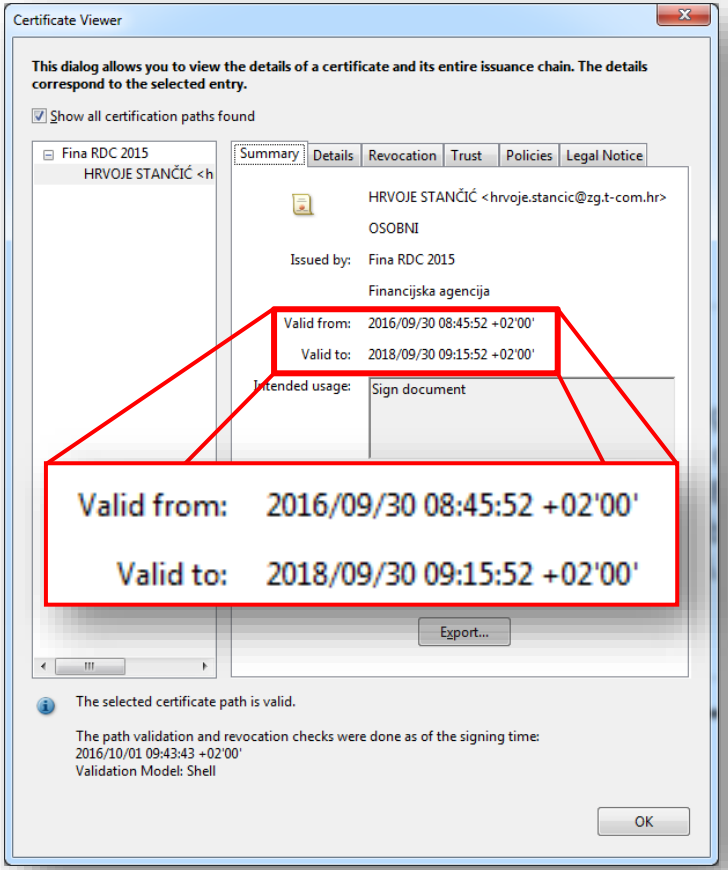
Contemporary approaches to archiving of digital originals: a review of concepts and technologies

Dr. **Arian Rajh**, associate professor
Faculty of Humanities and Social Sciences, Zagreb - arajh@m.ffzg.unizg.hr
Dr. **Hrvoje Stančić**, professor
Faculty of Humanities and Social Sciences, Zagreb - hstancic@m.ffzg.unizg.hr
Jelena Bolkovac, senior librarian, PhD candidate
Faculty of Mechanical Engineering and Naval Architecture, Zagreb - jelena.bolkovac@fsb.unizg.hr

Contents

1. Introduction

- Digital transformation of society – digital originals
- Digital signatures and seals
  - authenticity
  - trustworthiness
  - issue with validity period



2. Project I.Trust.Original

- Long-term preservation of the authenticity of digital originals using blockchain technology and artificial intelligence (2025-2029) - **I.Trust.Original**
  - funded by the European Union – NextGenerationEU fund
- Explores ways to extend the validity period of digital originals
- Harmonises technological concepts and archival principles
- 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

2. Project I.Trust.Original ...

- 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?

2. Project I.Trust.Original ...

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?

➢ Project positions itself at the intersection of **archival theory and technological innovation**

3. Research activities

- Research on the impact of technological development on the trustworthiness of digital sources
  - first activity of the project I.Trust.Original
  - collection and analysis of relevant sources
  - key topics include laws, bylaws, and regulations governing archival materials, as well as specifications and literature
  - 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

4. Systematisation of concepts and technologies

- Legal and regulatory frameworks
  - Croatian archival laws, Croatian specifications from the Government and the Croatian State Archives, EU regulations, and international recommendations and standards for all research areas
- Archival concepts
  - Provenance, related principles, context, documents, metadata, paradata, and archival bond

4. Systematisation of concepts and technologies ...

- Archival processing
  - Archival description – Records in Contexts (RiC) standard
- Digital preservation concepts
  - OAIS reference modes, PAIS, PAIMAS, data and metadata packaging frameworks (e.g. XFDU), specification for distributed preservation environments, and European-based specifications of information packages derived from the eArk project

4. Systematisation of concepts and technologies ...

- Digital signatures
  - regulations and standards governing the creation of trust services, standards for the creation and certification of trusted digital repositories, standards and technologies relevant to the realisation of qualified e-signatures
- Digital originals
  - concepts of archival documents, including their elements of form(s), content, context, technological environment, action, agents involved, and archival bonds
  - digital equivalents and other information objects that require long-term preservation

4. Systematisation of concepts and technologies ...

- Archival bond
  - conceptual and practical challenge
- Metadata
  - use, portability, reusability, standards, and interoperability
  - partly addressed in archival description-related research by examining the Records in Contexts standard
  - 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

4. Systematisation of concepts and technologies ...

- Trust and authenticity concepts
  - blockchain and digital ledger technologies (DLT)
- Artificial intelligence
  - applicability to other parts of the project.
  - models, processes, and prototypes will be deployed wherever feasible to improve archival processes, particularly in relation to large language models (LLMs)
- Other emerging technologies
  - new and emerging technologies, including post-quantum cryptography, cybersecurity, smart contracts, and novel interoperability frameworks

5. Technological trust and its benefits for archival science

- Digital preservation systems
  - typically aligned with the OAIS reference model
  - ensure bit-level integrity through fixity checks and redundancy, thus supporting one dimension of authenticity
- Metadata frameworks
  - support the preservation of authenticity by documenting context, provenance, and relationships between records
  - linked data and knowledge graph approaches enable reconstruction of the archival bond across distributed environments

5. Technological trust and its benefits for archival science ...

- Artificial intelligence and machine learning
  - enhance archival processes through automation, particularly in description, classification, preservation, and access
  - improve efficiency and scalability
- Blockchain and distributed ledger technologies
  - provide robust mechanisms for ensuring data integrity and tamper-evidence through cryptographic hashing and distributed consensus

5. Technological trust and its benefits for archival science ...

- Digital signatures and trust services
  - establish legally recognized mechanisms for asserting authenticity at the time of record creation and validation
- Other technological domains
  - include 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

6. Open questions

- 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
  - Laboratory for Research of Advanced Technologies and Innovation (**NEXTGEN-LAB**)
    - newly established
    - is being equipped with a **simulation environment to accelerate ageing** of digital signatures and digital timestamps

6. Open questions ...

- AI solution to (semi-)automate the establishment of archival bond
  - I.Trust.Original project aims to advance the TrustChain model developed during the InterPARES Trust project (2013-2019)
  - NEXTGEN-LAB
    - will provide an environment for research and development of blockchain and AI solutions
- Impact of trusted services relying on PKI during the long-term preservation of trustworthy digital records

7. Conclusion

- Blockchain and distributed ledger technologies
  - promising solution for registering the validity of signatures on-chain while they are still valid.
- Artificial intelligence
  - promising approach for (semi-)automating the establishment of an archival bond between documents already registered on the blockchain.
- Concept of digital original documents created simultaneously as human-readable and machine-readable documents
  - 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

8. Acknowledgements

- This presentation is based on the paper which 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.

Funded by the European Union

NextGenerationEU

9. 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
- 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

9. Literature ...

- 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.

9. Literature ...

- Massari, A., Peroni, S., Tomasi, F., & Heibi, I. (2026). Representing Provenance and Track Changes of Cultural Heritage Metadata in RDF: A survey
- 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

THANK YOU!

Contemporary approaches to archiving of digital originals: a review of concepts and technologies

Dr. **Arian Rajh**, associate professor
Faculty of Humanities and Social Sciences, Zagreb - arajh@m.ffzg.unizg.hr
Dr. **Hrvoje Stančić**, professor
Faculty of Humanities and Social Sciences, Zagreb - hstancic@m.ffzg.unizg.hr
Jelena Bolkovac, senior librarian, PhD candidate
Faculty of Mechanical Engineering and Naval Architecture, Zagreb - jelena.bolkovac@fsb.unizg.hr