment about the value of information both from the point of view of optimizing the operational efficiency of the organization and from the point of view of its long-term preservation" [4].

Disposition is the fate of a record after the active and semi-active stages of its lifecycle, as shown in Figure 1. The records lifecycle is divided into the active, semi-active and nonactive stages, based on the frequency of use of the record [6]. The active stage consists of the phases of creation, filing (the act of classifying a record according to a classification scheme), use, and selection through the application of a retention and disposition schedule. In the semi-active stage, records can be transferred to records centers where they are identified, organized, referenced and reach their final disposition. At the end of the active and semi-active stages, if a record is no longer needed, it must be destroyed. If, after those two stages, records

978-1-7281-0858-2/19/$31.00

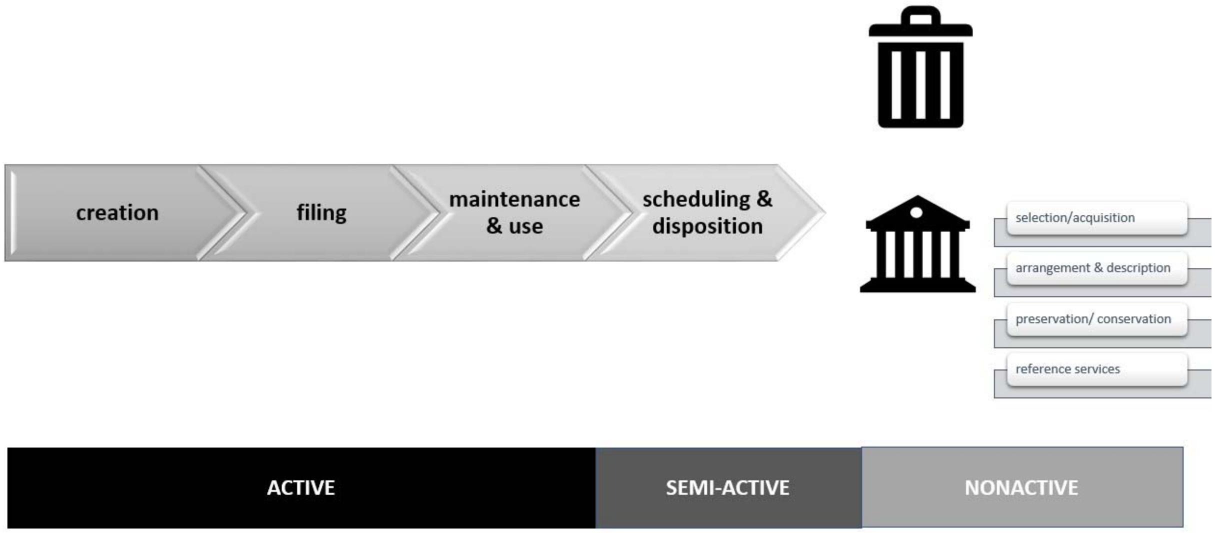

Figure 1 - Records life cycle

are considered historical, they are then transferred to an archival institution where they will be preserved throughout space and time. Whether the record will be destroyed or transferred to an archive, this procedure is documented according to a disposition authority. Although we affirm that appraisal is a function that cannot be automatized, there is a possibility of using disruptive technologies such as blockchain systems to perform records disposition.

Blockchain is defined as "an ever-growing, secure, shared record keeping system in which each user of the data holds a copy of the records, which can only be updated if all parties involved in a transaction agree to update" [7]. Blockchains are especially useful in situations where there is a trust issue between parties, given the technology's strong tamper-resistance. One of the biggest concerns regarding blockchain systems for the execution of archival functions, however, is its immutability and the problems caused by the impossibility of erasing information once recorded in the chain. Despite this, it would be possible to use blockchain technology to improve disposition procedures. In [8] the authors propose an individual development of a permissioned blockchain-based archival system in which the blockchain is used for storing the hashes of off-chain records. In that solution, they propose storing the records of transactions on the blockchain before the nonactive stage where the problem with erasing records persists. On the other hand, the disposition process should be tracked on a blockchain-based system. In contrast with that approach, in this work we propose a model for a smart contract, one of the most prominent blockchain applications, to control the disposition of records kept in an Electronic Document and Records Management System (EDRMS) using a blockchain platform.

It is possible to deploy smart contracts in environments other than blockchain platforms. However, immutability is not an attribute of smart contracts in those environments and we need immutability for our solution to increase trust between the parties (organizations and archival institutions). EDRMs could automatize the disposition process and select the records for destruction preparing the transfer of the valuable ones to the archival institutions. However, even EDRMs that follow international standards in their structures cannot provide integrity checks with the same quality provided by the blockchain hashes. In our solution, the hashes for the records and the records' metadata are a resource for their integrity check by the archival institution. We are aware that the hashes cannot guarantee the integrity of the records and their metadata, but they are a means to increase the trust between the parties and increase the transparency and accountability of the organization-archives relationship. We also highlight that hashes cannot assure the trustworthiness of records throughout space and time which might be a responsibility of the archival institution after the transferring.

The solution proposed in this paper uses a private permissioned blockchain platform operated by a consortium of different parties. The model contemplates organizational records produced in an ecosystem composed of the following stakeholders as shown in Table 1: organizations as creators of records and archival institutions as the receivers of historical records and disposition authority regulators. In our process, the organization, as the creator of a record, has authority over the creation, filling, maintenance, use and disposition of records. The archival institution is responsible for the acquisition of the records and for using the archival blockchain hashes to match the records transferred to the permanent repository, guaranteeing their integrity at the moment of transfer. For this solution to work, the organization must possess a well-established EDRMS using record classification schemes and all the required metadata (e.g. ISO 23081:2017, MoReq2010) in order to keep the context and semantics of the records. There is also an assumption that the access to records transferred to the archives is no longer restricted, meaning that the records can be accessed by the general public.

TABLE I. STAKEHOLDERS AND THEIR FUNCTIONS

Organization	***Archival Institution***
. Creation . Filling . Maintenance . Use . Disposition	. Acquisition . Trusted repository
Transfer the record to the archive Destroy records with no value	Conferring the integrity of records received matching hashes Storing records in secure storage keeping its integrity

III. USING SMART CONTRACTS FOR DISPOSITION PROCEDURES

Smart contracts are defined as pre-written, tamper-proof and unalterable computer programs representing a legal agreement or act of legal significance between two or more parties. A smart contract is stored, replicated and self-executed on a blockchain or distributed ledger, triggered by digitally verifiable events or conditions, and can take custody over and instruct transfer of blockchain-titled assets. Since smart contracts are self-executing computer programs stored in a decentralized ledger, human interaction is reduced, reducing the possibility of human error and diminishing risks such as security breaches and information loss. The goal of our solution is to guarantee that the records that are no longer considered will be selected for destruction, increasing the quality of defensible disposition and assuring the transfer of those considered historical to a trusted repository.

According to [9], the only way of creating and keeping authoritative records is establishing and preserving metadata depicting the following attributes: business context; dependencies and relationships among records and records systems; relationships to legal and social contexts; and relationships to agents who create, manage and use records. Most of those metadata must be generated at the point of a record's creation and capture in an EDRMS and must be preserved throughout its whole life cycle. It is important to highlight four aspects of the metadata that this solution aims to keep: the link between the metadata and the record, the increase of metadata throughout the record's existence, the metadata receiving the same treatment as the record itself, and the maintenance of user control and access permissions over the metadata and the record it is related to.

With regards to records that should be destroyed, the smart contract would automatically detect them according to their record classification code and erase them according to the established disposition authority, respecting the retention period. The blockchain holds a link to the record and its metadata which are stored on an erasable database – the temporary repository. In the case of historical records, one function of the smart contract will be to transfer the record to the trusted repository – the permanent repository. The link to the record and to its metadata will be recorded in the blockchain as a means to authenticate the chain of custody and verify its integrity when transferring to the archives. Part of the metadata is recorded over the record's use altering its logical structure. The updates regarding records' state and their metadata will also be registered in the temporary repository. The smart contract will call the EDRMS by the usage of an oracle service to start the transfer process. In this case, the record will be stored in the permanent repository, which is not erasable, and the links to the record and the metadata are stored in the smart contract. The permanent repository must be a trusted digital repository "whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future" [9]. It may be too soon to tell, but some research has speculated the potential of blockchain systems to become archival repositories [10].

IV. ARCHCONTRACT MODEL

We propose a system for the disposition of information based on blockchain technology and smart contracts. Smart contracts are used in many blockchain-based applications for control and processing purposes. Their tamper-proof and unalterable property makes them an ideal tool for ensuring the processing of information. As mentioned above, immutability is an attribute only assured in smart contracts developed in the blockchain environment. As described above, in the context of archival science this is a key feature. In our proposal, the main responsibility of the smart contract is the administration of the states of a record and the controlled state change only in a case where all requirements are fulfilled.

A. Concept of a temporary and permanent repository

Changes to the content of the record itself can only be made during the active stage when addition can be done to records such as the use of annotations. At the semi-active stage, the records use decreases and almost no aggregation is done. If a record is to enter the nonactive stage and be transferred to an archive, the only changes allowed in its structure are those for its preservation, such as format changes or storage migration. After crossing the archival 'threshold,' the record has to remain authentic, meaning it must keep its integrity and identity. Changes in metadata belonging to the record can only be done during the active and semi-active stage as long as all those changes are registered and can be accessed through space and time. Once a record and its metadata have reached the stage of full retention, both must remain unalterable. As stated in the first section, not all records will be kept for posterity and there are three options for the final disposition:

- Destruction – the record has to be completely deleted;
- Selective retention – a percentage of the records produced to support an activity will be preserved (e.g., 5% of the total records produced) for historical purposes; or
- Full retention – all the records supporting an activity must be kept forever.

Applying this to our context, we have to design a system providing these features of variable changeability during the lifecycle of a record. Another requirement is the full traceability of changes made. To fulfill these contrasting requirements our proposal provides a two-phase approach:

The temporary repository is used during the active and semi-active stages. The record is introduced during the active stage earliest when its record classification code is determined, and only annotation changes are foreseen. During the semi-active stage, single annotations and changes to its metadata might occur. The semi-active stage also includes the retention period of the record. The temporary repository has to fulfill the following requirements:

- Data entries have to be stored and retrieved without any change. This is needed because we rely on the possibility to calculate a hash value of an entry which has to remain the same if no changes to the entry are performed.
- The access address of a record has to remain the same for the whole time it is in the repository. Changes to the records or its metadata result in a new version including a new access address. The updated entry refers to the predecessor and all later versions of the same record in order to keep the archival bond.
- The access to entries has to be defined by a separate access control mechanism. Even if the access address to an entry is disclosed by some system access is only granted if the required permissions exist.
- It must be possible to delete data entries without any possibility of recovery. This requirement addresses the General Data Protection Regulation (GDPR) and related rights of a person or institution to delete data ("right to be forgotten," which includes a right to erasure as well as a right to de-indexing).

B. ArchContract

In our approach, the blockchain fulfills a documentary role. This means no archival data, as well as no metadata, will be stored directly on the blockchain. To still guarantee unchangeability hashes of those data are stored on the blockchain. Like in the use case of identity management the blockchain takes the role of an unalterable fingerprint of this data. We propose the use of a permissioned blockchain where only authorized users and organizations have writing permissions. The blockchain is provided and maintained by a consortium. In this approach, it can be decided later whether public reading will be permitted.

The proposed smart contract ArchContract acts as a registry for keeping track of the record and its metadata. Changes in stages or content of the record or metadata are reported to ArchContract and either confirmed and stored or rejected by ArchContract. To hold the versions of a record and its metadata together an ID (unchanging identification) for each record is used. This ID can also be used to verify if the latest version of an entry is used. It is important to understand that a smart contract cannot guarantee the correctness of the content of a record or its metadata. But even if malicious information is provided with the smart contract allows the tracking of the origin and ensuring consistency.

Since no archival data is stored on the blockchain this data has to be stored on the external repositories. The temporary repository guarantees the unchangeability of data and still allows the deletion of data entries. While the permanent repository, like a blockchain, only allows write once and read transactions. New versions of the record and its metadata are linked to their predecessors to guarantee an audit trail. All versions are held together by the unchangeable ID which is assigned at first entry.

The proposed ArchContract takes the responsibility of initiating the transfer as well as documenting the correct transfer by comparing the hashes of both documents on the temporary and permanent repository. Since no changes between both data are allowed both checksum hashes have to be the same. Figure 2 shows the overall process of our proposal depicting the three stages of a record and the interactions between ArchContract and the two repository systems.

1) A new record is stored in the temporary repository by an authorized user of the organization (e. g., the records manager).

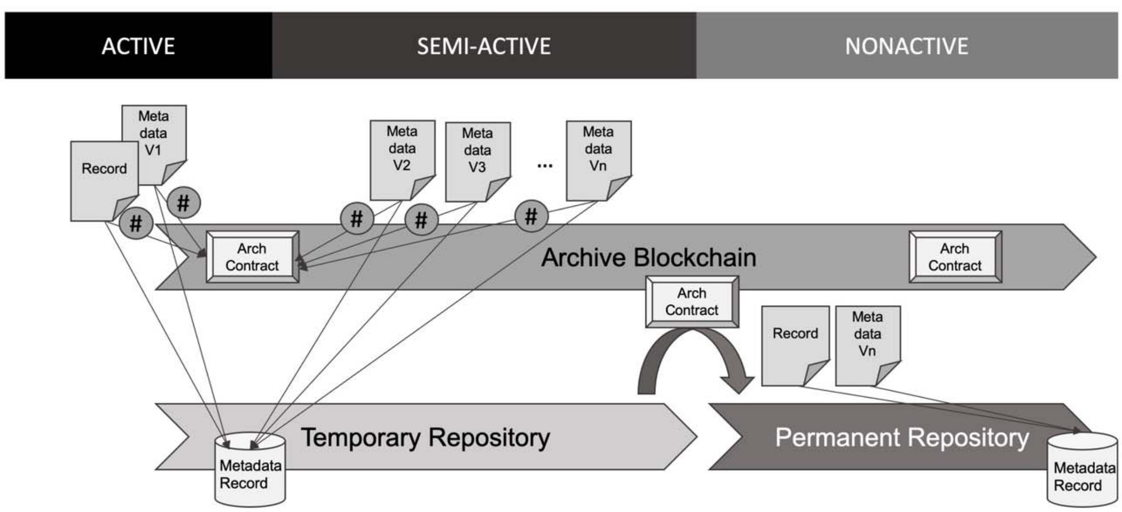

Figure 2 - Process of ArchContract

The temporary repository returns the hash (#) value of the record as an access address (similar to the InterPlanetary File System – IPFS).

2) The function *addRecord* of the ArchContract is called providing the hash of the record from 1), the hash of its metadata from 1) and the classification code of the record (ISO 23081:2017). This adds the record to the registry of ArchContract and returns the ID for the record.

3) An update of metadata results in a new entry in the temporary repository having a new address (hash). This update is reported to the ArchContract using a function *updateMetadata*. Both entries (old and new) are linked to allow the traceback at a future point.

4) An update of the record itself results in a new entry of the record as well as a new entry of the metadata (see 3). Both hold a reference to their predecessors.

5) The stage change of a record (e.g. from active to semi-active) can be initiated by an authorized user or organization. On a stage change, the address of the record and the metadata has to be provided to the ArchContract.

6) The transition from the semi-active stage to the nonactive stage (final disposition) is initiated by an archival authority. This can be guaranteed by the caller ID of this person which has to be registered in the ArchContract.

a) In the case of keeping the record according to its functional classification, the record and its metadata have to be stored at the permanent repository returning a unique address for the record and its metadata. We assume that these addresses are the same as in the temporary repository since the record has to be the same and therefore its hash has to be the same. The record and its metadata in the temporary repository are deleted. Since actual blockchains do not allow direct interaction with the "external world" an oracle service has to be used (see 4.3). This oracle service takes the unique identifier (hash) of the record and of its metadata on the temporary repository and executes the transfer to the permanent repository. Furthermore, the oracle is responsible for the deletion of the data in the temporary repository after the transfer has been confirmed. In addition, ArchContract keeps this information to be verified from outside. The ID of the record remains unchanged.

b) In the case of destroying the record, the semi-active stage includes the retention period which is defined by the record classification code. The deletion of the record and its metadata in the temporary repository is delegated to the same oracle service mentioned in a) and in addition, the fact of deletion is stored in ArchContract for later verification. To be compliant with the requirements of a defensible disposition a new record for this deletion has to be generated by the ArchContract documenting who authorized the disposition, under what authority, when it was carried out, and who witnessed it. This record is added to the permanent repository allowing no inferences to the content of the original record or its metadata.

C. Some Common Mistakes

ArchContract has to be called by authorized users. These users are identified by their access to their individual blockchain addresses (public and private key). Those rights are assigned by a consortium-based approach which excludes the sole decision-making authority of a single party. Additional users can be added using several mechanisms (e.g., voting or multi-signature). The system must ensure that users' access can be removed. Adding new records, updates of records and metadata as well as stage changes can only be performed by authorized users for these use cases. Since these rights are already handled by the blockchain a secure integration into ArchContract can be guaranteed. For example, the transition from semi-active to nonactive has to be performed by a representative of the archival institution.

Smart contracts are executed inside the blockchain system on many or all nodes of a blockchain. This guarantees the consensus on the execution and the results of this execution. Any interaction with the "world" outside the blockchain requires – at the actual state of the art – an additional service called an oracle. We assume that both repositories (temporary and permanent) provide this kind of oracle service. The task of the smart contract is to ensure that these oracle services are triggered by a change of stage.

V. Conclusions

Appraisal is one of the most relevant archival functions and a necessary step for disposition. Appraisal is certainly a function that demands human expertise; however, disposition procedures could be executed using blockchain systems smart contracts. The immutability of blockchain records has always been a concern for executing disposition, but only when documents are recorded on-chain. However, immutability can be an advantage in terms of supporting some of the archival functions and the requirements for records' authenticity.

In this paper, we proposed a solution for the disposition of off-chain records using a smart contract which we called ArchContract. The goal is to keep the link between the metadata and the records, the grown of metadata through records' existence, the metadata receiving the same treatment as the record itself and the same access permissions. Our solution consists of two phases being a temporary repository and a permanent repository. ArchContract acts as a registry for keeping track of the record and its metadata changes in state or content as a means to keep an audit trail. ArchContract also initiates the transferring of records from the temporary repository to the trusted permanent repository. All the procedures must be taken considering the use of an EDRMS compliant with regulations and best practices. By the minimization of manual interventions and the unalterable documentation in a blockchain the whole process of disposition gets more trustworthy.

We are aware that blockchain systems present some controversies and are not reliable recordkeeping systems. However, smart contracts present great capacity for implementing solutions related to archival functions, such as the one we propose in this work. Immutability is a problem regarding the disposition of records; however, it is also a great feature to support requirements to assure authenticity. The solution presented in this paper might demonstrate some exposures, but it is, at least, the beginning of a conversation between archival functions and the use of smart contracts.

REFERENCES

[1] V. L. Lemieux, “A typology of blockchain recordkeeping solutions and some reflections on their implications for the future of archival preservation,” in 2017 IEEE International Conference on Big Data (Big Data), 2017, pp. 2271–2278

[2] V. L. Lemieux, “Blockchain and Distributed Ledgers as Trusted Recordkeeping Systems: An Archival Theoretic Evaluation Framework,” IEEE Future Technologies Conference, 2017.

[3] V. L. Lemieux, “Trusting records: is Blockchain technology the answer?,” Records Management Journal, vol. 26, no. 2, pp. 110–139, Jul. 2016.

[4] A. Vellino and I. Alberts, “Assisting the appraisal of e-mail records with automatic classification,” Records Management Journal, vol. 26, no. 3, pp. 293–313, 2016.

[5] B. W. Dearstyne, The Archival Enterprise: modern archival principles, practices, and management techniques. Chicago and London: American Library Association, 1993.

[6] L. Duranti, “Records Lifecycle.” Rowman & Littlefield Publishers, Lanham, MD, UNITED STATES, 2015

[7] I. Bashir, Mastering Blockchain, 2nd ed. Birmingham, UK: Packt Publishing, 2018.

[8] A. Galiev, N. Prokopyev, S. Ishmukhametov, E. Stolov, R. Latypov, and I. Vlasov, “Archain: a Novel Blockchain Based Archival System,” in 2018 Second World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), London, 2018, pp. 84–89.

[9] “ISO 15489-1:2016 - Information and documentation -- Records management -- Part 1: Concepts and principles.” International Organization for Standardization (ISO), 2016.

[10] V. L. Lemieux, “Blockchain and Public Record Keeping: Of Temples, Prisons, and the (Re)Configuration of Power,” Front. Blockchain, vol. 2, 2019.