

# Analyzing Knowledge Graph Innovations and Emerging AI technologies for Cultural Heritage Data Management

Milen Kraev<sup>[0009-0009-8973-7765]</sup>, Detelin Luchev<sup>[0000-0003-0926-5796]</sup>

Institute of Mathematics and Informatics, Bulgarian Academy of Sciences,  
8, G. Bonchev Str., Sofia, Bulgaria  
kraevmilen@gmail.com, dml@math.bas.bg

**Abstract.** Knowledge graphs have become a vital tool for structuring and accessing cultural heritage data, offering new ways to connect and interpret historical information. As the volume and complexity of heritage data grow, integrating advanced computational methods becomes essential to enhance their accuracy, usability, and accessibility. This paper explores emerging trends in cultural heritage knowledge graphs, focusing on the role of artificial intelligence and machine learning in improving entity recognition, contextualization, and knowledge enrichment. It examines how natural language processing and deep learning techniques can refine data interpretation and automate updates, leading to dynamic, self-sustaining knowledge graphs. Additionally, the study highlights the integration of knowledge graphs with immersive technologies such as augmented reality and virtual reality, which offer interactive ways to engage with heritage content. Finally, it discusses the impact of linked open data initiatives in fostering cross-institutional collaboration and global accessibility. These advancements collectively redefine how cultural heritage is studied and experienced, making it more interconnected, interactive, and adaptable to new discoveries.

**Keywords:** Knowledge Graphs, Digital Cultural Heritage, Scientific Heritage.

## 1 Introduction

The preservation and study of cultural heritage in the digital era require innovative approaches to managing and analyzing vast amounts of historical data. Knowledge graphs (KGs) have emerged as a powerful tool for structuring and connecting diverse pieces of cultural information, enabling scholars and researchers to explore relationships between historical texts, artifacts, and traditions more effectively (Chaudhri et al., 2022; Akter & Rahoman, 2023). Unlike conventional databases, KGs capture complex interconnections within cultural heritage, allowing for a more nuanced understanding of historical narratives and their evolution over time (Huang et al., 2023; Wang et al., 2024; Tan, 2024; Maree, 2025).

As the scale and complexity of historical data continue to grow, artificial intelligence (AI) and machine learning (ML) have become essential in automating the extraction, organization, and interpretation of cultural information. Recent advancements in deep learning (DL) have further enhanced the capabilities of KGs by improving entity recognition, relationship detection, and semantic understanding (d'Amato et al., 2023).

One of the critical advantages of KGs in cultural heritage studies is their ability to bridge linguistic, geographical, and temporal gaps in historical research (Fan et al., 2023; Hosamo & Mazzetto, 2025). This is particularly relevant in the study of ancient texts and traditions, where variations in language, script, and historical context can complicate interpretation. Advanced natural language processing (NLP) models, including transformer-based architectures, contribute to this process by enhancing the extraction of meaning from historical documents, ensuring that cultural heritage remains accessible and interpretable across different linguistic and disciplinary boundaries. (Bruns, 2023).

This paper aims to analyze the latest innovations in KGs for cultural heritage and their potential to enhance data accessibility, usability, and engagement. The study is structured as follows: In chapter 2, it explores the role of AI and ML as well as DL techniques in automating the enrichment and dynamic updating of KGs. Chapter 3 reviews notable KGs worldwide and in Bulgaria. Chapter 4 discusses the intellectual property (IP) issue. Chapter 5 investigates the integration of KGs with immersive technologies such as augmented reality (AR) and virtual reality (VR), which enable more interactive engagement with heritage content.

## **2 Knowledge Graphs and Deep Learning Techniques for Enhanced Digital Cultural Heritage Management**

The integration of KGs and DL techniques offers significant improvements in managing digital cultural heritage by systematically structuring and analyzing fragmented historical data. AI-powered analysis further strengthens this process by refining the annotation and classification of cultural data, enabling the identification of connections between different historical elements (Bobasheva et al, 2022). KGs contribute to the verification and contextualization of historical information by assigning reliability scores to different sources, thereby minimizing ambiguities in historical research (Huang et al., 2023). The application of graph-based frameworks like GraphBRAIN (2024) enhances knowledge representation, offering greater efficiency, flexibility, and expressiveness in managing cultural heritage data (Ferilli et al., 2025).

Constructing KGs for cultural heritage presents unique challenges, particularly in improving knowledge retrieval and visualization. Traditional approaches often struggle with efficiently identifying entities and their relationships within historical and artistic contexts. To address these issues, advanced techniques integrate NLP and graph-based ML models, such as Graph Attention Networks (GAT) (Wang et al., 2024). A key aspect of this process involves collecting and encoding textual data related to historical artifacts. Pretrained language models, such as BERT, are often used to capture semantic nuances, ensuring that information is accurately processed (Devlin et al, 2019).

A bottom-up approach, beginning with targeted knowledge collection and ontology design, ensures that the resulting graph effectively represents historical artifacts and their contextual significance. Improving the usability of cultural heritage KGs is essential for making historical data more accessible to diverse audiences. While KGs effectively structure and link cultural information, their complexity often limits interaction to experts familiar with graph databases and ontology design. Efforts to address this challenge focus on refining user interfaces, enhancing visualization techniques, and implementing intuitive interaction mechanisms (Mulholland et al, 2024).

Combining KGs with digital twins, a more comprehensive framework can be developed to improve data accessibility, enable predictive maintenance, and support long-term preservation strategies (Hosamo & Mazzetto, 2025). KGs provide a structured representation of historical and architectural data, while digital twins create real-time, dynamic models that reflect the current state of heritage structures. A key advantage of this approach is its ability to facilitate proactive maintenance by identifying potential structural issues before they become critical.

### **3 Knowledge Graph Role in Cultural Heritage Preserving**

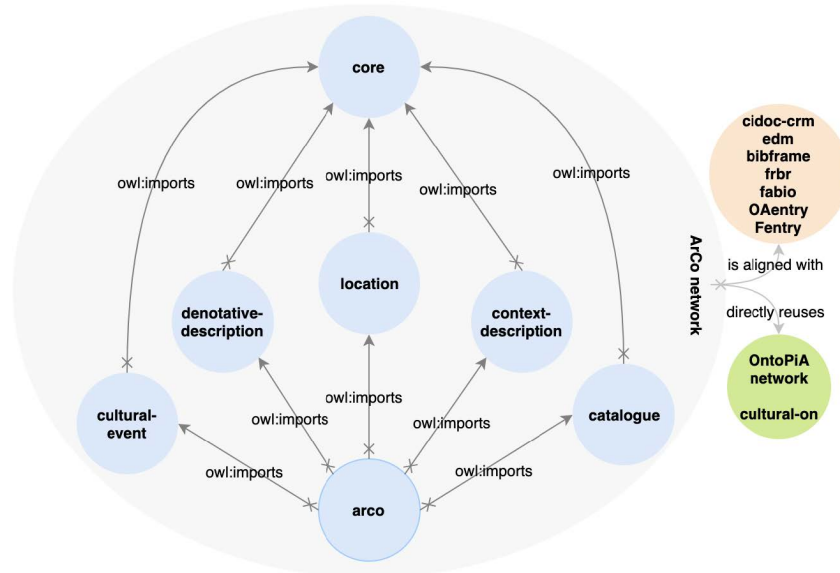
KGs have found widespread applications across various domains, from AI-driven systems to scientific research and cultural heritage preservation. Their ability to represent structured relationships makes them a powerful tool for data integration, inference, and knowledge discovery. This helps researchers, historians, and the public explore complex relationships between artifacts, events, and people across time and geography.

A vivid example is ClaOnto. This is an ontology designed to capture the complexities of ancient Chinese classics, including their various versions, interrelations, content, associated individuals, locations, periods, and languages. This comprehensive framework emphasizes the processes and evaluations involved in the transmission of these texts, contributing to a nuanced understanding and heightened awareness of their preservation. It comprises three modules: Basic Information, Transmission Information, and Evaluation Information. The practical application of ClaOnto follows a structured workflow for implementing an ontology-based KG, utilizing methods such as data mining, text analysis, and data visualization (see Fig. 1). By integrating ClaOnto, the evolution and dissemination of ancient Chinese classics can be better understood, supporting their preservation and continued relevance (Cui et al., 2024).

ArCo is a comprehensive KG designed to model, structure, and enrich data related to Italian cultural heritage. Built on linked data principles, it provides an interoperable framework that integrates information from national cultural institutions, enabling advanced querying, reasoning, and semantic enrichment of heritage resources. The development of the ArCo KG is centered around a structured and modular ontology framework designed to model Italian cultural heritage data comprehensively. The KG integrates data from ICCD (Istituto Centrale per il Catalogo e la Documentazione), ensuring that diverse cultural artifacts, historical sites, and artistic works are represented in a semantically rich manner. ArCo's architecture is built on top of LOD principles, allowing seamless integration with other cultural heritage datasets.



and knowledge management systems that structure and organize cultural and linguistic resources. By employing ontologies and knowledge graphs, these systems enable efficient representation of relationships between entities, concepts, and historical contexts.



**Fig. 2.** ArCo ontology network. (Carriero et al, 2019).

Such structured approach enhances the accessibility and usability of language data for researchers, educators, and the general public, fostering deeper engagement with cultural heritage materials. Linguistic technologies play a fundamental role in these infrastructures, supporting tasks such as text analysis, machine translation, and named entity recognition (Simov & Osenova, 2019).

A digital project called the Bulgarian Iconographical Digital Library (BIDL) aims to preserve and promote the rich history of Bulgarian Orthodox iconography. In-depth information about icons, such as metadata on iconographers, stylistic schools, techniques, materials, and locations, is gathered, digitized, and arranged by the project. The project improves accessibility, interoperability, and research potential by organizing this data in accordance with semantic web standards and connecting it to international repositories (Alexiev et al, 2020) (see Fig. 3). By acting as a digital preservation tool and scholarly resource, BIDL makes sure that Bulgaria's artistic and religious legacy is included in the global network of cultural knowledge.

Semantic web technologies can be used to efficiently structure and link cultural heritage data about icons. Details like authorship, iconographic school, techniques, materials, and geographic origin are captured by an appropriate model. After that, this data can be converted into formats that are compatible with other systems and meet international standards (Alexiev, 2018). The information is made more searchable, accessible, and reusable across systems by incorporating it into websites such as Europeana and

Wikidata. In order to improve metadata richness and connect with external datasets such as VIAF and Getty AAT, the method involves exporting data from the BIDL to Wikidata. The data also can be transformed into the Europeana Data Model (Alexiev et al, 2020), which makes digital cultural heritage platforms more accessible and interoperable. BIDL aligns closely with the principles of CIDOC Conceptual Reference Model (CIDOC-CRM), a standard ontology for cultural heritage information. BIDL structures iconographic data, such as iconographers, iconographic schools, materials, and locations into entities and relationships that reflect cultural and historical contexts, which is a core aim of CIDOC-CRM (Carboni, 2019).

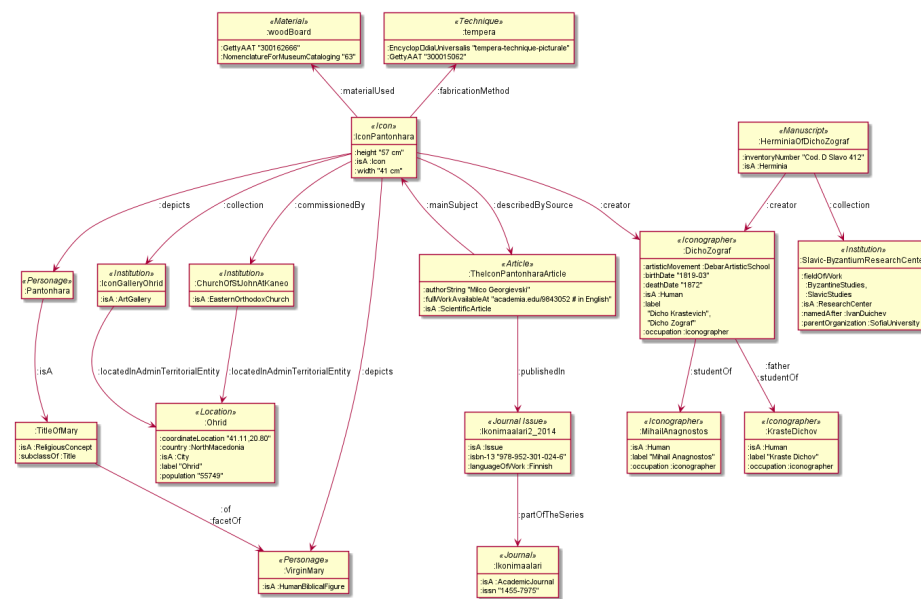


Fig. 3. Icon Knowledge Graph Model (Alexiev et al, 2020).

The Archaeological Map of Bulgaria is a national digital platform created to manage and record archaeological sites across the nation. By incorporating archaeological data into the initial phases of development projects, it facilitates infrastructure planning by guaranteeing that cultural heritage is recognized, evaluated, and conserved as needed.

The system facilitates effective workflows that include excavation, GIS-based field surveys, and preliminary assessments. It encourages cooperation between public authorities, developers, and archaeologists, balancing the needs of modern construction with the preservation of cultural heritage (Kecheva, 2018). This methodical approach is a prime example of how legal frameworks and digital tools can balance national development objectives with heritage preservation.

The Archaeological Map of Bulgaria is being modified to adhere to the FAIR principles (Findable, Accessible, Interoperable, Reusable) and incorporated into European research infrastructures such as ARIADNE and ARIADNEplus. The procedure entails harmonizing metadata, semantically enriching the data, and connecting the database to

well-known ontologies like CIDOC CRM (Kecheva, 2024). Although the KG technology is not yet fully utilized in this example, this project improves Bulgarian archaeological data's compatibility with other European heritage systems, increasing its accessibility for study and fostering a more cohesive digital cultural heritage network.

## **4 Challenges and Limitations**

Digitizing cultural heritage presents both opportunities and challenges, particularly in the context of IP rights. As cultural institutions and researchers work to preserve and make heritage materials accessible in digital formats, they must navigate complex legal frameworks governing copyrights, licensing, and ownership. Ensuring compliance with IP laws is essential for balancing the rights of creators, institutions, and the public while promoting the ethical use of digitized content. One of the key challenges in this process is determining the legal status of historical artifacts, documents, and traditional knowledge. Many cultural heritage materials exist in legal gray areas, where rights holders may be difficult to identify, or copyright protections may vary depending on jurisdiction. Institutions undertaking digitization efforts must develop clear policies on rights clearance, licensing agreements, and the responsible use of digital reproductions to avoid legal disputes. (Borissova, 2010). Cultural institutions must navigate the fine line between safeguarding the IP rights of original creators and fulfilling their mission to provide public access to cultural heritage. Implementing restrictive access controls to protect IP can limit educational and scholarly use, while open access policies may expose institutions to potential IP infringements (Todorov & Lutfiu., 2023).

## **5 Future Trends**

The future of cultural heritage KGs lies in advancing both their construction methodologies and practical applications. As digital preservation becomes a priority, integrating AI and ML techniques will be crucial for automating knowledge extraction, relationship detection, and contextual analysis. More sophisticated NLP models will enhance entity recognition from historical texts, ensuring that KGs capture nuanced relationships between artifacts, historical figures, and events with greater accuracy. One significant trend is the growing use of DL models, such as transformer-based architectures, to refine semantic understanding in cultural heritage data.

Knowledge fusion techniques will improve, allowing disparate datasets from museums, archives, and research institutions to be integrated into unified, interoperable KGs. The development of more dynamic and self-updating KGs will also shape the future. Traditional KGs often require manual updates, but AI-driven models will automate knowledge enrichment by continuously integrating new research findings, archaeological discoveries, and evolving interpretations of historical events. Another important trend is the enhancement of visualization and interactive applications (Stanković et al., 2024). As KGs become more comprehensive, they will be integrated into immersive technologies such as AR and VR. By combining KGs with 3D modeling and geospatial data, researchers and the public will gain new ways to interact with cultural heritage

(Tatić et al, 2024). As cultural heritage KGs expand, their integration with the Semantic Web and linked open data initiatives will become increasingly important. In the long term, the combination of AI, KGs, and digital preservation efforts will redefine how cultural heritage is studied and experienced (Georgiev & Nikolova, 2021).

## 6 Conclusions

The evolution of cultural heritage KGs is poised to revolutionize how historical and artistic information is preserved, analyzed, and shared. By integrating advanced AI and machine learning techniques, cultural heritage applications will become more automated, dynamic, and capable of handling complex semantic relationships. Improved NLP models will enhance entity recognition, while DL approaches will refine the contextualization of artifacts, ensuring that KGs remain accurate and continuously updated. Future trends indicate a shift toward more interactive and immersive applications, with AR and VR technologies enabling users to explore historical sites and artifacts in novel ways. Additionally, interdisciplinary collaborations will be essential for maintaining historical accuracy while leveraging cutting-edge computational methods. Open data initiatives and standardized ontologies will further enhance knowledge sharing across institutions, promoting accessibility and global engagement with cultural heritage. By providing structured, machine-readable representations of historical knowledge, KGs can support decision-making in heritage conservation, enabling policymakers to assess risks, prioritize restoration projects, and allocate resources more effectively.

## Acknowledgements.

This research is funded partly by CLaDA-BG, the Bulgarian National Interdisciplinary Research e-Infrastructure for Resources and Technologies in favour of the Bulgarian Language and Cultural Heritage, part of the EU infrastructures CLARIN and DARIAH, <https://clada-bg.eu/bg/>, Grant number DO1-324/01.12.2023 and Grant number DO1-97/26.06.2025, financed by the Bulgarian Ministry of Education and Science, Funding Procedure: The National Science Infrastructure Roadmap 2020-2027 (NSIF).

## References

- Akter, M.M., & Rahoman, M. (2023). A Systematic Review on Knowledge Graphs Classification and Their Various Usages. *Artificial Intelligence Evolution*, 4(2), 99-233. <https://doi.org/10.37256/aie.4220233605>
- Alexiev, V. (2018). Museum Linked Open Data: Ontologies, Datasets, Projects. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 8, 19-50. <https://doi.org/10.55630/dipp.2018.8.1>
- Alexiev, V., Tarkalanov, P., Georgiev, N., & Pavlova, L. (2020). Bulgarian Icons in Wikidata and EDM. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 10, 45–64. <https://doi.org/10.55630/dipp.2020.10.2>

- Bobasheva, A., Gandon, F., & Precioso, F. (2022). Learning and Reasoning for Cultural Metadata Quality: Coupling Symbolic AI and Machine Learning over a Semantic Web Knowledge Graph to Support Museum Curators in Improving the Quality of Cultural Metadata and Information Retrieval. *Journal on Computing and Cultural Heritage*, 15(3), Article 40, 1-23. <https://doi.org/10.1145/3485844>
- Borissova, V. (2010). *Digitizing cultural heritage in Bulgaria: A Survey of Intellectual Property-related Experiences and Practices*. World Intellectual Property Organization (WIPO). [https://www.wipo.int/export/sites/www/tk/en/docs/resources/borissova\\_report.pdf](https://www.wipo.int/export/sites/www/tk/en/docs/resources/borissova_report.pdf)
- Bruns, O. (2023). The persistence of temporality: Representation of time in cultural heritage knowledge graphs. In C. d'Amato & J. Z. Pan (Eds.) *Proceedings of the Doctoral Consortium at ISWC 2023 co-located with 22nd International Semantic Web Conference (ISWC 2023)* (CEUR Workshop Proceeding, Vol. 3678). [https://iswc2023.semanticweb.org/wp-content/uploads/2023/11/ISWC2023\\_paper\\_378.pdf](https://iswc2023.semanticweb.org/wp-content/uploads/2023/11/ISWC2023_paper_378.pdf)
- Carboni, L. (2019). An Ontological Approach to the Description of Visual and Iconographical Representations. *Heritage*, 2(2), 1191-1210. <https://doi.org/10.3390/heritage2020078>
- Carriero, V. A., Gangemi, A., Mancinelli, M. L., Marinucci, L., Nuzzolese, A. G., Prestutti, V., & Veninata, C. (2019). ArCo: The Italian Cultural Heritage Knowledge Graph. In Ghidini, C., et al. (Eds.), *The Semantic Web – ISWC 2019. ISWC 2019. Lecture Notes in Computer Science*, vol. 11779 (pp. 36–52). Springer. [https://doi.org/10.1007/978-3-030-30796-7\\_3](https://doi.org/10.1007/978-3-030-30796-7_3)
- Chaudhri, V., Baru, C., Chittar, N., Dong, X., Genesereth, M., Hendler, J., Kalyanpur, A., Lenat, D., Sequeda, J., Vrandečić, D., & Wang, K. (2022). Knowledge graphs: Introduction, history and perspectives. *AI Magazine*, 43(1), 17–29. <https://doi.org/10.1002/aaai.12033>
- Cui, Y., Yao S., Wu, J., & Lv, M. (2024). Linking past insights with contemporary understanding: an ontological and knowledge graph approach to the transmission of ancient Chinese classics. *Heritage Science*, 12, 150. <https://doi.org/10.1186/s40494-024-01504-x>
- d'Amato, C., Mahon, L., Monnin, P., & Stamou, G. (2023). Machine Learning and Knowledge Graphs: Existing Gaps and Future Research Challenges. *Transactions on Graph Data and Knowledge*, 1(1), 8:1-8:35. <https://doi.org/10.4230/TGDK.1.1.8>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). Association for Computational Linguistics. <https://aclanthology.org/N19-1423.pdf>
- Fan, T., Wang, H. & Hodel, T. (2023). CICHMKG: a large-scale and comprehensive Chinese intangible cultural heritage multimodal knowledge graph. *Heritage Science*, 11, 115. <https://doi.org/10.1186/s40494-023-00927-2>

- Ferilli, S., Bernasconi, E., Di Pierro, D., & Redavid, D. (2025). The GraphBRAIN Framework for Knowledge Graph Management and Its Applications to Cultural Heritage. In: Steffen, B. (Ed.) *Bridging the Gap Between AI and Reality. AISoLA 2023. Lecture Notes in Computer Science*, vol. 14129 (pp. 144–161). Springer. [https://doi.org/10.1007/978-3-031-73741-1\\_10](https://doi.org/10.1007/978-3-031-73741-1_10)
- Georgiev, V., & Nikolova, A. (2021). Virtual Reality Simulations for Presenting Cultural-historic Content in e-Learning for Kids. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 11, 267–272. <https://doi.org/10.55630/dipp.2021.11.24>
- Graphbrain. (2024). *Graphbrain – Language, Knowledge, Cognition*. <https://graphbrain.net/>
- Hosamo, H., & Mazzetto, S. (2025). Integrating Knowledge Graphs and Digital Twins for Heritage Building Conservation. *Buildings*, 15(1), 16. <https://doi.org/10.3390/buildings15010016>
- Huang, Y.Y., Yu, S.S., Chu, J.J., Fan, H.H., & Du, B.B. (2023). Using knowledge graphs and deep learning algorithms to enhance digital cultural heritage management. *Heritage Science*, 11, 204. <https://doi.org/10.1186/s40494-023-01042-y>
- Kecheva, N. (2018) Archaeological Map of Bulgaria – Transport and Pipeline Infrastructure Projects, *Internet Archaeology*, 51. <https://doi.org/10.11141/ia.51.2>
- Kecheva, N. (2024) The First Step towards FAIR-ness in Bulgarian Archaeology: the Archaeological Map of Bulgaria in ARIADNE and ARIADNEplus, *Internet Archaeology* 67. <https://doi.org/10.11141/ia.67.5>
- Li, W., Qi, G., & Qiu, J. (2020). Hybrid reasoning in knowledge graphs: Combining symbolic reasoning and statistical reasoning. *Semantic Web Journal*, 11(1), 53–62. <https://doi.org/10.3233/SW-190375>
- Maree, M. (2025). *Quantifying Relational Exploration in Cultural Heritage Knowledge Graphs with LLMs: A Neuro-Symbolic Approach*. arXiv:2501.06628. <https://doi.org/10.48550/arXiv.2501.06628>
- Mulholland, P., Kranenburg, P. V., Carvalho, J., & Daga, E. (2024). Supporting the end-user curation of cultural heritage knowledge graphs. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (pp. 35–44). Association for Computing Machinery. <https://doi.org/10.1145/3648188.3675132>
- Simov, K., & Osenova, P. (2019). Integrated language and knowledge resources for a Bulgarian-centric knowledge graph. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 9, 49–58.
- Stanković, R.S., Tatić, D., & Luchev, D. (2024). Towards an Integral Solution for Application of Information Technologies in Representation of Medieval Fortresses. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 14, 53–62. <https://doi.org/10.55630/dipp.2024.14.4>
- Tan, M.A. (2024). Knowledge Graph Construction and Refinement for Cultural Heritage Digital Libraries. In K. Taylor & A. Zimmermann (Eds.), *Proceedings of the Doctoral Consortium at ISWC 2024 co-located with the 23rd International Semantic Web Conference (ISWC 2024)* (CEUR WP, Vol-3884, pp. 1-8). CEUR Workshop Proceedings. <https://ceur-ws.org/Vol-3884/paper7.pdf>

- Tatić, D., & Stanković, R.S. (2024). Augmented reality software in presentation of cultural and historical heritage. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 14, 71–82. <https://doi.org/10.55630/dipp.2024.14.6>
- Todorov, T., & Lutfiu, S. (2023). Intellectual property protection of digital cultural heritage. *Digital Presentation and Preservation of Cultural and Scientific Heritage*, 13, 263–268. <https://doi.org/10.55630/dipp.2023.13.25>
- Wang, Y., Liu, J., Wang, W., Chen, J., Yang, X., Sang, L., Wen, Z., & Peng, Q. (2024). Construction of Cultural Heritage Knowledge Graph Based on Graph Attention Neural Network. *Applied Sciences*, 14(18), 8231. <https://doi.org/10.3390/app14188231>

Received: April 10, 2025

Reviewed: May 02, 2025

Finally Accepted: May 15, 2025

