

**FROM ARCHIVAL GOVERNANCE TO DIGITAL HERITAGE: SEMANTIC
MODELLING AND DIGITISATION OF BULGARIAN ARCHIVAL INSTITUTIONS
(1878-1912)**

Arsini Kolev

University of Library Studies and Information Technologies, Bulgaria, a.kolev@unibit.bg

ORCID ID: 0009-0002-6748-070X

Svetoslava Dimitrova

University of Library Studies and Information Technologies, Bulgaria, s.dimitrova@unibit.bg

ORCID ID: 0000-0001-5540-1539

Abstract: The paper examines the institutional development of Bulgarian archival governance during the formative period of the modern Bulgarian state, from 1878 to 1912, and proposes a conceptual framework for its semantic modelling and digitisation. The purpose of the study is to reconstruct the legal, administrative and documentary logic through which record-keeping and archival practices became part of the governance model after the Liberation. The methodology combines historical-institutional analysis with conceptual modelling. Normative acts, administrative regulations, procedural rules and institutional practices are interpreted as evidence of relations between administrative bodies, documentary processes, archival responsibilities and mechanisms of institutional continuity. The results identify a stable system of entities and relations: state and local institutions, archival units, regulatory instruments, document types and procedural actions such as creation, registration, classification, preservation and transfer. These components are transformed into a semantic framework that supports contextual description, structured metadata and interoperability with e-archives, digital repositories and knowledge infrastructures. The study concludes that early Bulgarian archival governance can be represented not only as a historical phenomenon, but also as a reusable digital heritage resource. The proposed model shifts the focus from document-level digitisation to system-level representation, preserving the institutional context in which documents were created and used. The paper recommends further technical implementation, validation with historical datasets, alignment with contemporary archival ontologies and the development of visualisation tools for administrative hierarchies and document flows. In this way, archival history contributes to access, research, education and the preservation of cultural memory in the digital age.

Keywords: archival governance, digitisation of cultural heritage, semantic modelling, e-archives, Bulgarian administrative history

1. INTRODUCTION

Archives have evolved from custodial repositories into infrastructures for access, interpretation and knowledge generation. In the context of extensive digitisation, archival systems are both mediators of historical knowledge and historical phenomena that must be critically examined, documented and digitally represented. Contemporary research therefore stresses the need for metadata frameworks, interoperability and long-term access to archival information (Conway, 2010; Gilliland, 2014; Hawkins, 2022; O'Neill & Stapleton, 2022).

The first decades after the Liberation of Bulgaria, from 1878 to 1912, were decisive for the creation of legislative frameworks, administrative institutions and record-keeping procedures. Newly formed state authorities established rules for the production, registration, classification, preservation and transfer of official documentation. In this setting the archive was not merely a record repository. It operated as part of the governance paradigm, supporting institutional legitimacy, legal accountability and administrative continuity (Cook, 1997).

Although this early archival system has significant historical value, it remains insufficiently represented in contemporary digital heritage environments. Digitisation often focuses on individual documents or collections, while the institutional logic that generated and managed those documents is not always modelled. This limits the possibility for structured search, reuse, comparison and knowledge extraction. The aim of this paper is to establish a conceptual framework for the semantic modelling and digitisation of early Bulgarian archival institutions. The contribution is threefold: it systematises archival governance in the formative period of the Bulgarian state; it proposes a model for representing institutional characteristics as interoperable digital structures; and it demonstrates how administrative memory may be transformed into a searchable digital heritage resource.

2. MATERIALS AND METHODS

The study applies a historical-institutional approach combined with conceptual semantic modelling. The historical scope covers the period 1878-1912, when the administrative foundations of the modern Bulgarian state were established. The source base includes normative regulations concerning governmental administration and documentation, procedural rules for record creation and preservation, regulations and directives related to archival activity, and institutional correspondence that reveals practical record-keeping procedures.

The analysis reconstructs the structural elements of archival governance: institutions such as ministries, departments, municipal authorities and archival organisations; administrative tasks including document creation, registration, classification, storage, transfer and preservation; document categories and record series; and the hierarchical circulation of documents within and among institutions. Particular attention is given to procedural logic, because document flows reflect the broader structure of state administration.

At the historical level, the materials are interpreted as evidence of how the state produced, authorised, organised and preserved its administrative memory. The investigation therefore moves from individual documents to the institutional conditions of their creation. This makes it possible to identify stable patterns in the relations between authority, office work, archival responsibility and legal continuity.

The semantic modelling stage converts the reconstructed historical elements into entities, relations and processes. The model includes administrative bodies, roles, archival units, regulatory acts, documentary outputs and document lifecycles. It establishes links between actors, functions and documents and identifies core metadata elements for digital representation. This approach is grounded in the broader development of cultural heritage linked data, CIDOC CRM-based modelling and contemporary archival ontologies (Davis & Heravi, 2021; Doerr, 2003; Giagnolini et al., 2025; International Council on Archives Expert Group on Archival Description, 2023; Koch et al., 2023).

The proposed digitisation framework is not limited to the creation of digital copies. It aims to represent the governance structure that produced the documents. For this reason, the model includes descriptive metadata, hierarchical structuring and semantic links between administrative functions, regulatory acts, documentary series and archival procedures. The expected outcome is enhanced search, contextual analysis and reuse of historical administrative data.

The applicability of the model was assessed through structural mapping of selected institutional examples from the period under investigation. The mapping aimed to test whether administrative hierarchies, regulatory frameworks and document circulation patterns could be represented within a coherent digital framework without losing historical context. The validation is conceptual rather than technical, but it establishes the basis for future implementation in e-archives, repositories and knowledge graphs.

3. RESULTS

The research resulted in the reconstruction of the institutional framework of Bulgarian archival governance and in the development of a conceptual structure for its semantic representation. The findings show that archival governance in the examined period was an integral part of state administration rather than an auxiliary activity. Creation, registration, classification, preservation and transfer of records were regulated through procedural norms and distributed institutional responsibilities.

The reconstructed model reveals an administratively structured hierarchy in which record-keeping responsibilities were shared among ministries, central agencies and local administrations. Rules for document registration and categorisation reflected the growing bureaucratic rationale of the state, while responsibilities for preservation and transfer demonstrated the connection between administrative authority and documentary control. This confirms that archival governance functioned as a regulatory mechanism for administrative continuity and legal legitimacy.

A second result is the identification of core entities required for semantic representation: administrative bodies, archival and record management units, regulatory instruments, document types, series of records and procedural actions. The relations between these components are essential. Regulatory acts guide institutions; institutions produce and manage documents; documents pass through procedural lifecycles; and archival bodies ensure long-term preservation. These relations make it possible to describe governance processes rather than isolated documents.

A third result is the proposed conceptual semantic framework. It supports the representation of institutional hierarchies, document lifecycles, regulatory frameworks and the relations between stakeholders, functions and documentary outputs. Such a framework can improve contextualisation and enable more advanced queries across institutional, functional and documentary dimensions. In line with developments in linked data and digital cultural heritage, the model enables the transformation of historical administrative information into a structured and reusable knowledge resource (Hawkins, 2022; Terras, 2015).

The findings also indicate that digitisation based only on the object level would be insufficient for the study of archival governance. A document can be searched and displayed as a digital item, but its full meaning depends on the institution, procedure and regulatory context in which it was created. The proposed model therefore preserves relationships between administrative function, documentary evidence and archival preservation. This strengthens the scientific and educational value of digital collections.

Overall, the results demonstrate that early Bulgarian archival governance can be transformed into a network-based representation suitable for interoperable digital environments. Historical administrative data may be used not only for description, but also for comparison, visualisation and knowledge extraction. The model thus links archival history with current priorities in digital heritage and digital humanities.

4. DISCUSSIONS

The findings confirm that the early Bulgarian archival system can be methodically reconstructed and represented as a semantic structure. This has implications both for archival history and for digital heritage studies. If archival objects are digitised without modelling the governance structures that produced them, the contextual integrity of the records may be weakened. A system-level approach preserves the connections among institutions, documents, rules and procedures.

The main methodological contribution is the integration of historical-institutional research with digital modelling. Digitisation is understood not only as the production of digital surrogates, but also as an interpretive activity that requires knowledge of the administrative framework in which records were created. This corresponds to the contemporary emphasis on interoperability, knowledge graphs, provenance and open access in cultural heritage infrastructures (Davis & Heravi, 2021; Giagnolini et al., 2025; O'Neill & Stapleton, 2022).

The approach also responds to a common problem in digital archives: the fragmentation of context. When metadata are limited to title, date, creator and physical description, they do not fully express the administrative processes behind the documents. By modelling actors, functions, regulatory sources and document lifecycles, the proposed framework makes the administrative meaning of archival materials more explicit and more reusable.

Several limitations should be acknowledged. The proposed framework is conceptual and requires technical implementation and empirical testing in an operational platform. The analysis is focused on a specific national and historical context; therefore, its transfer to other administrative traditions would require adaptation. The availability, completeness and physical condition of historical sources may also influence the level of detail in digital representation. Nevertheless, the model establishes a basis for future metadata crosswalks, comparative institutional analysis and visualisation of administrative hierarchies and document flows (Koch et al., 2023).

5. CONCLUSIONS

The paper analysed the institutional evolution of Bulgarian archival governance during the period 1878-1912 and proposed a conceptual framework for its semantic modelling and digitisation. The study showed that archives in the formative period of the modern Bulgarian state functioned as key components of governance, ensuring continuity, accountability and preservation of administrative memory.

The research demonstrated that the institutional logic of archival systems can be identified and converted into structured digital entities and relationships. This enables historical documentation to be represented not only as separate digital objects, but also as part of a connected governance system. The transition from document-level digitisation to system-level modelling improves contextual interpretation, semantic interoperability and knowledge extraction.

Future work should focus on implementing the model in an experimental e-archive environment, aligning it with contemporary archival and cultural heritage standards, and testing it with a larger body of historical sources. The development of visual and interactive tools would further support research, education and public access. In this sense, modelling archival governance contributes to the transformation of archives into advanced knowledge systems for the digital age.

ACKNOWLEDGEMENTS

This paper is a result of a scientific study conducted under the project “Research on the Archive Institution in the Governance Model of Bulgaria in the Period 1878-1912”, with contract No. KP-06-M85/2 of 05.12.2024, funded by the National Science Fund of the Ministry of Education and Science, based on a Competition for Fundamental Scientific Research for Young Scientists and Postdoctoral Students - 2024, led by Chief Assistant Arsini Kolev, PhD.

REFERENCES

- Conway, P. (2010). Preservation in the age of Google: Digitization, digital preservation, and dilemmas. *The Library Quarterly*, 80(1), 61-79. <https://doi.org/10.1086/648463>
- Cook, T. (1997). What is past is prologue: A history of archival ideas since 1898, and the future paradigm shift. *Archivaria*, 43, 17-63. <https://archivaria.ca/index.php/archivaria/article/view/12175>
- Davis, E. L., & Heravi, B. (2021). Linked data and cultural heritage: A systematic review of participation, collaboration, and motivation. *Journal on Computing and Cultural Heritage*, 14(2), Article 21. <https://doi.org/10.1145/3429458>
- Doerr, M. (2003). The CIDOC conceptual reference model: An ontological approach to semantic interoperability of metadata. *AI Magazine*, 24(3), 75-92. <https://doi.org/10.1609/aimag.v24i3.1720>
- Giagnolini, L., Koch, I., Tomasi, F., & Teixeira Lopes, C. (2025). Comparative insights into semantic archival modelling: Evaluating RiC-O and ArchOnto representation capabilities. *Journal of Documentation*, 81(4), 1003-1031. <https://doi.org/10.1108/JD-12-2024-0310>
- Gilliland, A. J. (2014). *Conceptualizing 21st-century archives*. Society of American Archivists.
- Hawkins, A. (2022). Archives, linked data and the digital humanities: Increasing access to digitised and born-digital archives via the semantic web. *Archival Science*, 22, 319-344. <https://doi.org/10.1007/s10502-021-09381-0>
- International Council on Archives Expert Group on Archival Description. (2023). *Records in contexts: Foundations of archival description* (Version 1.0). International Council on Archives.
- Koch, I., Teixeira Lopes, C., & Ribeiro, C. (2023). Moving from ISAD(G) to a CIDOC CRM-based linked data model in the Portuguese Archives. *Journal on Computing and Cultural Heritage*, 16(4), Article 71, 1-21. <https://doi.org/10.1145/3605910>
- O'Neill, B., & Stapleton, L. (2022). Digital cultural heritage standards: From silo to semantic web. *AI & Society*, 37, 891-903. <https://doi.org/10.1007/s00146-021-01371-1>
- Terras, M. (2015). Opening access to collections: The making and using of open digitised cultural content. *Online Information Review*, 39(5), 733-752. <https://doi.org/10.1108/OIR-06-2015-0193>