Design Considerations for a Generic Graph Database in Archival Document Management

Ilija Hristoski[1], Jelena Stojanov[2] and Željko Stojanov[2]

[1] *"St. Kliment Ohridski" University – Bitola / Faculty of Economics – Prilep, Prilepski Braniteli St. 143, MK-7500 Prilep, N. Macedonia*
[2] *University of Novi Sad / Technical Faculty "Mihajlo Pupin", Đure Đakovića St., 23000 Zrenjanin, Serbia*

ilija.hristoski@uklo.edu.mk; jelena.stojanov@uns.ac.rs; zeljko.stojanov@uns.ac.rs

Abstract:
In the digital age, managing archival documents in both physical and digital formats poses a significant challenge due to the sheer volume and complexity of historical records. Traditional relational databases often struggle to capture the intricate relationships and interconnectedness present in archival data. This research paper explores the utilization of property graphs, for creating a robust and efficient graph database system aimed at storing, organizing, and retrieving archival documents. By representing documents and their associated metadata as nodes and relationships in a graph, we propose a generic graph database design that can serve as a comprehensive framework for modeling and querying interconnected historical data, supporting efficient and adaptable archival document management. The paper contributes insights into designing graph databases that harmonize with the intricate nature of archived documents while catering to the evolving needs of modern archival practices.

Keywords:
Graph databases, archival document management, database design, graph database schema, graph data model, data modeling

1. Introduction

Archives consist of document collections or 'records' chosen for permanent retention due to their significance as proof or as sources for historical and other forms of research. Records emerge from the actions of individuals and organizations; they hold a functional role while in active circulation, and a portion of them are subsequently chosen and safeguarded as components of an archival compilation [1]. Archive collections usually possess unique attributes, underscoring the crucial need for their proper upkeep. Despite the fact they consist of historical records, archives provide an immense source of information. As Taryn Simon* once remarked, "*Archives exist because there's something that can't necessarily be articulated. Something is said in the gaps between all the information*." They demand meticulous storage and administration to ensure their safeguarding and persistence for present and future utilization. Besides specialist collecting institutions, which are also called 'archives', including national and local archives and record offices, archives are also kept by other institutions, e.g. museums, libraries, universities, faculties, schools, businesses, charities, arts organizations, and community groups. All of these subjects usually hold and maintain their own institutional records. Archives are highly valuable sources of information for research to nations and regions, organizations, communities, and individuals since they provide proof of historical occurrences. They narrate tales, document individuals and identities, and organizations' activities, constituting our recorded memory and being a vital component of our community, cultural, official, and unofficial history. This is completely in line with Hasan M. Elahi's† famous quote stating that "*We're all creating an archive of our own lives, whether we're aware of it or not*."

* Taryn Simon (b. 1975) is an American multidisciplinary artist who works in photography, text, sculpture, and performance.

† Hasan M. Elahi (b. 1972) is a Bangladeshi-born American interdisciplinary media artist whose work has an emphasis on technology and media and their social implications.

In general, the utilization of databases for archival purposes involves the organized storage, management, and retrieval of historical records and documents in a digital format. Databases are crucial tools for archivists and institutions looking to preserve, organize, and provide access to their archival collections.

More specifically, the utilization of graph databases for archival purposes offers a compelling alternative to other traditional types of databases, such as relational, object-oriented, and object-relational databases. Graph databases excel in managing and preserving archival collections due to their unique data modeling capabilities that are particularly well-suited to the complex, interconnected nature of historical records and documents, and their ability to represent complex relationships, hierarchical structures, rich metadata, and historical versions efficiently. Besides, their query optimization for graph traversal, scalability, and adaptability to evolving archival needs make them a valuable choice for preserving, managing, and exploring archival collections in the digital age. According to Gartner, it is anticipated that by 2025, graph database technologies will be integrated into 80% of data and analytics innovations, a substantial increase from the 10% adoption rate in 2021. This is expected to expedite decision-making processes throughout enterprises [2].

This research contributes to the overall body of knowledge related to building electronic archives, with a special emphasis on archiving office documents. The paper's main goal is to propose a generic framework suitable for designing a graph database that will facilitate archival document management. The resulting graph data model can be easily modified, enhanced, and adapted to meet any organization's specific requirements for archiving documents.

The rest of the paper is structured as follows. Section 2 offers a concise review of the most notable research conducted in recent years on this subject. In Section 3, we delve into the process of archiving documents, providing a brief but comprehensive overview. Moving forward to Section 4, we explore the world of graph databases, presenting a succinct overview of their essential characteristics and functionalities. This section lays the foundation for the subsequent discussion on our proposed graph data model. Section 5 presents the design of our proposed graph data model in detail, elucidating its structure, components, and how it addresses the specific requirements of document archiving and retrieval. Finally, in Section 6, we draw conclusions based on our research findings and the insights gathered throughout this paper, summarizing the key takeaways and potential implications of our proposed graph data model in the realm of document archiving.

2. Related research

In recent years, the utilization of graph databases for archival purposes has gained increasing prominence. This surge in interest is closely tied to the growing adoption of graph database technology across various domains. The proliferation of graph databases has not only transformed the landscape of data management but has also invigorated research efforts within the archival community. This emerging field is characterized by the convergence of archival science with computational techniques and technologies, indicating a significant shift in how archives are managed and leveraged in the digital age. Some of the most notable research is the work of Laclavík *et al*. from 2011 and 2012, related to the extraction and construction of a graph database (i.e. a social network) from an e-mail archive [3][4]. Based on the case of historical records about real-world individuals, there is research that strives to analyze and compare both graph-oriented databases and document-oriented databases in terms of flexibility, precision of record storage, as well as the ease of ingestion and accuracy of record presentation within the database [5]. In a simple experiment using a graph database, Conrad & Williams (2020) have demonstrated that it is possible to increase the number of access points to individual items in archival collections by leveraging existing machine-readable and searchable data and metadata to identify and display relationships between persons, places, dates, events, etc. across items and collections [6]. Further on, there is research based on the case study of Torre do Tombo National Archives in Portugal aimed at moving from hierarchical, ISAD-conforming descriptions to graph counterparts, which requires utilization of state-of-the-art technologies, data models and vocabularies, such as the design of the core data model for archival records represented as the ArchOnto ontology and its embodiment in the ArchGraph knowledge graph, including a graph database and an Object Graph Mapping library [7] [8].

All the previously mentioned research works corroborate the hypothesis that there exists a profound connection between physical archives and their digital counterparts in the form of graph databases and knowledge graphs. Furthermore, they serve as compelling evidence supporting what Marciano (2022) has termed ‘Computational Archival Science (CAS)’ [9].

3. Briefly on archiving documents

As per the Macedonian Law on Archive Material, archive material encompasses all original and reproduced/duplicated documentary material (in written, drawn, printed, photographed, filmed, audiovisual, typed, electronic, digital, optical, or any other recorded form) of enduring significance. This material holds importance for the nation, science, academic research, cultural heritage, custodians, and various other necessities [10]. According to this law, a document or record is recorded information generated or received in the course of commencing, conducting, or concluding an undertaking by a specific authority, legal entity, or individual. It must adequately encompass the content, context, and structure of that activity in a manner that, regardless of the document’s form or medium, offers proof of that particular activity.

Documentary material, as a source of archival content, encompasses all original and duplicated material (including written, drawn, printed, photographed, filmed, audiovisual, typescript, electronic, digital, optical, or any other recorded form). This category also includes books and other recorded documents generated during the routine activities of the custodians. It remains relevant to their ongoing work until the point at which the archival material is segregated from it.

The document or record is characterized by four fundamental attributes: authenticity, reliability, integrity, and usability.

a) *Authenticity*: A document is considered authentic when it can be substantiated as what it purports to be. It should be verifiable as having been created or dispatched by the claimed author or sender and at the claimed time of creation or dispatch;

b) *Reliability*: A document attains reliability when its content is completely trustworthy, accurately representing the activities or facts it encompasses;

c) *Integrity*: A document is deemed to have integrity when it remains whole and unaltered, free from any tampering or modification; and

d) *Usability*: A document is considered usable when it allows for easy identification, location, searchability, presentation, and interpretation.

Archival and documentary material can be categorized as conventional or unconventional. Conventional archival and documentary material refers to records that are documented on paper and can be read without the need for any special devices or equipment. Non-conventional archival and documentary material encompasses records that are documented on specialized mediums, such as microfilm, optical-magnetic media (e.g., hard disks), compact disks (CDs), DVDs, flash memory, or holographic disks. Specialized devices are necessary to access and read the content within this material. Non-conventional material includes electronic documents generated using computers and other electronic devices in digital formats, data stored in databases that serve as the basis for creating conventional documents through data processing, and digital images obtained by digitization of traditional documents.

Archival work encompasses several essential processes, including the selection of archival material from the documentary material, organizing and arranging both archival and documentary material, documenting (inventory and description) of the archival material, categorizing the archival content, listing the documentary material whose retention period has expired, disposing of the identified documentary material, providing accommodation, storage, and protection for archival and documentary material in proper facilities equipped with standard equipment and technical-technological resources. Finally, it involves the transfer of archival material to permanent storage within the State Archives.

The basis of our work are office and archival operations of public institutions and services (including universities and faculties), state bodies and institutions, public enterprises, and alike. According to the Decree on office and archival operations, in all of these entities, archival operations include a large number of activities related to managing documents/records, such as receiving, reviewing, and sorting by organizational units and officials, filing the received document in a ledger and auxiliary record books, combining multiple documents/records, delivery, administrative-technical processing that includes signing by an authorized official, putting an official seal and/or stamp, sending, separating and classifying, and archiving [11][12]. The research methodology we have used was based on a thorough and systematic document analysis of relevant legal documents ([10], [11], [12]), which helped to gather information, identify patterns, and draw relevant conclusions.

4. Briefly on graph databases

In graph theory, a graph is a mathematical representation of a set of objects (vertices or nodes) connected by relationships (edges). A graph database represents a specialized type of database management system that is designed to store, manage, and query data using graph structures. As defined by Angles & Gutierrez in 2008, a graph database can be described as a database in which the data structures for both the schema and instances are represented as a (labeled) (directed) graph or generalizations of the graph data structure. In such a database, data manipulation is achieved through graph-oriented operations and type constructors, and it includes integrity constraints suitable for the graph's structure [13]. In a more formal context, a graph database schema takes the form of a graph $G_{db} = (N, E, \psi, V, \lambda)$, where N is a set of nodes and E is a set of edges; ψ is an incidence function $E \rightarrow N \times N$; V is a set of labels and λ is a labeling function $N \cup E \rightarrow V$.

Any graph database leverages all the graph concepts to model, store, and retrieve data in a way that is particularly suited for data with complex and interconnected relationships.

Informally, a graph database can be defined using graph theory concepts in the following manner:

- *Vertices (Nodes)*: In a graph database, each data element is represented as a vertex or node. Nodes can store attributes or properties that define the characteristics of the data element. These properties provide additional information about the data and can be used for querying and analysis;
- *Edges (Relationships)*: Edges represent the connections or relationships between nodes in the graph. These relationships can have different types and can carry additional information in the form of properties. Edges in a graph database can be directed (going from one node to another) or undirected (bidirectional);
- *Labels and Types*: Graph databases often allow nodes and edges to be labeled or typed. Labels or types categorize nodes and edges into different groups, allowing for efficient querying and manipulation based on these categories;
- *Properties (Attributes)*: Both nodes and edges in a graph database can have properties, which are key-value pairs that provide detailed information about the data;
- *Graph traversal*: Graph databases support the traversal operation, which involves navigating through the graph by following relationships between nodes. Traversal allows for efficient querying of complex patterns and paths in the data. Graph traversal algorithms, such as depth-first search (DFS) and breadth-first search (BFS), are commonly used to explore the graph's structure and relationships.

Graph databases have the intrinsic capacity to handle more data, operate at higher speeds, and most significantly, facilitate more connections. They are inherently designed to efficiently store, process, and query relationships, making them highly proficient in managing intricate, interconnected data and complex queries. Importantly, this proficiency remains consistent regardless of the overall dataset size. By replacing the limitations of the conventional relational model with flexibility and speed, these attributes have contributed significantly to their widespread adoption and their recent surge in popularity [14].

5. Proposed graph data model

Designing an effective data model is foundational for graph databases. This section explores the complexities of modeling archival documents as nodes and their relationships to other object types within a graph database. The proposed graph database model is presented in Figure 1.

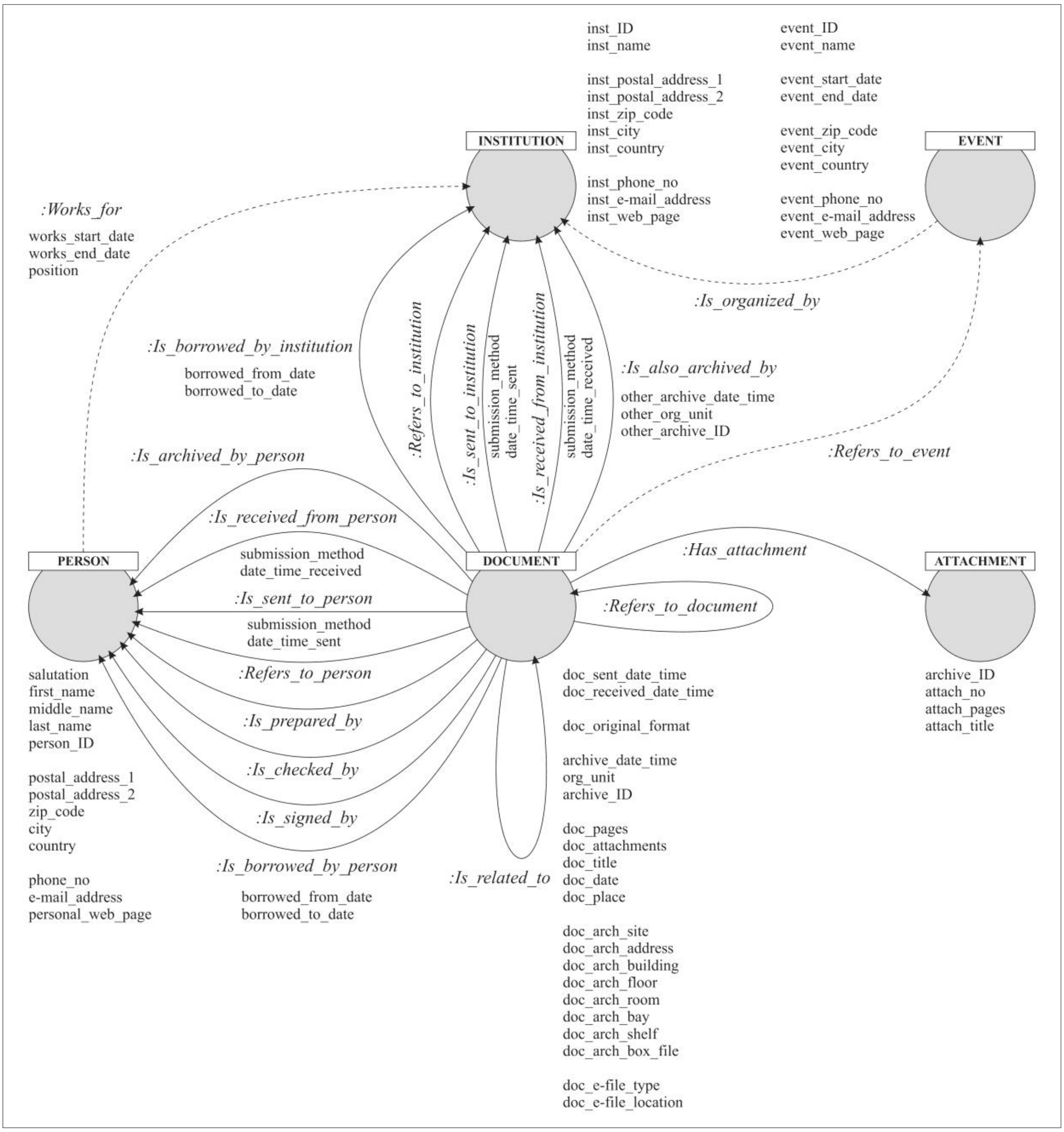

Figure 1: Graph database model (Source: Authors' representation)

In this model, the primary focal point is the 'DOCUMENT' node type. As a representation of an actual archived document, it encompasses various attributes including details like the date and time of reception (*doc_received_date_time*) and archiving (*archive_date_time*), the organizational unit (*org_unit*) associated with it, and a distinctive archiving identifier (*archive_ID*). Additionally, if the document was submitted to the organization via postal mail or electronic mail, the date and time of submission are also recorded (*doc_sent_date_time*). Furthermore, this node encompasses general attributes that provide a comprehensive description of the document, including the number of pages (*doc_pages*), the count of document attachments (*doc_attachments*), the official document title (*doc_title*), as well as the document's date of issuance (*doc_date*) and the place where it was issued (*doc_place*). Finally, the 'DOCUMENT' node type maintains a record of the physical location where the archived document is stored. This physical location adheres to a structured hierarchy, encompassing various levels, including the document archival site (*doc_arch_site*), address (*doc_arch_address*), building (*doc_arch_building*), floor (*doc_arch_floor*), room (*doc_arch_room*), bay (*doc_arch_bay*), shelf (*doc_arch_shelf*), and box_file (*doc_arch_box_file*).

Two reflexive relationships are associated with the 'DOCUMENT' node type: the '*:Referes_to_document*' and the '*:Is_related_to*' relationships. The first relationship establishes a link between a specific document and one or more other documents that are referred to within the text of that document. The second relationship signifies the connection between a particular document and one or more related documents. These reflexive relationships enhance the model's ability to represent and navigate the intricate web of document references and relationships within the archival system.

Closely linked to the 'DOCUMENT' node type is the 'ATTACHMENT' node type. Within the graph data model, each attachment associated with a document is depicted as a distinct 'ATTACHMENT' node, and their connection is illustrated through the presence of the '*:Has_attachment*' relationship. The relationship between these two node types is established through the shared value of the unique archiving identifier attribute (*archive_ID*). In instances where a document includes multiple attachments, each attachment is assigned a sequential number (*attach_no*) for identification. Additionally, for each attachment, the model records the number of pages (*attach_pages*) it comprises, as well as its title (*attach_title*).

In the real-world processes of physically creating and handling documents, various individuals play different roles, all of which are symbolically represented by the 'PERSON' node type within the graph data model. Besides the *person_ID* attribute that uniquely identifies each person, this node encompasses all typically documented attributes that store essential information about individuals involved in these processes, providing a comprehensive profile of each person involved in document-related activities.

Numerous relationships can be established between the 'DOCUMENT' and 'PERSON' node types. For instance, during the document creation process, a document '*:Is_prepared_by*' a person, a document '*:Is_checked_by*' a person, and a document '*:Is_signed_by*' a person. These three relationships help capture the various roles and responsibilities individuals have in the document's lifecycle, offering a detailed view of the document's history and the people involved. In the context of document handling, a specific document can either be received from a particular person (*:Is_received_from_person*) or sent to a particular person (*:Is_sent_to_person*). These two relationships can incorporate additional attributes such as *submission_method*, *date_time_received*, and *date_time_sent*, respectively, which provide additional contextual information about the manner and timing of document reception and transmission, enhancing the comprehensiveness of the archival representation. Certainly, in the context of the graph data model, establishing other relationships between the 'DOCUMENT' and 'PERSON' node types is entirely feasible. For instance, the '*:Refers_to_person*' relationships can signify instances where a document makes references to specific individuals, whilst the '*:Is_borrowed_by_person*' relationships could be used to represent situations where a previously archived document is borrowed by specific individuals. The latter one is supposed to have attributes such as '*borrowed_from_date*' and '*borrowed_to_date*'.

Yet another node type, labeled 'INSTITUTION', is required in the graph data model. Each institution is distinctively identified by the value of the *inst_ID* attribute and it can encompass other pertinent data that are customarily recorded for institutions. This node type enables the model to effectively represent and manage information related to various institutions involved in document-related processes, enriching the overall context of the archival system.

The two basic relationships that can link the 'DOCUMENT' node type to the 'INSTITUTION' node type are '*:Is_sent_to_institution*' and '*:Is_received_from_institution*' that include attributes such as *submission_method*, *date_time_received*, and *date_time_sent*, respectively, which provide additional contextual information about the manner and timing of document reception and transmission. As in the case of the 'PERSON' data type, establishing other relationships between the 'DOCUMENT' and 'INSTITUTION' node types is entirely feasible. For instance, the '*:Refers_to_institution*' relationships can signify instances where a document makes references to specific institutions, whilst the '*:Is_borrowed_by_institution*' relationships could be used to represent situations where a previously archived document is borrowed by specific institutions. The latter one is supposed to have attributes such as '*borrowed_from_date*' and '*borrowed_to_date*'.

The model also accommodates complex scenarios in which a single document is archived by multiple institutions. This is accomplished by introducing a relationship called '*:Is_also_archived_by*' that connects the 'DOCUMENT' and 'INSTITUTION' node types. This relationship encompasses properties derived from the other institution's timestamp, which includes details such as the date and time of archiving by the other institution (*other_archive_date_time*), the organizational unit of the other institution (*other_org_unit*), and the unique archiving identifier of the other institution (*other_archive_ID*).

In addition to the node types already mentioned, the graph data model can be further enhanced by incorporating additional node types and relationships, which can elevate the graph from a standard representation to a knowledge graph. One such node type is the 'EVENT' node type, which encompasses distinct properties such as the event unique identifier (*event_ID*), event name (*event_name*), and other pertinent data like event start date (event_start_date), event end date (event_end_date), and so on. This 'EVENT' node is interconnected with the 'DOCUMENT' node type through the '*:Refers_to_event*' relationship. This extension enriches the model by associating documents with specific events, fostering a deeper understanding of the contextual relationships between documents and significant events. Another valuable enhancement to the existing graph data model is the introduction of the '*:Works_for*' relationship, establishing connections between the 'PERSON' and 'INSTITUTION' node types. By including this relationship, the model gains the ability to represent the professional associations between individuals and institutions, providing valuable insights into employment history and roles within organizations. This relationship can encompass attributes like the start date of employment (*works_start_date*), the end date of employment (*works_end_date*), and the individual's position (*position*) within the institution.

6. Conclusions

Utilizing databases for archival purposes is a strategic approach to efficiently manage, preserve, and provide access to historical records and documents. A well-designed archival database can facilitate research, ensure the long-term preservation of valuable materials, and contribute to the broader dissemination of historical knowledge. Graph databases are particularly well-suited for archival document management because they include scenarios where relationships and connections between data elements (i.e. archived documents) are crucial.

The proposed property graph model serves as a foundational theoretical framework that can pave the way for the development of more sophisticated data models within the domain of document archiving. As such, it provides a versatile and extensible blueprint upon which future models can be built, offering the potential to address evolving needs in document management, preservation, and retrieval. By embracing this theoretical framework, researchers and practitioners can explore innovative approaches, incorporate additional node types and relationships, and adapt the model to suit specific archival contexts, ultimately advancing the field of document archiving and archival science as a whole.

The implementation of the proposed property graph model in any of the existing graph databases (e.g., Neo4j, TigerGraph, Amazon Neptune, …) can be highly beneficial because all graph databases provide query languages (e.g., Cypher for Neo4j, GSQL for TigerGraph, Gremlin for Amazon Neptune, …) that can allow users to express complex graph patterns and retrieve specific information from the graph. Queries can range from simple node or edge lookups to intricate path and relationship queries. In addition, graph databases employ indexing techniques to speed up queries. These indexes help locate nodes, edges, and relationships quickly, making queries that involve traversing the graph more efficient. Indexing strategies are critical to maintaining good database performance as the dataset grows, which is one of the fundamental intrinsic characteristics of archives.

Some potential limitations of using graph databases for archival purposes are related to the growth in the volume of archived documents and associated data. These include issues such as scalability, complexity of queries, query performance, storage requirements, maintenance, and costs. Additionally, there are challenges related to data migration when transitioning from traditional document management systems to a graph-based approach, as well as concerns about interoperability when integrating the graph model with existing archival systems or external databases.

Future work related to the proposed design can encompass further refinement of the schema to accommodate additional metadata attributes or node types specific to certain archival domains. It can involve: (a) Adapting the model to support different document types, such as images, audio, or video, and considering the unique metadata associated with each; (b) Incorporating semantic technologies like RDF to enhance the model's ability to represent and query the relationships with greater precision; and (c) Continuous assessment of the graph data model's real-world effectiveness and efficiency, involving thorough validation and benchmarking against archival standards.

References:

[1] The National Archives, Archive Principles and Practice: an introduction to archives for non-archivists, 2016. URL: https://cdn.nationalarchives.gov.uk/documents/archives/archive-principles-and-practice-an-introduction-to-archives-for-non-archivists.pdf

[2] J. Hare, M. Beyer, P. den Hamer, A. Jaffri, R. Sallam, Understanding When Graph Analytics Are Best for Your Business Use Case, Gartner Research, 2022. URL: https://www.gartner.com/en/documents/4013850.

[3] M. Laclavík, Š. Dlugolinský, M. Kvassay, L. Hluchý, Email Social Network Extraction and Search, NextMail 2011 Workshop, WI-IAT 2011, Lyon, France, in: Proceedings of the 2011 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, IEEE Computer Society, 2011, pp. 373–376. doi: 10.1109/WI-IAT.2011.30.

[4] M. Laclavík, Š. Dlugolinský, M. Šeleng, M. Ciglan, L. Hluchý, Emails as Graph: Relation Discovery in Email Archive, WWW 2012 EMAIL '12 Workshop, Lyon, France, in: Proceedings of the 21st International Conference on World Wide Web, Association for Computing Machinery pp. 841–846. doi: 10.1145/2187980.2188210.

[5] M. Healy, T. O'Connor, J. Keating, Comparison of Graph- and Collection-Based Representations of Early Modern Biographical Archives, in: T. Andrews, F. Diehr, T. Efer, A. Kuczera, J. van Zundert (Eds.), Graph Technologies in the Humanities 2020, CEUR Workshop Proceedings 2020, Vol-3110, Vienna, Austria, 2020, pp. 60–85.

[6] M. Conrad, L. Williams, Elevating "Everyday" Voices and People in Archives through the Application of Graph Database Technology, in: Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 2020, pp. 1861−1865. doi: 10.1109/BigData50022.2020.9377921.

[7] I. Koch, N. Freitas, C. Ribeiro, C. T. Lopes, J. R. da Silva, Knowledge Graph Implementation of Archival Descriptions through CIDOC-CRM, in: A. Doucet, A. Isaac, K. Golub, T. Aalberg, A. Jatowt, (Eds.), Digital Libraries for Open Knowledge, in: Proceedings of the 23rd International Conference on Theory and Practice of Digital Libraries, TPDL 2019, Oslo, Norway, volume 11799 of Lecture Notes in Computer Science, Springer Nature Switzerland AG, Cham, Switzerland, 2019, pp. 99–106. doi: 10.1007/978-3-030-30760-8_8.

[8] L. Costa, N. Freitas, J. R. da Silva, An Evaluation of Graph Databases and Object-Graph Mappers in CIDOC CRM-Compliant Digital Archives, Journal on Computing and Cultural Heritage, 15.3 (2022): 1−18. doi: 10.1145/3485847.

[9] R. Marciano, AFTERWORD: Towards a new Discipline of Computational Archival Science (CAS), in: L. Jaillant (Ed.), Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections, Bielefeld University Press, Bielefeld, Germany, 2022, pp. 205–218. doi: 10.14361/9783839455845.

[10] State Archives of the Republic of Macedonia, New Draft-Law on Archive Material, 2012. URL: https://arhiv.mk/wp-content/uploads/2021/01/predlogzakon_za_arhivski_materijal.pdf

[11] Government of the Republic of Macedonia, Decree on Office and Archival Operations [in Macedonian], Official Gazette of Republic of Macedonia, No. 1 from January 3rd, 2014, pp. 4–8. URL: https://www.slvesnik.com.mk/Issues/fc82ee320a774f22a861ae2a5392965e.pdf

[12] State Archives of the Republic of Macedonia, Instructions on Methods and Techniques for Managing Archive and Documentary Materials in Office and Archive Work [in Macedonian], Official Gazette of Republic of Macedonia, No. 99 from July 2nd, 2014, pp. 23−51. URL: https://www.slvesnik.com.mk/Issues/86bef7078a6b4be483a8f93b79f75402.pdf

[13] R. Angles, C. Gutierrez, Survey of Graph Database Models, ACM Computing Surveys, 40.1 (2008): 1−39. doi: 10.1145/1322432.1322433.

[14] I. Robinson, J. Webber, E. Eifrem, Graph Databases: New Opportunities for Connected Data, 2nd. ed., O'Reilly, Sebastopol, CA, 2015.