ARCHEION
2022, T. CXXIII, s. 10–30
ISSN 0066-6041
e-ISSN 2658-1264

DOI 10.4467/26581264ARC.22.015.16668

Why a world gone digital needs archival theory more than ever before?

Luciana Duranti

Uniwersytet Kolumbii Brytyjskiej w Vancouver / University of British Columbia, Vancouver, British Columbia (Canada)
luciana.duranti@ubc.ca, ORCID 0000-0001-7895-1066

ABSTRACT
The paper is a transcript of a lecture delivered on 12 October 2022 at the University of Warsaw by Professor Luciana Duranti, researcher and lecturer in archival science and diplomatics at the School of Information of the University of British Columbia in Vancouver, B.C. Canada. The lecture was part of a series of open lectures organised by the national archives, entitled 'Konarski Lectures', focusing on contemporary archival science. The author is a specialist in the fields of archival science, diplomatics and electronic records, and Principal Investigator and director of the InterPARES Project (www.interpares.org and www.interparestrustai.org). The lecture addressed the most important challenges facing archival science in the age of digital transformation. In answering the central question regarding the need to develop archival theory in a digital world, the author highlighted the essential research areas that require new theoretical developments. These include ensuring reliability, accuracy and authenticity, that is, trustworthiness, of electronic records, in particular assessing the suitability of electronic signature technology and blockchain architecture for these purposes, and the application of AI to generating, managing and sharing electronic records. The latter is a key problem area of the international interdisciplinary archival research project InterPARES Trust AI, which aims to design, develop and use artificial intelligence to ensure the availability and use of trustworthy electronic records of public entities. The problem of ensuring the authenticity (i.e. the identity and integrity) and the accessibility of digital records stored in the cloud, particularly in the commercial cloud, was also addressed.

KEYWORDS
preserving electronic records, trustworthiness of digital records, Project InterPARES Trust AI, cloud storage, artificial intelligence, blockchain technology

Dlaczego cyfrowy świat potrzebuje teorii archiwalnej bardziej niż kiedykolwiek wcześniej?

STRESZCZENIE
Tekst jest zapisem wykładu prof. Luciany Duranti, badaczki i wykładowczyni archiwistyki i dyplomatyki w School of Information na Uniwersytecie Kolumbii Brytyjskiej w Vancouver (Kanada), który odbył się 12 października 2022 r. na Uniwersytecie Warszawskim w ramach cyklu organizowanych przez archiwa państwowe otwartych wykładów pn. „Konarski Lectures", poświęconych współczesnej archiwistyce. Autorka jest specjalistką w dziedzinie

SŁOWA KLUCZOWE
przechowywanie dokumentów elektronicznych, wiarygodność dokumentacji cyfrowej,

archiwistyki, dyplomatyki i dokumentu elektronicznego, pełni funkcję kierownika badań i dyrektora Projektu InterPARES (www.interpares.org oraz www.interparestrustai.org). Wykład poświęcony został najważniejszym wyzwaniom stojącym przed nauką o archiwach w dobie transformacji cyfrowej. Odpowiadając na tytułowe pytanie o potrzebę rozwijania teorii archiwalnej w świecie cyfrowym, Autorka wskazała najważniejsze obszary badawcze, wymagające nowych ustaleń teoretycznych. Należą do nich m.in. zapewnienie niezawodności, dokładności i autentyczności, to znaczy wiarygodności, dokumentu elektronicznego, w szczególności ocena przydatności technologii podpisu elektronicznego i architektury blockchain oraz zastosowania AI do wytwarzania, zarządzania i udostępniania dokumentacji elektronicznej. Te właśnie zagadnienia są kluczowymi obszarami problemowymi międzynarodowego projektu interdyscyplinarnych badań archiwistycznych InterPARESTrustAI, którego celem jest projektowanie, rozwijanie i wykorzystanie sztucznej inteligencji do zapewnienia dostępności i wiarygodności dokumentacji elektronicznej podmiotów publicznych. Poruszony został także problem zapewnienia autentyczności (tzn. identyfikacji i integralności) oraz dostępności dokumentów cyfrowych przechowywanych w chmurze, w szczególności na zasadach komercyjnych.

Project InterPARESTrustAI, przechowywanie w chmurze, sztuczna inteligencja, technologia blockchain

I will talk for a long time; I hope it will not be too long, but this has been the subject of my work for the past 25 years.

So, the premise is that records and archives are not just an infrastructure; they are a "critical" infrastructure, just like electricity and water, because, through this infrastructure, the beliefs, values and facts of our society are upheld, and our human institutions are supported. This means that societies need to be able to trust the records implicitly, without having to examine each one to see whether it can be trusted. For this to happen in the digital environment, the records must be made explicitly trustworthy by the archivists.

In fact, in the digital environment, the standard of trustworthiness is that of the ordinary marketplace, that is, buyer beware. Thus, in order to make the records explicitly trustworthy, archivists must be able to substitute what they do not have with something else. In a digital environment, we no longer have original records. We no longer have immutable fonds. But we can introduce metadata and security. This is the greatest challenge of our times.

As I imagine Professor Konarski would have done, let's start with definitions, so that you will know what I'm talking about. I use the terms "records" and "archival documents" as synonyms. A record or an archival document is any document made or received in the course of activity and kept for further action or reference. Because every record is first a document, that is information affixed to a medium, a record has stable content and fixed form. I will explain later what stable content and fixed form mean in the digital environment. Because of the

circumstances of their creation, because they are means to a purpose, instruments of activity, records are natural, they are by-products rather than products, they are interrelated, they are linked by an archival bond to other records, and they are impartial with regard to the questions that will be asked of them in the future. To preserve a record means to ensure its physical or technological stabilisation and the protection of its nature, of its intellectual content, and of the relationship that it has with all the records of the same activity. What is different between analogue records and digital records is that, in the digital environment, the content, the structure, and the form of the records are no longer inextricably linked. The record as a stored entity, as encoding, is distinct from its manifestation on a computer screen, and the digital components of the record must also be considered, in addition to its documentary form. For example, consider an e-mail: it will have digital components, such as the header, the message, the block signature, maybe a digital signature, and maybe attachments, that are all in separate places in the electronic system. When you open the email, you see these components all together, but, when you close it, they go to their different places. Whenever you recall a record, you have to reconstitute it. In so doing, you create a copy; there is no original anymore. This is important, because what you preserve overtime is the digital components of the record; it's not necessarily what you see on the screen, though it should look the same every time you recall the record on the screen. When we retrieve records, we in fact generate copies. There are no originals in the digital environment. What does this mean? It means that it is not possible to preserve digital records. We can only preserve our ability to reproduce or recreate them. You reproduce a record when you have a traditional record, like a report or a memo, which you can close and then show again; but you recreate a record when such entity doesn't exist as such in the system, as in the case of a relational database: when you ask the same query that first produced what you saw on the screen, you should be able to produce the same record. This is a recreation of a record. Thus, digital preservation is the process of generating and maintaining authentic copies of digital material and keeping them accessible through different generations of technology over time, irrespective of where they are stored.

In the end, it all amounts to a question of authenticity. Can we trust the records now? The archival canon says that a record is authentic if it is what it says it is, what it purports to be. So the question is, how do we know what the record purports to be? Well, you can look for this information in many places.

You can look at the back of the record (i.e. the verso), at the classification code, at the subject line, or you can even look at the medium. We know, if one studied the Papal records of medieval times, that, if the parchment on which a record is written is candid and almost transparent, the record is likely to be a privilege. If the parchment is yellowish, it's likely to be an executive order. It is the medium that sometimes tells you what something is or who has the custody of it, because that can also be revealed by the material on which a record is written.

What a record is purported to be depends on the culture of the place and time, on the discipline in the context of which authenticity has to be established, on technological context or on the law. Civil law – I understand Poland has a civil law system – determines authenticity by focusing on who issued the record, rather than on the record itself; common law instead focuses on the record itself. Thus, in civil law, we presume authentic any record made by a sovereign authority or in its name, that is, any public record; or one made by a delegate, such as a notary or lawyer. But, what if the person that issues the record is not the sovereign authority? Well, then we go and look at diplomatic authenticity. We know – we are all students or professors of archival science here – that diplomatics was born to prove the authenticity of records, developed a methodology to demonstrate it, and it did establish a scientific methodology for determining it. This methodology was based on form, that is, on the rules of representation used to convey a message, and on the record's degree of perfection – whether it is a draft, a copy or an original. The form, of course, is physical and intellectual, and if it corresponds to the presumed or declared time, place, and order, then the record in that form is considered authentic. What diplomatics does is to establish authenticity on the basis of the record or what appears on the face of the record, what you see.

This is very different from archival authenticity. We know that archival science doesn't look at the face of the record. Archival science includes authenticity among the qualities or characteristics of every record. Every archival document or record is authentic by definition. Together with naturalness, impartiality and interrelatedness, authenticity is a characteristic of all records. Why? Because archival science looks at records in relation to their creator, and says that, if the creator used the record in the course of the usual and ordinary business, which was carried out trusting such record, even if the record is, diplomatically, a forgery, it is authentic with respect to the creator. Thus, for archives, the records are authentic when they are made or received and kept for the need to act

through them, and when they are preserved as evidence of facts and acts by the creator or its legitimate successor. Thus, archival science, by linking the record to the context of creation and preservation, extends authenticity from being a property of the record itself to being a property of procedure, and ties it to unbroken custody.

What happens in the digital environment? In archival science, there was no question that the identity of a record, and therefore its authenticity, resided in the provenance and the documentary context of the record. But this turned out to be linked to the immutability of a record that is affixed to a permanent medium, that is, to integrity – whether what you have today is exactly the same material thing that you had 3 centuries ago. However, in the digital environment, authenticity cannot be assessed only on the basis of context. In fact, even if the relationships among the records established at creation remained intact over time, the documentary component of the record – remember that a record is made-up of a document and its relationships – could lose integrity, which is a quality of the record. We never paid attention to it before, because content, structure and form were inextricably linked at the time, but they no longer are – content data, composition data, and form data are separately stored digital components. Thus, in the InterPARES project, we returned to diplomatic authenticity and looked separately at the identity and the integrity of the record.

What is the identity of a record? Identity refers to the attributes of a record that uniquely characterise it and distinguish it from all the other records. They include, in the digital environment, the author, who is the person issuing the record, the addressee – the person for whom the record was intended, the writer – the person who articulated the content of the record, the originator – the person on whose digital account the record was generated, and the creator, the person in whose fonds the record exists. Those persons must be all identified in the metadata for every single digital record. The other elements of identity are the dates of creation, that is, the date of making the record, receiving it, and filing it, and the date of transmission; the matter or action in which a record participates; the expression of the relationships with other records, which might be a classification code; and an indication of any attachment. But, when we talk about integrity, how do we define integrity? Integrity refers to the quality of being complete and unaltered in all essential respects. That is the key term here – essential respects. We have never been really fussy about it. What if we had a record on parchment with holes in it? What if the ink passed through to the

other side of a paper document? As long as we could read the record, it was just good enough. What is good enough in the digital environment?

In the digital environment we must have bitwise integrity, also defined as data integrity. It is the kind of integrity that the digital signature protects. You change one bit, and the document won't open any longer. Integrity thus means that the data in the document are not modified, either intentionally or accidentally. The original bits are in a complete and unaltered state from the time of capture. They have the exact same order and value, as a small change could result in a very different value.

In the digital environment we must also consider duplication integrity. As I said at the beginning, all we can do to "preserve" digital records is to make copies that can be trusted. Our primary activity as digital archivists is duplication. Duplication integrity means that the process of creating a copy does not modify a record, either intentionally or accidentally, because the output is an exact bit copy of the original data set. So form data, content data, and composition data are all the same. However, duplication integrity is linked to time, because every time you open something, you create a copy. So at least onemetadata changes, the time. Also, in the digital environment, when we say duplication, what do we really mean? We might mean we make a copy. A copy is a selective duplicate. A PDF is a copy. You only copy what you can see. Which means, if there are metadata, they are not in the copy. It doesn't include confirmation of completeness. It provides an incomplete picture of the digital environment. An image, on the other hand, is called a forensic duplicate, as it is a bit by bit reproduction of the storage medium. The storage medium and its content, including ambient data, that is, a snapshot of every file that you have open; swap space, which is the virtual memory with all your passwords and encryption keys; and slack space, with all the material you have deleted, would be included in the duplicate, a very unethical thing to do. Many have done this. I know universities where the hard drive of the heads of departments have been automatically copied by technological services to preserve their activity, but in fact they preserve much more than that.

Also the duplication process needs to have integrity. It must respect basic principles: The principle of non-interference, which means that whatever method you use to reproduce or recreate a digital document does not change the digital entities, and the principle of identifiable interference, which means that duplication does change the entities, but the changes are identifiable are identified, with the attachment of paradata. I don't know whether you use

this term, but paradata is the data related to the persons who manage, keep or preserve the records and the consequences of theirs actions. So any archivist who makes any change to the records should be named; the time when the change is made should be recorded; and so on.

In a way, what we need to do is to be able to authenticate the records at any given time. Authentication is a declaration of authenticity based either on direct knowledge (e.g. I have written the record), a material proof (e.g. I have 10 identical copies of the same record), an inference, or a deduction. What are the bases for authenticating digital records? There is a combination of them. A chain of legitimate custody remains a ground for inferring authenticity and authenticate the record; itis what used to be called the unbroken chain of custody. If I trust the custodian(s) over time, I trust the record. Also, the digital chain of information is a base for authentication. It is the information linked to the record, that is, the metadata added throughout the life of the record: not only the original identity metadata, but also integrity metadata related to every time that the technological environment changes and every time anybody interferes with the record. Finally, or a declaration made by an expert, who bases it on the trustworthiness of the system that hosts the record throughout its life, and the procedures and processes controlling its preservation and use.

What does "trustworthiness of the system" mean? Some countries, Canada among them, consider system integrity the basis for declaring authenticity if one can prove that nobody has interfered with the system by looking at the various logs through time. Then one can infer that the records are authentic. The reason, again, is that digital records are always new, and that we can only preserve our ability to reproduce them. This is based on two considerations. There are some records that are the counterpart of traditional paper records (say, a memo or a report). The physical form in which these records are stored is different from the documentary form we see on the screen. It is also different from the physical form of the record in a computer processor. Then there are digital records which have no analogue precedent, such as experiential records and interactive records. With these records, it is not possible to preserve, in digital form, a copy intended for human use. Their copies can only be preserved for machine use.

Thus, in the context of the InterPARES project, we came up with definitions of two types of record: the stored record and the manifested record. When we think of a paper record, we can hold it in our hand; it is a single entity, a single thing. But when we look at a digital record, there are really two entities there. One

is what we see – the manifested record. The other one is what is in the system. The record that we preserve is the one stored in the system. It is different from the manifested record, which is a copy of the record in a form that is readable either by humans or by an automated system.

The general requirement for preserving digital records is that, regardless of the stored record, which is represented in bits in digital storage, and of how its encoding may change within the system, it must be possible to generate from it a manifested record which has exactly all the attributes of the first effective (i.e. complete and capable of reaching its consequences) version of the manifested record that we saw. As you see, this is very different from bitwise integrity, which doesn't allow changing even a single bit. That's what the digital signature ensures: bitwise integrity. However, as we change systems through time and technological advances, the bits must change. What must not change is the record that we see on the screen every time we open a stored record.

The practical issue here is: how feasible is the making of authentic reproductions? Because of this difficulty, technological experts have tried to solve it through technology by developing technology dependent authentication. That's what they've always tried to do since the issue of authenticity arose. First, they tried with the digital signature. What does the digital signature do? As mentioned, it protects bitwise integrity. If any bit, or their order, is changed, the document doesn't open, you can't see it. It verifies the origin of a record – basically, it tells you from where that record came, and it makes the record indisputable and incontestable, which means you cannot say "I never sent you that record". It has been given legal value by a number of legislative acts, including the European directive on electronic records. It is enabled through a complex and very costly public key infrastructure, and ensures authenticity of information across space, when moving from one person to another, from one organisation to another. It does not ensure the authenticity of the records through time, though, as it is subject to obsolescence much faster than the actual record to which it is linked, and that makes the problem of preservation much more complicated, because the digital signature cannot be migrated to a new system together with the record to which it is attached. You have to detach it first and then reattach a digital signature to the migrated record. Also, certificates linked to the digital signature have an expiration date of 5 years; then what?

Theory helps us; theory has always helped us. In this specific case, diplomatics theory tells us that the digital signature has the function of a seal, not of

a signature, because it is attached to a record that is complete without it, and this means that it can be removed and substituted with metadata. When you receive something with a digital signature, you can detach the digital signature and add to the metadata that the document had been received with a digital signature attached and that the signature was removed by the recipient after verification at a specific time. What else is needed? If the records manager or the archivist authenticates the record, the original authentication is no longer needed. This, in fact, is implemented in North America.

Now let's move to another technology that is very popular, especially in Eastern Europe: blockchain. In Estonia, for example, blockchain is a very popular technology. What is it? It's the technology that enables Bitcoin. It is a ledger, that is, an information store that keeps a final and definitive trace of transactions. Please keep in mind, that it does not keep the transactions themselves. It does not keep the records of the transaction. It keeps the hash algorithm of the record that carries out the transaction. Records as such are kept somewhere else. They are not on the blockchain, as the blockchain only holds algorithms. How is record authenticity verified? There is a record stored; a hash algorithm is made from it and compared with the one on the blockchain. If they are identical, the record is authentic. Blockchain relies on a distributed network, which means that all the servers that have the same record in it, or better, the same trace of record in it, are identical. They work based on decentralised consensus; that means that each server has to approve that trace of a record as authentic. What this means is that, if somebody wants to attack and destroy records or a system, they cannot possibly attack all of them. In a decentralised system, one can attack a few servers, but then there are others. This is what this idea of guaranteeing the authenticity of the record is based on.

These validated sets of algorithms are held in blocks, which are linked in a chain that cannot be tampered with and to which one can only add things. One cannot delete anything from a blockchain, only add. How does it work? Let's say that you start with one block of 1000 transactions and no block can have more than 1000 algorithms of 1000 records. When this 1000 number is reached, then one takes a hash of this block and starts another block with it, then adds another 1000 transactions and takes a hash of this second block that included the hash of the previous one, and so on. This is a chain that cannot be tampered with in any way.

There are three types of blockchain. There is the public blockchain where anyone can participate in reading, writing and editing, without any permissions. These blockchains are open, and they're transparent, like Bitcoin, Ethereum, etc. Of course there is no privacy, there are no controls, everything is in the open. Then, there are private blockchains, where only one organisation can put its records, and there might be public access or restricted access, but only that organisation can actually upload the hash algorithms of its records. Now, of course, what this type of blockchain does is to eliminate the security coming from the absence of a single, central point of attack; they do have a central point of attack. So what's the point of having the blockchain?

Then there is a consortium, which is a private blockchain that removes the individual autonomy of bringing changes to the blockchain, same as in the private blockchain, but it operates under a group of institutions that are all similar. There can be a blockchain for the banks, a blockchain for the energy organisations, a blockchain for insurance companies, a blockchain for archives, etc. How is blockchain used? It is used to confirm the integrity of a record which is kept elsewhere. It is also used to confirm that the record existed or was created before being hashed on the blockchain. It certainly could not have been created after that. It also confirms the sequence of uploading of the records' hash to the blockchain. Since the blockchain is shared by so many different bodies, the uploaded data is not constituted of traces of records that are related to each other. They are in chronological order of uploading, regardless of who uploads them.

As you can imagine, the archival bond is gone here, as the relationship between the records is not maintained on the blockchain; it's definitely not in the traces of the records. Is the blockchain a record keeping system? No, of course it is not. First of all, it doesn't hold any records, only the hash of records. It may include smart contracts, which are actual contracts made in code within the system, but they are not considered records yet, as they don't have signature or a date. The records still must be stored and managed off chain. This is good because, if they were on the blockchain, they would be immutable. One could say – but wasn't that the point? Making the records immutable? Indeed, immutability is the attraction of the blockchain. It is what ensures data integrity, because nothing can be changed, nothing can be removed. It is also the biggest problem of blockchain, because, if you're dealing with active or current records, you cannot update or correct any wrong data. Also, you cannot protect privacy, you cannot exercise the right to be forgotten, you cannot destroy any record you no

longer need. Further, any record-making system upgrade would basically annul everything you have done so far, as any change in the records would invalidate the blockchain. A new blockchain would have to be started. Note that the creation of a blockchain is an expensive endeavour; note also that blockchain is among the most polluting technologies in existence.

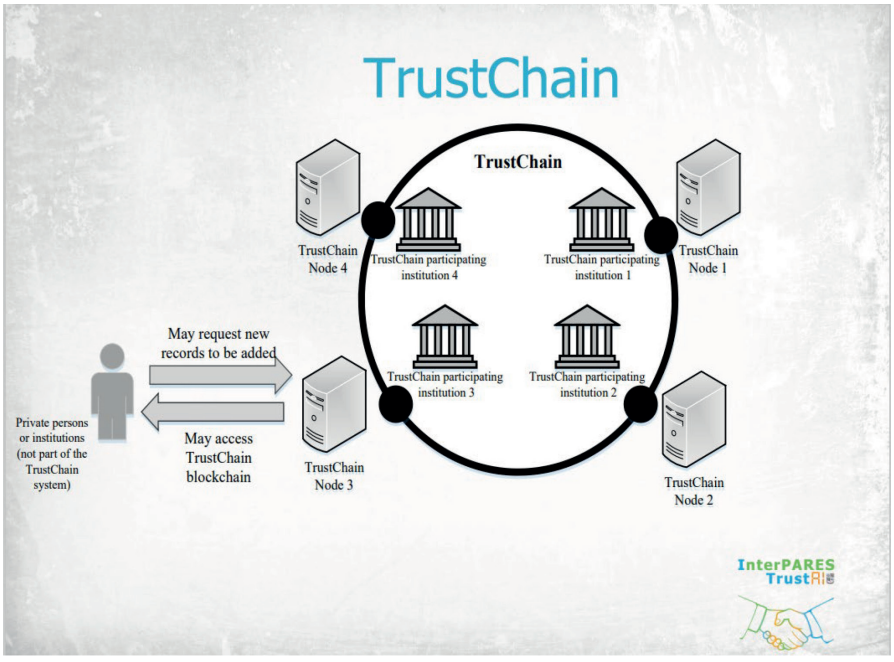
In addition, when it comes to records identified for permanent preservation, any transfer to a preservation system, any migration, any increment of the records aggregation would invalidate the blockchain. You'd have to start over making hashes of all the records. In addition, the hash on the blockchain does not allow for links to the hash of the related records, so – as already stated – there is no archival bond. The interrelationship between the records disappears. The hash of metadata is not there either, so there is no context. If the metadata were embedded into the record at time of creation, one could be tempted to put them on the face of the record and make a hash of the whole thing together. But there would be no possibility to add any metadata related to the technological environment changes, to the responsibilities changes, or whatever other changes.

In addition, there are authenticity problems, because how does one prove that the record was authentic in the first place, before being put on the blockchain? One could have uploaded a forgery on the blockchain to start with. How does one preserve contextual evidence? How does one handle the decentralised nature of the blockchain? The blockchain is trans-jurisdictional. If all these servers all over the world contain the hash of the same records, they each fall under the legislation of their respective location. Decentralisation is a problem, because the processing happens with technical components that are in the custody of different actors. Some components may be under the control of single organisations, other under the control of business partners, and yet others under the control of unknown third party actors. What could happen is that an organisation's records could be in the custody of thousands of independent actors over which records creators have little or no control. The consensus mechanism is also a problem, as it might be not within the decision making sphere of the records creator. These decisions could be made by remote third party developers, so the reliability of the upload of records to the blockchain could be difficult to establish.

What about partial decentralisation, then? The InterPARES project doesn't ever reject anything in principle. Whatever comes up, any possibility of finding a solution, we try it. So, while in principle all those I have described are things that do not work, let's nonetheless try to work with the blockchain and see what

happens. Thus, we built a blockchain-based system called TrustChain. We applied the fundamental concepts of blockchain – the hash algorithm, the idea of blocks, the distributed consensus. We then made some underlying assumptions. The blockchain is in the cloud environment, not in a physical place in any one institution. So, first of all we decided it had to be on a private cloud, not a public cloud. Second, the TrustChain had to be a consortium, an aggregation of institutions of the same kind, and only nodes that are approved would be able to write, while everybody could read.

The InterPARES TrustChain relies on the involvement of a group of trusted institutions, a consortium of archives, and the record keeping systems in the creating offices that transfer their records to such institution so that they work together: the creating office has the records in a record keeping system that is linked to the preservation system, and these records are hashed on the same blockchain. This way one would have integrity, time of creation, correct sequence of records, because one would upload series of records in the fonds. No repudiation because the validity of the signature certificate would not expire, because one could make an hash of the certificate of the digital signature and upload it on the blockchain, where it would stay for 3 years or forever.



Source: Own elaboration.

Inside the circle in the figure, you have the archival institutions. Outside the circle you have the record keeping technologies of every single creating body that transfers records to the institutions, and then, external to them, you have individuals who want to have access to the materials, and can access the materials regularly.

There are still some issues with this model. The archival bond remains a problem, because files are uploaded at a specific time; and as we know, the records within the creating office continue adding up, maybe within the same file, maybe across files. The metadata would then change, but no change is possible in the blockchain. So maybe we should have a supporting system, a parallel system, which makes it possible to update all the metadata continuously and to also update the records that should end in a blockchain.

Nevertheless, to adopt blockchain would be far too risky for archivists. This technology is at an early stage of development. The challenges and limitations include privacy; if a consortium of organisations share the same thing, how do you control the privacy element? Then, compliance; the rules might be different from one institution to another, from one records creator to another. Governance: Who controls the whole thing? Scalability is one big issue, and I will now explain why. Security is another, and I will later explain why.

Let's talk about scalability. Authenticity of the traces of the records in the blockchain is guaranteed by the fact that there are so many actors involved and each one has to approve what is uploaded to the blockchain. Do you realise how long it takes to do that? Every time that one uploads something, before it is actually accessible, every single member of the blockchain has to approve that upload. Thus, time is a problem, and this is all due to security, because the blockchain uses a proof of work – the consensus mechanism I mentioned before. The calculator has to solve, mathematically, very difficult puzzles on each new block, before approving it to the ledger. This proof of work is data that is very costly and time consuming to produce, because everyone must be satisfied with the requirements and producing such proof of work is a random process with low probability. Thus, there is lots of trial and error going on; calculating the proof of work is called mining and takes lots of time. Each time one uploads something, there is a random value in the block header. This value is an arbitrary number that can be used only once in a communication. So, if someone in the chain – the entire chain – has 51% of the computing power, they control basically everything, and, all of a sudden, we have centralization instead of decentralisation, as this

entity can modify the transaction and the transaction data, can stop the block from verifying the transaction and can stop mining in every block. Thus, at this stage, the blockchain is not really a good idea. Maybe, over time, things will be developed further, but at this stage blockchain is not a recommendation that I would make.

We have seen that technological authentication is not working for us. What else can we do? Let's go back to authenticity for a moment. The fundamental difference between the authenticity of analogue and digital records is in the fact that, while the authenticity of analogue records can be proven and verified on their face and only exceptionally one needs circumstantial or extrinsic evidence, authenticity of digital records cannot. Assessment of digital materials' authenticity is always an inference. It is based on extrinsic elements, such as significant properties included in the identity and integrity metadata, and it relies on circumstantial evidence, such as the integrity of the system and the policies that control the system. If a policy says that only a specific individual can have access to the system and the logs of the system prove that only that person accessed it, then while authenticity remains an inference, it's quite reliable. It is also based on the technology that encrypts and secures the access to the system.

So what if we used AI? Artificial intelligence? This is what we are exploring in the latest phase of InterPARES. Artificial intelligence systems are computing systems that use algorithms capable of carrying out complex tasks that normally are carried out by human intelligence, such as processing large quantities of information, calculating and predicting, learning and adapting responses to a changing situation, recognising and classifying objects. The question that we ask in our research project is: can we develop artificial intelligence systems for carrying out, in a competent and efficient way, archival functions, such as preservation, while respecting the nature of the records and assuring their trustworthiness?

We know that artificial intelligence systems have big issues, because they provide evidence based on probabilities, not on facts. They provide evidence that is not interpretable or transparent, and it is only as good as the data that we provide, so, if the data are not good, we are in trouble. AI has outcomes that may have a disproportionate impact on some groups of people. It may challenge the autonomy and privacy of people. Most of all, it is very hard to assign responsibility when artificial intelligence makes the decision. Plus, AI decisions are based on past human decisions and, when it comes to human affairs, tomorrow rarely

resembles today. Finally, data and numbers cannot say what is and is not moral, or socially desirable. Therefore, there is a declaration of principles that should be respected when developing artificial intelligence tools. They are: respect for the wellbeing and the autonomy of people and their privacy; solidarity, democratic participation, equity, diversity and inclusion, caution, responsibility and sustainable development.

What is the past experience of archives with AI? Archives have been looking at artificial intelligence for a long time. They have considered, though, either a specific tool in a specific context or a single set of records. For example, they have used recurrent neural networks for classification of large numbers of records; recommendation systems to make all documents searchable, through written text recognition; chatbots that emulate human conversation to find connected information; and named entity recognition to create visualisation tools for all types of data.

The issue is that, so far, archives have relied on off the shelf tools, already existing, not designed for archives, and these limit what challenges can be met. Such tools make the needs of archives subservient to the field of machine learning. There are many tangible instances of biases when artificial intelligence has been used for archives. This raises the question whether off the shelf tools are a good idea, and what artificial intelligence would look like if the power relation between AI and archives were reversed, with archival theory informing the creation of artificial intelligence tools. That is what InterPARES 5 is all about. 'InterPARES trust AI', or 'I trust AI' project has the purpose of designing, developing and leveraging artificial intelligence to support the ongoing availability and accessibility of trustworthy public records by forming a sustainable ongoing partnership between academia and archives, etc. Its objectives are to identify specific AI technologies that can address critical records and archives challenges; to determine the benefits and risks of using those technologies; to ensure that archival concepts and principles inform the development of responsible artificial intelligence; and to validate the outcomes through case studies and demonstration.

We are carrying out many studies. At this stage, we have about 74 studies going on, about 40 countries participating and 89 organisational partners. We have about 200 researchers from all relevant fields: Machine learning, artificial intelligence, archives, history, law, and others. They focus on all aspects of archival functions, such as creation, appraisal, arrangement, description etc.

The expected outcomes are to improve the existing tools and to create new tools that will address archival needs, such as, for example, machine translation, image recognition and description. There is a project being implemented at the state archives of Milan, where archivists are digitising thousands of parchments, but attaching metadata to each of them would require enormous amount of resources, personnel and technology. Thus, our researchers are developing an optical character recognition tool to automatically create the metadata for all the digitised material, text summarization and classification, and text style transfer for language civilization. We have, certainly in Canada, inventories created a century ago using language that is totally inappropriate, especially with reference to for indigenous people, women, etc. So, there are AI tools that can be used to identify all the inappropriate language in them and, although the original inventories would remain intact in the archives translate it into a language that is acceptable for the finding aids handed to the public for the purpose of identifying the materials they need.

It might be possible in the future to use artificial intelligence based on archival concepts to authenticate archival materials and to detect any interference with it. Still, we need to link any such tool to a much more sophisticated cyber protection agent to be able to protect archives, current and historical, from hostile powers, and to prove that we have successfully done so. Alternatively, we can keep the records permanently offline. If the records are offline, nobody can interfere with them. We could also maintain a complete identical reproduction of the fonds in a secure physical offline location. That's what the UK TNA does. It would not provide privacy and confidentiality, but at least we would have a set of authentic records.

What about records in the cloud? During InterPARES 4, we spent five years studying records in a cloud environment. The first survey told us that the reason why archives, as well as records creators, choose to keep records online, is first of all, economic. Collaboration, efficiency and performance, increased storage, all those are factors; but mostly, the reason is money saving. However, there are issues with records in the cloud that are related to a long list of things. I will cover them one by one.

Let's start with data ownership. When a user entrusts records to a provider and uses its platform, the provider generates additional data related to the actions that they carry out: how they process the material, where they store it, how they change the technology overtime, etc. While the content that you upload to the

provider remains yours, whatever is created by the provider doesn't belong to you, but to the provider and you have no access to it. Thus, before using a provider, you have to be sure that any contract that you make with the provider spells out very clearly what happens to the data that they create about your records. Because you will need those data to prove the authenticity of your own records, based on the integrity metadata that these people will have created. Also, we know that availability of the infrastructure is a fact, while access to the records is a right; but you cannot have access if you don't have availability. In a cloud environment, availability of the stored records implies availability of the infrastructure, which means, for example, that the system is expected to be available 100% of the time. That never happens. We have studied lots of contracts offered by cloud providers, and nobody guarantees availability beyond 98% of the time. Do you realise what this means in terms of how many days in a year the infrastructure where your records are would not be available? That's a big issue. If records creators or individuals who need the records under the Freedom of Information Act within a specific time request access to these records when the system, the technology, is not available, then they are in trouble.

Then there is reliability. Reliability means behaving consistently with expectations. What does that imply? Well, one expectation we would have is that, if 10 researchers at the same time ask for access to the same record, they will see the same thing. However, that's not guaranteed at all. As the records move from a server to another, different metadata are attached to them – be they related to technology, location, use, or other.

The biggest problem, however, relates to retention and disposition. How do you verify that the provider actually complies with your records retention and disposition schedule? You might believe that, since you don't see the records any longer, the provider must have destroyed them. This is not the case. It is too costly and too time consuming for a provider to retrieve all the copies of the records that they have disseminated throughout the data centres. What they do is cut the links to the records, but the records still remain in the servers, for as long as the data center will exist. It is difficult to access them; but this doesn't mean it is impossible for those who are determined to access them – just not through legal channels. In addition, transfer from one system to another for permanent retention may involve loss of authenticity. Let's say that a series of records has been scheduled to be transferred from the records creator cloud to the Archives cloud for permanent storage. Such transfer is a very complex process

and records may lose authenticity in that process. In case of records that are to be actually destroyed, a breach of confidentiality or privacy may occur. As I said, there may be the persistence of copies, but certainly there will be the persistence of the metadata, which are separate from the records that would eventually be destroyed. Furthermore, the persistence of the metadata generated by the provider about the user data is a given. Providers will never destroy their own metadata about what they have done with your records; thus, if you need records to be destroyed – for example, in Canada, by law we must destroy the records of juvenile criminals five years after they have been clean – but all the metadata about those records remains, it is the same as if the records were not destroyed at all. Please keep in mind that those metadata are generated by the provider and don't belong to the user.

Storage and maintenance impact the quality of records and their ability to serve as evidence, especially in case of legal jurisdictions where the authenticity of the record is an inference based on the integrity of the system. As said earlier, Canadian law says that, if we can prove the integrity of the system where the records are, we should assume that the records are authentic. However, how are we going to prove the integrity of the system of our cloud provider, if we have no access to any of their data about those systems? Also, the contract doesn't specify what providers do when the technology changes and the format of the data has to be changed. As you can see, what we say after five years of research is not that you shouldn't use the cloud provider; it's that, when you write a contract with the cloud provider, you have to spell out all of these things. We have actually developed a model contract, suitable for any kind of institution. It is freely available on the website and easy to find.

The providers call anything related to keeping data or records "backup procedure" – to them it's all just backup. And the standard contract says that you are responsible for the backup, not the provider. Thus, security is another issue, probably the biggest one; protecting the system and the records from unauthorised access, use, alteration or destruction. Now please follow my reasoning. In a world where the integrity of a system is the basis to infer the integrity of the record, from which one infers the authenticity of the record and its trustworthiness, security equals authenticity. That's where we got to. At this point, we can't prove any authenticity. We can only prove absolute security, and that's what we have to deal with. Individuals may enforce security with something they know – passwords, with something they own – tokens, or something they

are – biometrics of eyes, fingerprints, private keys. A cloud provider enforces security through encryption, and should produce audit trails and access logs; and should capture, maintain, and make available metadata associated with access, retrieval, use, and management; I say "should." I don't say that they do. They don't. They should. It should be in the contract, spelled out clearly, with penalties if they don't do this, because that's where security resides.

Let's remember what was the main reason to use the cloud – money. Saving money. The moment you start requiring all those security measures, your budget doubles, and obviously so, because the most important encryption is not the encryption in place, it's the encryption during transmission. It is during transmission that the records are intercepted – when they move from the creator of the records to the archives for preservation, on a regular basis. If you use the cloud as an archival institution, it is the creator that moves its material to your cloud. Encryption during transmission is a big security issue, which links directly to data location and cross-border data flow. The cloud is a platform for mobile applications; accessible from smart devices, while the records can be anywhere – in data centres in various parts of the world. The location of the record is the criterion for determining what law applies in case of litigation. National strategies used to require that the records be kept within the borders of the country where those records were created. However, imagine how expensive it is to have a data centre in the middle of Europe, or in the middle of the United States or Canada, where the land is expensive. So, where would data centres be located? In Latin America, in Africa, in the middle of Asia, in places where the land doesn't cost as much. If you are required to have your data centres in Poland, it becomes a problem, because the reason why you wanted to use the cloud in the first place – which was to save money – would no longer justify your choice. You no longer save any money. Thus, the international strategy no longer requires this. It calls for multilateral agreements among countries to collaborate in the area of security. It's called the new safe harbour. That means that countries agree with each other that, while the records may be in Brazil, or in South Africa, or elsewhere, those that are mine fall under my jurisdiction, those that are yours fall under yours, and we take care of security together so that we don't get in each other's way. Such multilateral agreements among countries do exist. Of course, this is easy to do at a national level; but the moment you go down to the level of province or city, it gets much more complicated.

The next issue is contract termination. If the provider ceases to exist or terminates one or more of its services, the records will be deleted or inaccessible. Don't expect the provider to give you the records; it costs them a lot of money to retrieve your own records and pass them to you. Free services do not have an established duration, and providers may terminate the service unilaterally. Sometimes, people who create their own archives assume they would use a free service to store all their records: if the provider decides to close the door, say goodbye to your records. Even if data is given back to the user, it is not guaranteed that it would be in a usable and interoperable format. It may be in a format you can't read, or in a format that cannot interact with your own records. If the contract is terminated by the user, it is not guaranteed that the provider would give you back your data, and if they do, it may be very expensive to transfer it all in a format that is accessible to you. And you definitely would not have the right to access the metadata generated by the cloud provider, and may have no guarantee that the provider will destroy copies of the data held in the data centre. So, even if they give the data back to you, they may have copies, many of them.

Now let's consider records preservation. Preserving records in the cloud is a black box process, that is, you have no idea of what is going on. Providers may not know where the records are; they usually don't. They can and do subcontract some of their services to other providers, who potentially maintain servers or are registered as providers in other countries. One cannot expect that the same hardware and software would remain in service for as long as the records must be preserved. Standards provide information about preservation format, but there is no guarantee that providers would respect the standards; there is also no way of ensuring and verifying authenticity.

Should we give up on the cloud then? Not necessarily. We would have to examine the security measures agreed upon in the contract. Contracts are very important. We can still determine authenticity on the basis of significant properties. Those include the attributes of identity and integrity, but may also mean the logs. Logs are not accessible when they are in a commercial cloud. Thus, if the records have to be in a cloud environment, it should be a private cloud, protected by a strong contract with the provider. This means you don't save any money, but the reason you may want to do it is that it allows for better access from anywhere in the world, for collaboration, as well as efficiency. These are all good reasons. Saving money, however, is definitely not one of them. One thing we must not forget, especially in this time, is cyber security; the records are certainly

much safer in a dark repository than they are online. There is no question about that.

In conclusion. Archivists are not yet attuned, as a profession, to the risk of cyberattacks. We tend to think about cyberspace as something very abstract. But in reality, cyberspace is a very material space – the servers, with records on them, are made of metal. So cyberattacks are very likely. The reason why archivists are not sensitized to this is that they have not been preserving much digital material, and especially sensitive digital material. Nonetheless, they must plan for the protection of the records created today. And if it is true that, with digital records, preservation starts at the moment of creation, cyber protection must start long before then. You have to design a system of cyber protection possibly before you start designing the record keeping system and definitely before starting to create the records, because you cannot stick it on a system afterwards. Things do not work that way. First you need to have all your archival theory in place to understand what you need to protect and preserve, and then you act on the basis of an archival understanding of digital technology. Never the intellectual foundations of our science and discipline have been as essential as today to the survival of our values and beliefs as a democratic open society.

Internet sources

InterPARES Trust AI – Artificial Intelligence, https://interparestrustai.org/, accessed 6 December 2022.

The InterPARES Project, http://www.interpares.org/, accessed 6 December 2022.

YouTube. "Konarski Lectures: L. Duranti 'Why a World Gone Digital Needs Archival Theory More than Ever Before.", https://www.youtube.com/watch?v=vbOY3TXNan8, accessed 6 December 2022.