Prejeto / Received: 4. 5. 2026

1.01 Izvirni znanstveni članek 1.01 Scientific article

https://doi.org/10.54356/MA/MRQL1967

CONTEMPORARY APPROACHES TO ARCHIVING OF DIGITAL ORIGINALS: A REVIEW OF CONCEPTS AND TECHNOLOGIES

Arian Rajh, PhD
Associate Professor, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia
arajh@ffzg.hr

Hrvoje Stančić, PhD
Full Professor, Faculty of Humanities and Social Sciences, University of Zagreb, Croatia
hstancic@ffzg.hr

Jelena Bolkovac
Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia
jelena.bolkovac@fsb.unizg.hr

Abstract:

Digital materials are increasingly prevalent as a means of record creation. This paper presents interim results from the research project "Long-term preservation of the authenticity of digital originals using blockchain technology and artificial intelligence (I.Trust.Original)", funded by the European Union's NextGenerationEU fund, and focuses on concepts and tools for archiving digital materials.

The authors review the concepts of the digital original and the archival bond and examine their application within the Open Archival Information System (OAIS) framework. Blockchain and distributed ledger technologies (DLT) are analysed as mechanisms for creating digital originals and for ensuring integrity, transparency, and trust. The paper also explores the role of artificial intelligence (AI), including large language models (LLMs), in archival processes and in supporting blockchain-based preservation. Semantic web and linked data are regarded as tools for enhancing interoperability and discoverability, while smart contracts are recognised in the context of rights management.

The paper aims to map these concepts and technologies and assess their impact on the archiving of digital originals, highlighting key benefits and limitations. It emphasises the interconnectedness of archival theory and technological innovation, and outlines directions for future research and practical implementation.

Keywords:

archives, artificial intelligence, blockchain, digital original, semantic technologies

Izvleček:

Sodobni pristopi k arhiviranju digitalnih izvirnikov: pregled konceptov in tehnologij

Digitalni zapisi vse pogosteje prevladujejo pri ustvarjanju dokumentarnega gradiva. Prispevek predstavlja vmesne rezultate raziskovalnega projekta »Dolgoročno ohranjanje avtentičnosti digitalnih izvirnikov z uporabo tehnologije veriženja blokov in umetne inteligence (I.Trust.Original)«, ki ga financira sklad NextGenerationEU Evropske unije, ter se osredotoča na koncepte in orodja za arhiviranje digitalnega gradiva.

Avtorji obravnavajo pojma »digitalni izvirnik« in »arhivske povezave« ter preučujejo njuno uporabo v okviru sistema odprtih arhivskih informacijskih sistemov (OAIS). Tehnologije verižnega bloka (blockchain) in razpršene evidence (DLT) analizirajo kot mehanizme za ustvarjanje digitalnih izvirnikov ter zagotavljanje integritete, preglednosti in zaupanja. Članek raziskuje tudi vlogo umetne inteligence (AI) – vključno z velikimi jezikovnimi modeli (LLM) – v arhivskih procesih in pri podpori ohranjanja na podlagi tehnologije veriženja blokov. Semantični splet in povezani podatki so predstavljeni kot orodja za izboljšanje medsebojne povezljivosti in odkrivanja digitalnih izvirnikov, medtem ko se pametne pogodbe pa so obravnavane v kontekstu upravljanja pravic.

Namen članka je predstaviti te koncepte in tehnologije ter oceniti njihov vpliv na arhiviranje digitalnih izvirnikov, pri čemer izpostavlja ključne prednosti in omejitve. Poudarja medsebojno povezanost arhivske teorije in tehnoloških inovacij ter nakazuje smernice za prihodnje raziskave in praktično izvajanje.

Ključne besede:

arhivi, umetna inteligenca, veriženje blokov, digitalni izvirnik, semantične tehnologije

1 Introduction

1.1 The I.Trust.Original project

The general digital transformation of society increases the need to work with digital originals and, consequently, to preserve them in accordance with archival rules and retention periods. Advanced qualified digital signatures and seals are widely accepted and guarantee the authenticity and trustworthiness of documents in everyday business operations. However, to determine the signatory's identity, it relies on signature certificates whose validity period is significantly shorter than the prescribed retention period in the digital archive. The current solution to this problem is to add new digital signatures or time stamps. This approach only temporarily addresses the issue. The research project “Long-term preservation of the authenticity of digital originals using blockchain technology and artificial intelligence” (2025–2029), or the I.Trust.Original project for short, funded by the European Union – NextGenerationEU fund, explores ways to extend the validity period of digital originals even after the expiration of the mechanisms that ensure their authenticity, although this period remains significantly shorter than that prescribed by law regulating retention periods.

The project is part of a wider landscape of international research initiatives addressing the challenges of digital preservation, authenticity, and trust in digital environments. It builds upon foundational work such as the “International Research on Permanent Authentic Records in Electronic Systems (InterPARES)” projects (1999–2026), which established key principles for preserving authentic digital records. More specifically, the I.Trust.Original project continues research initially undertaken during the InterPARES Trust (ITrust; 2013–2019) and InterPARES Trust AI (ITrustAI; 2021–2026) projects. It also extends the research of the four consecutive one-year projects “Application of blockchain technology in heritage environment (NFT@FFZG)” (2022–2025), which focused on a specific form of digital originals – the unique, one-original-only form.

The I.Trust.Original project, through complex simulations investigating the use of blockchain technology to preserve the trustworthiness of digital originals and the application of artificial intelligence (AI) algorithms to automate their classification, which considers special archival classifications (archival bond) once ingested into the digital archive, seeks to advance the integration of archival theory with emerging technologies. Despite its focus on technological trust, the project recognises that trust in digital records arises not only from innovative technological mechanisms but also from their alignment with archival principles such as authenticity, integrity, provenance, and contextual relationships.

1.2 Aim of the I.Trust.Original project

The I.Trust.Original project aims to explore the possibilities offered by emerging digital technologies such as blockchain and artificial intelligence in the context of the long-term preservation of the trustworthiness of digital originals. The project divides its central research problem into three parts, reflecting the different approaches required to address each aspect. These three parts can also be explored simultaneously, as they are not sequential.

The first part of the project aims to investigate the applicability of blockchain technology in ensuring the long-term preservation of the trustworthiness of digital originals. The second part explores the potential of artificial intelligence in establishing the archival bond. The third part aims to examine the feasibility of creating unique digital originals and to develop a model for their long-term preservation.

The defined project aim leads to the following fundamental research questions:

1. How can blockchain technology be applied to the long-term preservation of the trustworthiness of digital originals without the need for their periodic (re)signing?
2. How can machine learning and artificial intelligence methods be used to classify and (semi-)automatically establish an archival bond between related documents?
3. What is the optimal model for the creation, description, and long-term preservation of unique digital originals?
4. What are the security, legal, ethical and technical advantages and limitations of the proposed innovative solutions compared to traditional digital archiving methods?

By uniting the aim and research questions, the I.Trust.Original project positions itself at the intersection of archival theory and technological innovation. It builds on existing work by evaluating emerging technologies and embedding them within a coherent conceptual framework centred on the digital original. In this way, the project seeks to contribute to the development of new models and practices for ensuring the long-term trustworthiness of digital records in increasingly complex and dynamic digital environments.

1.3 Research activities

This paper addresses the first activity of the I.Trust.Original project: “Research on the impact of technological development on the trustworthiness of digital sources”, and presents interim results after a six-month period. This activity is planned to continue

throughout the four-year project. According to the project plan, research begins with the collection and analysis of relevant sources. Most of the work in this activity occurs during the first year. During the remaining period, the activity will continue at reduced intensity to allow ongoing monitoring of technological developments. Key topics include laws, bylaws, and regulations governing archival materials. The research also examines the long-term preservation of authentic documents. The analysis will consider archival theory related to the preservation of digital documents that are authentic, reliable, usable, accurate, and legally valid as evidence in court. The research will also examine models for creating and managing digital originals, such as paper documents and deeds issued as single originals. Practices unique to the digital environment will also be analysed. The study will further explore the relationship between archival concepts and transformative or disruptive technologies. A significant part of the research focuses on distributed systems, blockchain, and related distributed ledger technologies (DLT). Another area is the application of artificial intelligence (AI) in the creation, maintenance, and use of digital documents and data. The archival description process and the Records in Contexts (RiC) standard will be examined. This standard enhances information about archival materials by using semantic web concepts, linked data, and automated reasoning. The potential for integrating this standard with blockchain technology, AI, and other transformative technologies will be assessed. A substantial part of the initial research phase, and the project as a whole, will examine the concept of the archival bond. This will be discussed in detail later. The establishment and maintenance of the archival bond will be studied in relation to both paper-based concepts and diplomatics, with cutting-edge technologies such as distributed systems, linked data, and AI also considered. Existing digitisation process models will be analysed, including the model from the “Guidelines for the Digitisation of Documentary and Archival Materials” published by the Croatian State Archives. Other models and optimisation strategies will also be considered.

Based on initial findings, a digitisation and maintenance process for digitally born documents will be developed, integrating new principles and technologies to produce trustworthy digital originals eligible for acceptance as evidence. The analysis will inform the selection of types, formats, and documents for testing trustworthy digitisation and for training AI in archival bond research. The research will involve graduate (master’s) and PhD students, who will collaboratively investigate specific project segments for their theses.

2 Research methods

2.1 Literature review

An extensive, comprehensive collaborative literature review is being conducted due to the complexity of the research topic and the project's interdisciplinary nature. It is used to systematically identify existing knowledge by examining scholarly publications, legal frameworks, policies, standards, and other relevant sources to establish a theoretical and empirical context. This process involves the application of clearly defined selection criteria and analytical techniques, enabling not only the organisation of existing knowledge but also critical evaluation of different perspectives. In doing so, it will enable the identification of gaps in current knowledge, the mapping of key concepts, the contrasting of perspectives, and the identification of emerging trends (vom Brocke et al., 2009; Snyder, 2016).

In the context of research on digital records and archival practices, the literature review is particularly important due to the rapidly evolving nature of technologies and the interdisciplinary scope of the field. It enables the integration of insights from archival

theory, information science, law, and computer science, thereby supporting a comprehensive understanding of complex issues such as authenticity, trustworthiness, and long-term preservation. Additionally, it will provide a foundation for assessing how emerging technologies interact with established archival principles.

Furthermore, a longitudinal approach to the literature review, where sources are continuously monitored and incorporated into the research over time, enables project activities to remain responsive to technological and regulatory developments. As a result, the literature review will function not only as an initial research phase but also as an ongoing methodological component that informs and supports all current and subsequent stages of the project.

2.2 Analysis

Analysis of the relevant literature will be conducted by combining document analysis, conceptual and comparative analysis, and framework-based evaluation to address the various approaches and the interdisciplinary scope of the project. Researchers on the project will thoroughly review theoretical and professional literature. In addition to theoretical and professional papers and larger texts, during the document analysis, researchers will conduct the systematic examination of legal and regulatory frameworks, professional standards, and technical specifications relevant to archival practice and digital preservation. This includes analysis of national and international legislation, institutional guidelines, and standards governing digital preservation, metadata, and trust services.

Conceptual analysis will further enable critical examination of fundamental archival and technological concepts, including provenance, archival bond, and digital originals. This approach will clarify terminology, identify conceptual shifts, and evaluate how traditional archival principles are applied in digital contexts. Comparative analysis will assess different models, standards, and technological approaches, highlighting their similarities, differences, strengths, and limitations. It will also support evaluation of various frameworks and solutions across domains such as digital preservation, metadata management, and distributed systems.

Framework-based evaluation will further support the analysis by systematically examining established models and standards, such as archival description frameworks and digital preservation architectures, in relation to emerging technological environments. The interdependencies between archival, legal, and technological domains will also be examined. By exploring how these domains intersect and interact, the research will highlight the challenges of ensuring trustworthiness in digital records.

Finally, emerging technologies will be analysed, including blockchain, NFTs, artificial intelligence, and semantic web solutions, with an assessment of their applicability, potential benefits, and limitations within the archival context. Together, these analytical methods will enable a comprehensive and nuanced examination of the research problem, supporting the development of informed models and recommendations.

3 Project activity initiation and research setup

3.1 Systematisation of concepts and technologies

The I.Trust.Original project addresses several research areas: legal and regulatory frameworks, archival concepts, archival processing (archival description), digitisation, digital preservation, digital originals, metadata, distributed systems, trust and authenticity concepts, AI, description logics, semantic web and linked data, as well as other emerging technologies. Some research areas overlap. For example, archival description intersects with archival concepts (such as provenance and the archival bond), metadata, description logics, the semantic web, and linked data. The digital originals area overlaps with digital preservation concepts, distributed systems, trust and authenticity mechanisms, AI, and other emerging technologies. Legal and regulatory frameworks intersect with nearly all other areas. First, these areas will be studied individually. Next, an analysis of their interconnections and more complex concepts will follow.

3.1.1 Legal and regulatory frameworks

Researchers will thoroughly analyse Croatian archival laws and other relevant Croatian laws and bylaws, if necessary. They will also study Croatian specifications from the Government and the Croatian State Archives, EU regulations, and international recommendations and standards for all research areas. For example, in the analysis of trust and authenticity, the eIDAS regulation and relevant ETSI standards will be examined. For digitalisation models, the Ordinance on the management of documentary materials held outside state archives and the Guidelines for the digitisation of documentary and archival materials will be reviewed. The entire ISO OAIS standards suite, along with various metadata standards, will be examined for digital preservation concepts and solutions. Metadata standards will also be considered when handling metadata and paradata activities. Legal and ethical considerations will include studies on AI and emerging technologies. The aforementioned regulations and standards are only the beginning.

3.1.2 Archival concepts

Researchers will scrutinise “hardcore” archival concepts of provenance (Bearman & Lytle, 1985), related principles (Zhang, 2012), context, documents, metadata, paradata (Andersson, Huvila & Sköld, 2024), and archival bond (Duranti, 1997; Cameron, Franks & Hamidzadeh, 2023; Xie, 2024, p. 42; Thibodeau, Richmond & Beauchamp, 2025, p. 6046) within today’s archival information environment. The functioning of the provenance concept will be examined, and its semantic shift from a single-hierarchy reflection of one creator’s documentary legacy to a complex network of broadly described ambient contexts (Hurley, 2005), or the various contexts in which archival materials are involved, will be considered. The concept of the document will be tested to ensure its functionality in both paper and digital environments, as well as its relevance to today’s social information sphere. The concepts of subsidiary information created and used to control and maintain archival content – metadata and paradata – will be analysed and organised during this phase of the project. This new knowledge will be used to update the metadata set for the theoretical TrustChain model developed during the InterPARES Trust project, aligning it with current technological and scientific developments. An initial investigation by the associated researchers from Brazil examined whether and how the TrustChain model could be adopted in its current form, focusing on Brazilian institutions and the model’s alignment with the two main document management frameworks in Brazil.

Regulatory compliance analysis and strategic evaluation of the TrustChain architecture were conducted against the requirements of e-ARQ Brasil (V2.0) and MoReq-Jus (2nd Edition). Its complementarity with the OAIS model and RDC-Arq was also studied, with a focus on mitigating long-term risks (De Sordi & Vasco, 2026). Finally, archival bond, a linkage among archival materials related to the same administrative process, linking various files that need to be associated, and the role of "relationship decider-agents" will be analysed for application to digital materials (Xie, 2024, pp. 42–45).

3.1.3 Archival processing

Special attention will be given to the study of archival description, leading to improved use and reproducibility of archival findings. This component is linked to themes such as description logics, semantic web, and linked data. The research will focus on the Records in Contexts (RiC) standard, which will be applied to descriptive practices aligned with emerging semantic web technologies. It will also assist researchers in clarifying concepts and understanding the applicability of its ontology and the embedded descriptive logic framework. This component may be used alongside the AI component. For example, it could facilitate communication between users (in natural language) and triple stores containing archival knowledge, by translating user language into SPARQL queries and enabling semantic retrieval of archival materials. Researchers will review digitisation and tokenisation in archival processing – the conversion of analogue recordkeeping and archival materials into digital formats for business, use, and preservation purposes, along with the tokenisation of archival and museum digital surrogates and digital originals. They will examine the results of the tokenisation process within heritage institutions (Stančić et al., 2025) and assess their applicability for long-term preservation strategies.

3.1.4 Digital preservation concepts

Digital preservation concepts will be investigated, ranging from the OAIS reference model to producer-archive interaction standards such as PAIS and its methodological counterpart PAIMAS, as well as data and metadata packaging frameworks (e.g., XFDU). The OAIS standard was recently updated, so its changes will be considered. The OAIS-derivative specification for the distributed preservation environments (PaaST) will also be examined in the light of new developments in the blockchain domain. These standards will significantly contribute to establishing preservation-safe environments, which are part of the equation examined by this project. This component will also include European-based specifications of information packages derived from the eArk project (Schlarb et al., 2016).

3.1.5 Digital signatures

Digital signatures and their validity during long-term digital preservation are the focal point of the I.Trust.Original project. The regulations and standards governing the creation of trust services will be investigated alongside standards for the creation and certification of trusted digital repositories. The standards and technologies relevant to the realisation of qualified e-signatures – public key infrastructure (PKI), digital certificates, and digital timestamps – will be assessed in the context of preserving trustworthy (reliable, authentic, and accurate) digital records.

3.1.6 Digital originals

Researchers will dedicate themselves to studying the concepts of archival documents, including their elements of form(s), content, context, technological environment, action, agents involved, and archival bonds (Lemieux, 2024, p. 144). They will also examine their digital equivalents and other information objects that require long-term preservation. The concept of digital originals will be considered from the perspective of creating a digital original in a single copy. This form is used for documents such as promissory notes, bills of lading etc., where only one original needs to be created (Stančić, 2023), as well as in the context of NFT use in heritage institutions (Stančić et al., 2025; Bilogrivić & Stublić, 2023; Stublić, Bilogrivić & Zlodi, 2023).

3.1.7 Archival bond

The archival bond is a concept in archival science first theorised by Giorgio Cencetti (1939), who viewed archives as more than a simple collection of documents, but as an organic whole. Duranti (1997) defines the archival bond as “the network of relationships that each record has with the records belonging in the same aggregation”. The activity during which documents are produced can last for an extended period, which is significant because the archival bond remains mutable until the activity it is part of concludes.

This crucial concept is central to preserving the context and interrelationships of digital records, ensuring their functionality and trustworthiness in a digital environment that differs greatly from traditional archives. The preservation of the archival bond in traditional archives is closely tied to two fundamental archival principles: *respect des fonds* and original order (Schellenberg, 1956; Duchein, 1983). *Respect des fonds* ensures that records created by the same institution, organisation, or individual are kept together and not mixed with those of other creators. Original order maintains the internal arrangement established during the records’ active use. Together, these principles safeguard the relationships among records, ensuring that their functional and procedural contexts remain intact (Duchein, 1983; Millar, 2017). Analogue records are fixed in form and location, whereas the digital environment is inherently dynamic, requiring relationships to be explicitly represented and maintained (Bearman, 1994; Duranti, 2010).

Concepts such as metadata schemas, relational databases, and linked data frameworks attempt to reconstruct or maintain these bonds by encoding relationships between records. For example, descriptive metadata can document the procedural context of a record, while system architectures can enforce connections between related digital objects. Nevertheless, these solutions are only as effective as the standards and practices that govern their implementation. The archival bond, therefore, remains both a conceptual and practical challenge, requiring not only technical infrastructure but also a deep understanding of the functions and activities that give rise to records (Duranti, 2010; Lemieux, 2016).

3.1.8 Metadata

The use, reusability, standards, and interoperability of metadata will also be examined. This aspect will be partly addressed in archival description-related research by examining the Records in Contexts standard. However, the research will also cover metadata standards such as METS (also addressed in the digital preservation concepts

component), PREMIS, EAD, and the metadata specification produced by the Croatian State Archives in 2022.

3.1.9 Trust and authenticity concepts

Blockchain and digital ledger technologies (DLT) will be assessed as potential technologies for implementation in the archival sector and for use in the long-term preservation of the validity of digitally signed records. Researchers will therefore examine various types of blockchains (Lemieux, 2017, pp. 2273–2276), their specific issues (Lemieux, 2024, p. 143), and their advantages to propose a suitable solution for recordkeeping services. The technical specification for electronic record management systems, mandatory in public institutions, anticipates reliance (Government of the Republic of Croatia, 2024, p. 32) on blockchain and distributed ledger technologies (Van Hijfte, 2026, p. 45), and the results of this project should support the overall Croatian digital strategy. This component is also related to the theme of distributed systems (Van Hijfte, 2026, p. 17).

3.1.10 Artificial intelligence

The research will include a significant AI component, with researchers analysing its applicability to other parts of the project. Models, processes, and prototypes will be deployed wherever feasible within the project's financial and time constraints to improve archival processes, particularly in relation to large language models (LLMs). Special attention will be given to the use of AI in (semi-)automating the establishment of archival bond.

3.1.11 Other emerging technologies

Researchers will examine new and emerging technologies, including post-quantum cryptography, cybersecurity, smart contracts, and novel interoperability frameworks, to assess their impact on the models being developed and tested, as well as their potential for incorporation into the proposed solutions.

3.2 Technological trust and its benefits for archival science

Digital preservation systems, typically aligned with the OAIS reference model, address the long-term storage, management, and accessibility of digital objects (International Organization for Standardization, 2025; Lavoie, 2014). These systems ensure bit-level integrity through fixity checks and redundancy, thus supporting one dimension of authenticity. Nevertheless, their reliance on institutional trust, procedural controls, and metadata limits their capacity to provide independent and persistent verification of authenticity. In particular, OAIS-based systems do not inherently resolve the challenge of maintaining the validity of digital signatures over time without reprocessing or intervention.

Metadata frameworks and semantic technologies further support the preservation of authenticity by documenting context, provenance, and relationships between records (Massari et al., 2025). Linked data and knowledge graph approaches extend this capability by enabling the reconstruction of the archival bond across distributed environments (McKenna, Debruyne & O'Sullivan , 2022). However, metadata remains inherently declarative and dependent on the trustworthiness of the managing institution. As such, it cannot, by itself, provide cryptographically verifiable guarantees of

authenticity or ensure that such guarantees remain valid over extended preservation periods.

Artificial intelligence and machine learning technologies enhance archival processes through automation, particularly in description, classification, preservation, and access (Colavizza et al., 2021; Hutchinson, 2019). While these tools improve efficiency and scalability, they do not address the fundamental issue of authenticity verification. Moreover, the opacity and probabilistic nature of AI-generated outputs may introduce additional challenges for evidential reliability, potentially undermining trust in archival metadata and processes.

In contrast, blockchain and distributed ledger technologies provide robust mechanisms for ensuring data integrity and tamper-evidence through cryptographic hashing and distributed consensus. However, these approaches often equate authenticity with immutability and data integrity verification, thereby overlooking the broader archival requirements for context, provenance, and archival bonds (Lemieux, 2016). As a result, they offer only a partial solution to the problem. Artificial intelligence is showing potential as a means of addressing the mutability of archival bond in relation to the immutability of blockchain. Employing a range of different digital technologies to address a single problem is often necessary (Rolan et al., 2018).

Digital signatures and trust services, particularly within the framework of the eIDAS regulation, establish legally recognised mechanisms for asserting authenticity at the time of record creation and validation (European Parliament & Council of the European Union, 2024). However, these mechanisms are inherently time-bound, as their validity depends on external infrastructures such as certification authorities and evolving cryptographic standards. Consequently, they do not guarantee the long-term, intervention-free verifiability required for archival preservation, highlighting a critical disconnect between legal validity and archival authenticity.

Other technological domains, including cloud computing and digital forensics, contribute to the preservation of digital records by enhancing storage scalability and enabling the capture of technical metadata and original context (Kirschenbaum et al., 2010). While these approaches support aspects of authenticity, such as fixity and provenance, they do not directly address the problem of persistent, independent verification of digitally signed records over time.

3.3 Open questions

Building on the TrustChain model developed during the InterPARES Trust project, the I.Trust.Original project aims to advance the model and complement it with an AI solution to (semi-)automate the establishment of archival bond. It remains to be seen how digitally signed records behave in a digital preservation environment after the expiration of the digital certificates on which they rely to confirm the identity of the signatory and the time the document was signed. Therefore, the newly established “Laboratory for Research of Advanced Technologies and Innovation (NEXTGEN-LAB)” will be equipped with a simulation environment to accelerate the ageing of digital signatures and digital timestamps. It will also provide an environment for research and development of blockchain and AI solutions. This will enable the research team to gain deeper insight into various aspects of the impact of trusted services relying on PKI during the long-term preservation of trustworthy digital records. This will be especially relevant for digital originals – documents created on-chain and relying on distributed consensus mechanisms for their trustworthiness.

4 Discussion

This project problematises several tensions: between traditional archival science and new technologies; between hierarchical archival stability and increasingly dynamic digital realities; and between known and established archival forms – such as documents, digital documents, and digital originals – and broader socially relevant information.

New technologies are transforming how archival documents and data are created, structured, and managed. Technology is not merely a fixed tool or a mutable characteristic of archival information – such as a medium assumed to be replaceable without consequence – but rather a force that shapes information itself (Barad, 2007). For this reason, technology should be incorporated into preservation strategies by design. However, technologies alone cannot solve problems that are only partly technological. Archival theory must also engage with these changes, and practice must respond accordingly.

How should this interconnectedness, interweaving of archival and technological concepts, work? Digital archives, for example, need to be reconfigured from simpler to more complex and capable OAIS versions to address the complications posed by new information types and technologies. They should preserve large volumes of dynamic information and documents and support our trust in these digital objects. Next, new processing techniques and the archival standard RiC should support contemporary practices by shifting from a simplified hierarchical archival representation towards the utilisation of more complex, networked forms of social information that better meet our business and social needs.

In the context of interconnectedness of archival theory and emerging technologies discussed in this paper, the archival bond concept, as another example, takes on new dimensions. Blockchain and distributed ledger technologies (DLT) can serve as mechanisms for creating digital originals and ensuring data integrity, transparency, and trust in digital records (Zheng et al., 2018; Lemieux, 2016). Similarly, artificial intelligence can help recognise implicit relationships within large datasets. However, neither technology can fully substitute for the archival principle that records must be comprehended within their functional and procedural context (Duranti, 2010). Preserving the archival bond in digital environments ultimately depends on integrating archival knowledge with technologies. This makes sure that relationships between records remain trustworthy and meaningful over time. This example illustrates the problem's complexity and shows that integrating archival theory with current technologies requires mutual adjustments. There are more such examples, which the iTrust.Original project will examine.

5 Conclusion and further work

Digitally signed documents lose their validity once the digital certificates they depend on expire. This presents a serious challenge for the long-term preservation of such records in digital archives. Blockchain and distributed ledger technologies offer a promising solution for registering the validity of signatures on-chain while they are still valid, as well as for recording certificate chains to establish the validity of already expired digital signatures retrospectively. Blockchain is an immutable structure, and once information is registered, it cannot be altered. Therefore, it is not an adequate solution for maintaining mutable information about the archival bond. Artificial intelligence offers a promising approach for (semi-)automating the establishment of an archival bond

between documents already registered on the blockchain. The concept of digital original documents created simultaneously as human-readable and machine-readable documents, confirming the uniqueness of a single digital original and distinguishing it from all identical digital copies, offers a promising solution for both the digitisation of paper documents created as single originals and the creation of digital art, music, video, and other copyrighted works in digital form. The solutions for preserving digital information should be sought in the interplay between archival-theoretical and emerging technological concepts.

Acknowledgements: This paper was written as part of the institutional research project "Long-term preservation of the authenticity of digital originals using blockchain technology and artificial intelligence (I.Trust.Original)" at the University of Zagreb, Faculty of Humanities and Social Sciences; funded by the European Union – NextGenerationEU (2025–2029). It relates to the project activity "Research on the impact of technological development on the trustworthiness of digital sources".

6 Literature

Andersson, L., Huvila, I., & Sköld, O. (2024). An Introduction to Paradata. In I. Huvila, L. Andersson, & O. Sköld (Eds), Perspectives on Paradata (Vol. 13, pp. 1–14). Springer International Publishing. https://doi.org/10.1007/978-3-031-53946-6_1.

Barad, K. M. (2007). Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning. Duke University Press.

Bearman, D. A. (1994). Managing Electronic Mail. Archives & Manuscripts, 22(1), 28–50.

Bearman, D. A., & Lytle, R. H. (1985). The Power of the Principle of Provenance. Archivaria, 21. https://archivaria.ca/index.php/archivaria/article/view/11231.

Bilogrivić, M., & Stublić, H. (2023). Application of Blockchain Technology and NFTs in a Museum Environment. 2023 46th MIPRO ICT and Electronics Convention (MIPRO), 1353-1358. https://doi.org/10.23919/MIPRO57284.2023.10159780.

Brocke, J. vom, Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). Reconstructing the Giant: On the Importance of Rigour in Documenting the Literature Search Process. ECIS 2009 Proceedings. 161. https://aisel.aisnet.org/ecis2009/161.

Cameron, S., Franks, P., & Hamidzadeh, B. (2023). Positioning Paradata: A Conceptual Frame for AI Processual Documentation in Archives and Recordkeeping Contexts. https://doi.org/10.14288/1.0438178.

Cencetti, G. (1939). Il fondamento teorico della dottrina archivistica [The Basic Theory of Archival Science]. Archivi, 6(1), 7-13.

Colavizza, G., Blanke, T., Jeurgens, C., & Noordegraaf, J. (2022). Archives and AI: An Overview of Current Debates and Future Perspectives. Journal on Computing and Cultural Heritage, 15(1), 1–15. https://doi.org/10.1145/3479010.

Consultative Committee for Space Data Systems. (2014). Producer-Archive Interface Specification (CCSDS 651.1-B-1). CCSDS. https://ccsds.org/publications/bluebooks/entry/3044/ .

Consultative Committee for Space Data Systems. (2004). Producer-Archive Interface Methodology Abstract Standard (PAIMAS) (CCSDS 651.0-M-1). CCSDS. https://ccsds.org/publications/magentabooks/entry/3040/.

Croatian State Archives. (2022). Specifikacija metapodaraka za opis dokumentarnog i arhivskog gradiva. (Guidelines for the Digitisation of Documentary and Archival Materials). https://www.arhiv.hr/Portals/0/Specifikacija.pdf.

De Sordi, N. A. D., & Vasco, J. A. C. (2026). Trustchain e a preservação digital de longo prazo: Análise de aderência aos modelos de requisitos e-Arq Brasil, MoReq-Jus e ao modelo OAIS. Revista Brasileira de Preservação Digital, 7, e026006. https://doi.org/10.20396/rebpred.v7i00.21102.

Duchein, M. (1983). Theoretical Principles and Practical Problems of Respect des fonds in Archival Science. Archivaria, 16, 64–82.

Duranti, L. (1997). The Archival Bond. Archives and Museum Informatics, 11(3–4), 213–218. https://doi.org/10.1023/A:1009025127463.

Duranti, L. (2010). Concepts and Principles for the Management of Electronic Records, or Records Management Theory is Archival Diplomatics. Records Management Journal, 20(1), 78–95. https://doi.org/10.1108/09565691011039852.

European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1183 of 11 April 2024 Amending Regulation (EU) No 910/2014 as Regards Establishing the European Digital Identity Framework. EUR-Lex. https://eur-lex.europa.eu/eli/reg/2024/1183/oj/eng.

Government of the Republic of Croatia. (2024, October 22). Tehnička specifikacija informacijskog sustava elektroničkog uredskog poslovanja (Electronic Record Management Systems Technical Specification) [Specifikacija]. MPUDT. https://mpudt.gov.hr/o-ministarstvu/ustrojstvo-6300/uprava-za-politicki-sustav-i-opcu-upravu/uredsko-poslovanje-26231/26231.

Hurley, C. (2005). Parallel provenance. Archives & Manuscripts, 33(2), 52–91.

Hutchinson, T. (2020). Natural Language Processing and Machine Learning as Practical Toolsets for Archival Processing. Records Management Journal, 30(2), 155–174. https://doi.org/10.1108/RMJ-09-2019-0055

International Council of Archives. (2023). Records in Contexts – Conceptual model (Version 1.0.). https://www.ica.org/app/uploads/2023/12/RiC-CM-1.0.pdf

International Organization for Standardization. (2010). Space Data and Information Transfer Systems – XML Formatted Data Unit (XFDU) Structure and Construction Rules (ISO 13527:2010). https://www.iso.org/standard/53985.html

International Organization for Standardization. (2025). Space Data System Practices – Reference model for an open archival information system (OAIS) (ISO 14721:2025). https://www.iso.org/standard/87471.html

Kirschenbaum, M. G., Ovenden, R., Redwine, G., & Donahue, R. (2010). Digital Forensics and Born-digital Content in Cultural Heritage Collections. Council on Library and Information Resources.

Lavoie, B. (2014). The Open Archival Information System (OAIS) Reference Model: Introductory Guide (2nd Edition). DPC Technology Watch Report 14-02. York, UK: Digital Preservation Coalition. https://dx.doi.org/10.7207/twr14-02.

Lemieux, V. L. (2016). Trusting Tecords: Is Blockchain Technology the Answer? Records Management Journal, 26(2), 110–139. https://doi.org/10.1108/RMJ-12-2015-0042

Lemieux, V. L. (2017). A Typology of Blockchain Recordkeeping Solutions and Some Reflections on Their Implications for the Future of Archival Preservation. 2017 IEEE International Conference on Big Data (Big Data), 2271–2278. https://doi.org/10.1109/BigData.2017.8258180

Lemieux, V. L. (2024). Blockchain: The Theory and Practice of Trustworthy Archival Documents. In Archival Science in Interdisciplinary Theory and Practice. Rowman & Littlefield.

Massari, A., Peroni, S., Tomasi, F., & Heibi, I. (2026). Representing Provenance and Track Changes of Cultural Heritage Metadata in RDF: A survey of existing approaches. Digital Scholarship in the Humanities, 41(Supplement_1), i196-i212. https://doi.org/10.1093/llc/fqaf076

McKenna, L., Debruyne, C., & O'Sullivan, D. (2022). Using Linked Data to Create Provenance-Rich Metadata Interlinks: The Design and Evaluation of the NAISC-L Interlinking Framework for Libraries, Archives and Museums. AI & SOCIETY, 37(3), 921-947. https://doi.org/10.1007/s00146-021-01373-z

Millar, L. (2017). Archives: Principles and Practices (Second edition). Facet Publishing.

Rolan, G., Humphries, G., Jeffrey, L., Samaras, E., Antsoupova, T., & Stuart, K. (2019). More human than human? Artificial Intelligence in the Archive. Archives and Manuscripts, 47(2), 179–203. https://doi.org/10.1080/01576895.2018.1502088.

Schellenberg, T. R. (1956). Modern Archives: Principles and Techniques. SAA. https://babel.hathitrust.org/cgi/pt?id=mdp.39015071452539&seq=10&q1=manager.

Schlarb, S., Schmidt, R., Karl, R., Bartha, M., Rorden, J., Delve, J., & Aas, K. (2016). An OAIS-oriented System for Fast Package Creation, Search, and Access. iPRES 2016 - Proceedings of the 13th International Conference on Digital Preservation, 85–95. https://phaidra.univie.ac.at/o:502851.

Snyder, H. (2019). Literature Review as a Research Methodology: An Overview and Guidelines. Journal of Business Research, 104, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Stančić, H. (2023): Digitalizacija i upravljanje digitalnim izvornicima. Zbornik 54. savjetovanja hrvatskih arhivista: Utjecaj digitalizacije na arhivsku praksu, 7-30. https://urn.nsk.hr/urn:nbn:hr:131:583191.

Stančić, H., Trbušić, Ž., & Rajh, A. (2025). Model of Application of Blockchain Technology FTand Nfts in Heritage Environment. 2025 IEEE International conference on cyber humanities (IEEE-CH), 1–6. https://doi.org/10.1109/IEEE-CH65308.2025.11279589.

Stublić, H., Bilogrivić, M., & Zlodi, G. (2023). Blockchain and NFTs in the Cultural Heritage Domain: A Review of Current Research Topics. Heritage, 6(4), 3801-3819. https://doi.org/10.3390/heritage6040202.

Thibodeau, Kenneth, et al. (2018). Preservation as a Service for Trust (PaaST): Functional and Data Requirements for Digital Preservation – version 1.0. https://interparestrust.org/assets/public/dissemination/PreservationasaServiceforTrust1_0-FINAL.pdf.

Thibodeau, K., Richmond, A., & Beauchamp, M. (2025). Archival Research Theory: Putting Smart Technology to Work for Researchers. 2025 IEEE International Conference on Big Data (BigData), 6044–6051. https://doi.org/10.1109/BigData66926.2025.11401002.

Van Hijfte, S. (2026). Blockchain Platforms: A Look at the Underbelly of Distributed Platforms. Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-00979-1.

Xie, S. (2024). Digital Archival Bond: A Universal Bond for Digital Trustworthiness. In Archival Science in Interdisciplinary Theory and Practice. Rowman & Littlefield.

Zhang, J. (2012). Original Order in Digital Archives. Archivaria, 74 (Special issue on arrangement and description of archives), 167–193.

Zheng, X., Mukkamala, R. R., Vatrapu, R., & Ordieres-Mere, J. (2018). Blockchain-based Personal Health Data Sharing System Using Cloud Storage. 2018 IEEE 20th International Conference on E-Health Networking, Applications and Services (Healthcom), 1–6. https://doi.org/10.1109/HealthCom.2018.8531125.

POVZETEK

SODOBNI PRISTOPI K ARHIVIRANJU DIGITALNIH IZVIRNIKOV: PREGLED KONCEPTOV IN TEHNOLOGIJ

Dr. Arian Rajh
Izredni profesor, Filozofska fakulteta Univerze v Zagrebu, Hrvaška
arajh@ffzg.hr

Dr. Hrvoje Stančić
Redni profesor, Filozofska fakulteta Univerze v Zagrebu, Hrvaška
hstancic@ffzg.hr

Jelena Bolkovac
Fakulteta strojništva in ladjedelništva, Univerza v Zagrebu, Hrvaška
jelena.bolkovac@fsb.unizg.hr

Prispevek se nanaša na nalogo »Raziskava vpliva tehnološkega razvoja na verodostojnost digitalnih virov« v okviru projekta I.Trust.Original. Projekt financira sklad Evropske unije NextGenerationEU (2025–2029), izvajal pa se bo na Filozofski fakulteti Univerze v Zagrebu.

Ker postaja digitalno gradivo vse pogostejša oblika ustvarjanja zapisov, se arhivi soočajo z vedno več izzivi, povezanimi z verodostojnostjo, avtentičnostjo, celovitostjo in dolgoročnim ohranjanjem tega gradiva. Prispevek obravnava koncepte in tehnološke pristope k arhiviranju digitalnih zapisov ter analizira vpliv sodobnih tehnologij na verodostojnost in dolgoročno dostopnost digitalnih zapisov.

Avtorji razpravljajo o konceptu digitalnega izvirnika – ali ta sploh lahko obstaja ali pa gre s tehnološkega vidika vedno zgolj za kopijo oziroma reprezentacijo? Nadalje preučujejo uporabo konceptov iz referenčnega modela odprtega arhivskega informacijskega sistema (OAIS), ki predstavlja uveljavljen okvir za digitalno hrambo, v povezavi s konceptom digitalnega izvirnika. Analizirajo tehnologijo verige blokov (blockchain) in tehnologijo razpršene evidence (DLT – distributed ledger technology) kot možna mehanizma za ustvarjanje digitalnih izvirnikov in zagotavljanje celovitosti podatkov, preglednosti ter verodostojnosti za vse vrste digitalnih zapisov.

Glede na velik obseg zajetega digitalnega gradiva preučuje prispevek tudi vlogo umetne inteligence (UI) – vključno z velikimi jezikovnimi modeli (LLM) – v postopkih popisovanja, vrednotenja, dolgoročnega ohranjanja digitalnih izvirnikov in omogočanja dostopa do njih. UI je obravnavana tudi kot možna podpora tehnologiji veriženja blokov pri izvajanju digitalne hrambe.

Tehnologije semantičnega spleta in povezanih podatkov (linked data) so predstavljene kot orodja za izboljšanje interoperabilnosti, kontekstualizacije in odkrivanja digitalnih izvirnikov v različnih sistemih in institucijah. Nazadnje je obravnavana uporaba pametnih pogodb na področju upravljanja pravic, kar skupaj s prej omenjenimi tehnologijami ustvarja sodoben tehnološki ekosistem arhivov.

Avtorji si prizadevajo pojasniti vse te koncepte in tehnologije ter prikazati njihov vpliv na arhiviranje digitalnih izvirnikov, pri čemer izpostavljajo tako njihove prednosti kot tudi omejitve znotraj arhivskega konteksta. Poudarjajo potrebo po povezovanju arhivske

teorije in tehnoloških inovacij za napredek digitalnega arhiviranja ter opozarjajo na etične vidike in institucionalno odgovornost. Z opredelitvijo ključnih trendov, izzivov in odprtih vprašanj nakazuje prispevek smeri prihodnjih raziskav in praktičnih rešitev na področju digitalnega arhiviranja.

About the authors:

Arian Rajh, PhD, is an associate professor of archival science at the Faculty of Humanities and Social Sciences, University of Zagreb, Croatia. He received his PhD in 2010 from the University of Zagreb. His research interests include archival theory, description, data preservation, metadata, and the semantic web. A certified professional archivist with 20 years of experience, he specialises in contemporary archival materials. He has led numerous national and EU projects, including digitisation, digital dossiers, ISO OAIS systems, research projects such as InterPARES, and blockchain/NFT in heritage. He teaches core archival science courses, has published over 75 papers, and contributes to archival legislation in Croatia.

Hrvoje Stančić, PhD, is a tenured professor of archival science at the Faculty of Humanities and Social Sciences, University of Zagreb, Croatia. He received his PhD in 2006 from the University of Zagreb. His research interests include digital preservation, the trustworthiness of digital records, and the application of blockchain technology and artificial intelligence in archives. He is a researcher and a member of the Executive Committee of the InterPARES Trust AI project (2021-2026) and principal investigator at the I.Trust.Original project (2025-2029).

Jelena Bolkovac is a senior librarian at the Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia. She is currently a PhD student at the Faculty of Humanities and Social Sciences, University of Zagreb, Croatia, in the field of Information and Communication Sciences. Her research interests include bibliometrics, AI, and open science.

O avtorjih:

Dr. Arian Rajh je izredni profesor arhivistike na Filozofski fakulteti Univerze v Zagrebu. Doktoriral je leta 2010 na isti univerzi. Njegovi raziskovalni interesi vključujejo arhivsko teorijo, popisovanje, hrambo podatkov, metapodatke in semantični splet. Je arhivist z dvajsetletnimi izkušnjami in strokovnjak za sodobno arhivsko gradivo. Vodil je številne nacionalne in evropske projekte, med njimi digitizacijo, digitalne spisovne zbirke, sisteme ISO OAIS, raziskovalne projekte, kot je InterPARES, in projekte s področja tehnologije veriženja blokov/nezamenljivih žetonov (blockchain/NFT) v kulturni dediščini. Predava temeljne arhivske predmete in je avtor več kot 75 znanstvenih objav.

Dr. Hrvoje Stančić je redni profesor arhivistike na Filozofski fakulteti Univerze v Zagrebu. Doktoriral je leta 2006. Njegovi raziskovalni interesi vključujejo digitalno hrambo, verodostojnost digitalnih zapisov in uporabo tehnologije veriženja blokov (blockchain) ter umetne inteligence v arhivih. Je raziskovalec in član izvršnega odbora projekta InterPARES Trust AI (2021–2026) ter vodja projekta I.Trust.Original (2025–2029).

Jelena Bolkovac je višja knjižničarka na Fakulteti strojništva in ladjedelništva Univerze v Zagrebu. Trenutno je doktorska študentka na Filozofski fakulteti Univerze v Zagrebu na področju informacijskih in komunikacijskih znanosti. Njena raziskovalna področja vključujejo bibliometrijo, umetno inteligenco in odprto znanost.