

Ivančica Sabadin¹
Peter Pavel Klasinc²

INFLUENCE OF SEMANTIC WEB ON ARCHIVAL SCIENCE

Abstract

Purpose: *The purpose of this paper is to describe the concept of the Semantic Web, as well as the technologies necessary for its implementation. The paper analyses current research on the use of Semantic Web technologies in the GLAM field. The aim is to find out what are the results of the research and what are the advantages and disadvantages of using Semantic Web technologies in the field of GLAM.*

Method/approach: *A descriptive method was used to describe a brief history of the Web, the concept and technologies of the Semantic Web, and the Records in Contexts model. A literature review method was used to analyse previous research on the use of the Semantic Web in the field of GLAM (Galleries, Libraries, Archives, Museums) and in the field of archival studies.*

Results: *Twenty articles from the databases ProQuest, SCOPUS, Emerald Insight and UNPAYWALL were analysed in the context of the Semantic Web and its application in the field of archival studies. Eleven articles related to GLAM and nine articles related to archives were identified and analysed.*

Conclusions/findings: *More and more institutions in the field of GLAM are publishing their inventories online and, in this context, there is a growing body of research on the use of Semantic Web technologies to facilitate the identification of material and to increase accessibility and usability.*

Keywords: *Semantic Web, Linked Data, Archives, Records in Context*

1 Ivančica Sabadin, PhD student of Archival Sciences at Alma Mater Europaea – European Centre Maribor, Slovenia, e-mail: ivancica.sabadin@almamater.si.

2 Assoc. prof. dr. Peter Pavel Klasinc, Director of the International Institute of Archival Sciences Trieste-Maribor in Trieste (Italy). Dean of the study programs of Archival sciences at AMEU ECM, peter.klasinc@almamater.si, +386 41 673 785.

1. INTRODUCTION

The Web, or the “World Wide Web” as we know it today, was invented in 1989 by the British scientist Sir Tim Berners-Lee. The original Web was designed to help scientists research. The basic feature of the web is a hyper-text file that contains links to other files and can be viewed using a special tool or browser. The first website address was <http://info.cern.ch/hyper-text/WWW/TheProject.html> and is still accessible on the web. By the end of 1993, there were about 500 web servers. The year 1994, or “Year of the Web”, was particularly important for the development of the Web, with two conferences on Web technologies. By the end of 1994, there were 10,000 servers on the Web - 2,000 of which were commercial - and 10 million users (CERN, n. d.).

The concept of the Semantic Web was mentioned in an article written in 2000 by the creator of the Web, Sir Tim Berners Lee. He said that the goal of the Semantic Web is to be able to process data in a way that is similar to databases and mathematical formulas (Berners- Lee et al., 2000). As previously discussed, the foundation of the Web relies on the creation of hyper-links between documents, yet no such links exist between data. The Web is primarily designed to facilitate human use, not computer applications and data management. For this reason, the Semantic Web, as an extension of the standard Web, allows the structure of the Web to be extended and machine-readable semantics to be created. In this way, we are creating a Web that makes sense and helps us find information, make decisions and solve problems.

To create the Semantic Web or the Web of Data, we need a large amount of data in a standard format that can be accessed and managed by the tools of the Semantic Web. In addition, it is not enough to have access to the data, but the links between the data must also be available (Song, 2014, 6). *“Linked Data refers to a set of best practices for publishing structured data on the Web”* (Linked Data, n. d.). Several technologies can be used to create linked data, but the three basic technologies are: XML, RDF and ontologies. *“Extensible Markup Language, abbreviated XML, describes a class of data objects called XML documents and partially describes the behaviour of computer*

programs which process them" (W3C, 2008). The foundation of XML comprises an entity that can possess connections to other entities and is a component of the XML file's physical arrangement. The logical structure consists of declarations, elements, comments, character references and processing commands (W3C, 2008). An RDF model is generated from the XML file.

"RDF is a data model used to organize semantic data in the Semantic Web" (Liu and Hong, 2021, 1). The output of the RDF model is a knowledge graph containing entities, attributes and relationships. The foundation of RDF and the Semantic Web are entity-attribute-relationship triples. *"In the knowledge graph, triples are called knowledge"* (Liu and Hong, 2021, 1).

"Ontology is the branch of metaphysics dealing with the nature of being" (Oxford reference, 2023). *"Ontologies define a common vocabulary that is used for sharing information in a certain domain. They consist of machine-interpretable definitions of concepts in the domain and the relationships among them. Developing ontologies is useful for the sharing of common understanding of the structure of information among people and software agents"* (Azwari, 2016, 20).

In 2012, the International Council on Archives (ICA) began work on a new standard for describing records based on archival principles. The first version of the new conceptual model, called Records in Contexts (RIC), was published in September 2016. A new full version of the conceptual model would follow five years later, in July 2021. The RIC consists of three documents: The Records in Contexts - Introduction to Archival Description (RiC-IAD), the Records in Contexts Conceptual Model version 0.2 (RiC-CM 0.2) and the Records in Contexts-Ontology (ICA RiC-O). The RIC model takes into account the latest developments in the field of information and communication technologies, as well as archival principles, and is based on the concepts and technologies of the Semantic Web and Linked Data.

2. METHODOLOGY AND LIMITATIONS OF THE RESEARCH

The purpose of this paper is to describe the concept of the Semantic Web and the technologies required to make it a reality. The paper analyses the use of Semantic Web technologies in the GLAM domain, aiming to deter-

mine the research in this area and the results, advantages, and disadvantages of using Semantic Web technologies.

The research examines scientific articles from various databases, including ProQuest, SCOPUS, Emerald Insight, and UNPAYWALL. The study was conducted using the Boolean AND operator to search for articles containing the keywords 'semantic web', 'GLAM', and 'archives' with the following inclusion and exclusion criteria.

Inclusion criteria:

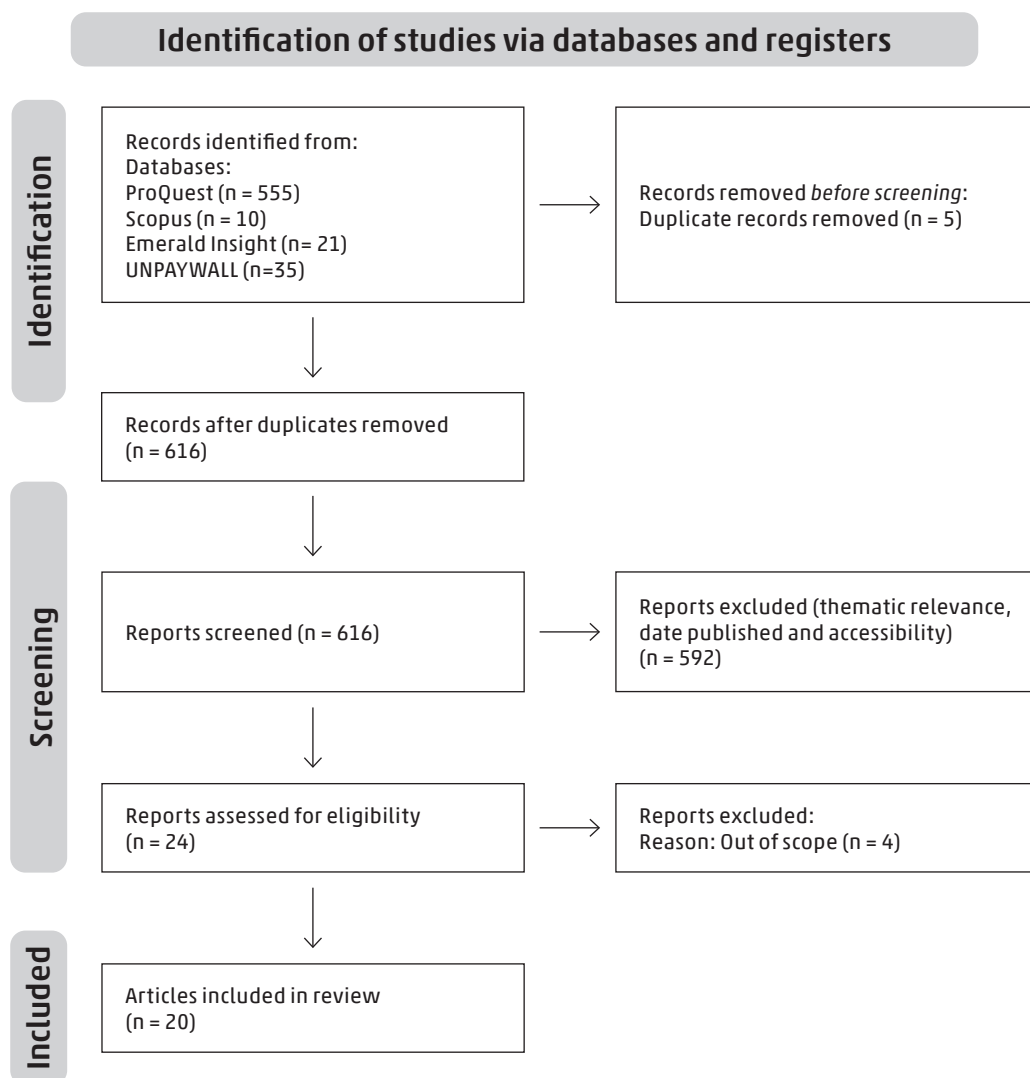
1. Published in English, Slovenian or Croatian languages
2. Published between 2013 and 2023
3. The accessibility of the full text
4. Thematic relevance

Exclusion criteria:

1. Published in other languages
2. Published before 2013
3. Only abstract is available
4. No relevance to the subject.

A total of 24 scientific articles were retrieved, four of which were excluded because they were not in line with the research topic (see Figure 1).

Figure 1: Prisma flow chart (Page et al., 2021, 5)



The analysed articles are presented in an organised manner, with a focus on those that describe the use of Semantic Web technologies in the 'GLAM' domain, encompassing galleries, libraries, archives, and museums. This is followed by an analysis of the articles that describe the use of Semantic Web technologies in archives.

3. THE IMPACT OF SEMANTIC WEB TECHNOLOGIES ON THE GLAM DOMAIN

"The web is emerging as a preferred platform to publish open data and to inter-link it together. This has result in the Web of Documents to be emerged as Web of Data" (Lytras et al., 2019, 3).

"Memory institutions such as Galleries, Libraries, Archives and Museums (GLAM) are the main players that collect, organize, disseminate and preserve CHI³ in various cultural domains" (Wijesundara and Sugimoto, 2018, 59).

During the past decade, there has been a substantial rise in the quantity of digital collections being published by GLAM institutions, remarked Gustavo Candela (2023, 2). These collections contain diverse materials such as maps, images, OCR text, sound, and video.

There has been a rising interest in implementing Semantic Web technologies in the GLAM sector, particularly in the utilization of linked data. Summers and Salo (2013, 5) believe that the reasons for this interest are part of the institutional missions of these organizations: collection, preservation and providing access to records. The authors emphasised the need to create descriptions of the cultural values held by the institutions and the need to share the elements of description. The authors also highlighted the specificity of archives compared to museums and libraries: *"there is too much material to be described at the item level, so description tends to be done at the higher, more abstract level of series or subseries"* (Summers and Salo, 2013, 5). For the authors, the area of enabling permanent identification and digital preservation is the most important area that libraries, archives and museums have in common with linked data. Cultural institutions are responsible for preserving cultural assets for use both now and in the future. *"An essential part of this custodial responsibility is the proper identification of the artifacts that make up their collections"* (Summers and Salo, 2013, 7).

"The application of the Semantic Web principles by GLAM organisations has several benefits including: i) the description of the information with machine-read-

able, standard and controlled vocabularies; ii) the enrichment with external repositories; and iii) the reuse by third-party actors" (Candela, 2023, 1).

"Linked Data provides a viable means of archival datafication capable of implementing the FAIR⁴ Guiding Principles, creating machine-readable, interoperable, extensible Archival Linked Data suited to interrogation and analysis using digital humanities research methods. Using Linked Data, archival data (catalogue data, metadata, data extracted from the contents of born-digital and digitised archives) can be embedded into the web, enriching and further contextualising archival data, and making it easier to discover, access, and utilise" (Hawkins, 2021, 321).

Mariana Ziku (2020, 2-3) highlighted the need to shift the GLAM field towards interdisciplinary and data-driven approaches. The author contends that the concept of the Semantic Web is driving the implementation of computational techniques in the management and analysis of cultural heritage, which is highlighting the importance of providing open, organized, and interoperable data.

Standards for the implementation of Semantic Web technologies have either been established or are presently being developed in cultural institutions. For libraries, the Library Reference Model (LRM) has been developed by IFLA.⁵ Golub et al. (2021, 216) state that the LRM model corresponds to the ISO standard CIDOC-CRM⁶, which was developed by the International Documentation Committee of the International Council of Museums (ICOM) in 2006. It is important to note that CIDOC-CRM concentrates on events and processes, whereas the IFLA LRM predominantly models process outcomes (Golub et al., 2021, 216).

Melo et al. (2023, 556) highlight that the RIC-O⁷ and CIDOC-CRM ontologies are being used to represent archival information. The RIC-O model is an OWL⁸ ontology that represents archival records and their contextual enti-

4 FAIR - the Findability, Accessibility, Interoperability, and Reuse. More information available at: <https://www.go-fair.org/fair-principles/>

5 IFLA - International Federation of Library Associations and Institutions

6 CIDOC-CRM - CIDOC Conceptual Reference Model

7 RIC - O - Record in Context Ontology

8 OWL - The W3C Web Ontology Language

ties. A tool has been developed to transform ICA⁹ records from the French National Archives into a contextual standard known as ICA RiC. *"CIDOC-CRM, which is a model widely used in the Heritage domain, and based on well-documented experiments of modelling Museums, Archaeology, and Architecture domains"* (Melo et al., 2023, 556). Melo et al. (2023, 556) claim that the use of CIDOC-CRM can facilitate semantic integration and interoperability due to the wealth of platforms that provide information in the CIDOC-CRM standard across different domains. The use of the CIDOC-CRM model ensures that, firstly, there is a wealth of information available in the field of cultural heritage for integration and, secondly, there are numerous platforms available for the exploration and retrieval of migrated information (Melo et al., 2023, 576).

3.1 EXAMPLES OF THE USE OF SEMANTIC WEB TECHNOLOGIES IN THE GLAM DOMAIN

Due to the growing use of Linked Data and Semantic Web technologies, a rising number of research studies and projects within this sphere are now accessible. In Sweden, Golub et al. (2021) conducted research on the digital access to cultural heritage across 91 museums. The study utilized 21 different criteria to analyse the feasibility of retrieving information about museum materials. It has been found that users primarily search by subject, posing a significant obstacle for museums as users must be familiar with both the subject itself and the method of information-seeking utilised. This is compounded by the challenge of understanding the exact meaning of terms, synonyms, homonyms and multiple meanings. The authors conclude that *"there is a strong need for the implementation of established controlled vocabularies in museums more widely, not only in Sweden"* (Golub et al., 2021, 242).

Meghini et al. (2019) conducted research on the utilization of narratives on Europeana. The authors perceive the narrative *"as a semantic network, meaningful to the user, consisting of events related to one another, to the entities that compose the events (e.g., agents, places, time, physical objects) and to the digital objects through semantic relations"* (Meghini et al., 2019, 8). The

9 ICA – International Council on Archives

authors note that while the Europeana portal features a metadata-based search engine for digital objects, it does not currently support event-based search or user-friendly visualisation. Within the project, a tool for creating and visualising narratives or NBVT¹⁰ has been developed. *"The NBVT takes as input resources inserted manually by the user or imported automatically from Wikidata"* (Meghini et al., 2019, 11). Wikidata utilizes Semantic Web and Linked Data technologies to systematically gather structured data from related projects, such as Wikipedia, Wikisource, and Wikibooks. Meghini et al. (2019, 11) reported that the portal hosts over 25 million entities and facilitates data export via Wikidata API and Wikidata Query Service, running SPARQL¹¹ in the background. The primary objective of the NBVT tool is to enrich libraries with a semantic event network, providing relevant context to the digital objects they hold.

Bianchini et al. (2021) provided an overview of VIAF, a crucial tool for identifying entities (people, locations, works and expressions) associated with a bibliographical domain. Bianchini et al. (2021, 2) concluded that for improved integration of libraries within the Semantic Web, it is necessary to incorporate larger stakeholder groups including non-national agencies, museums, archives and users, while employing a bottom-up approach. *"Wikidata is a freely available hosted platform that anyone—including libraries—can use to create, publish, and use Linked Open Data (LOD)"* (Bianchini et al., 2021, 2). Interest in using Wikidata to publish Linked Open Data (LOD) for GLAM (Galleries, Libraries, Archives and Museums) institutions is gradually increasing. *"Libraries' interest in Wikidata is usually focused on LOD and semantic discovery"* (Bianchini et al., 2021, 5). The authors say that libraries show great interest in Wikidata due to LOD and semantic insights. In this paper the comparison between VIAF and Wikidata is described in detail. Initially, VIAF and Wikidata datasets are analysed to compare VIAF and Wikidata clusters to establish their relationships. Based on relevant websites and literature, several general aspects are compared, including scope, aims and identification, and differences and similarities are explored and

10 NBVT - Narrative building and visualising tool

11 SPARQL - SPARQL Protocol and RDF Query Language (recursive acronym)

characterised. *"The semantic web has offered new important tools and chances to libraries, archives, museums and other cultural institutions, and their data are recognized as a relevant asset for building the backbone of the semantic web as to the control of entities of bibliographic and cultural interest. After eight years of existence, Wikidata is playing a relevant role in the publication, aggregation, and control of bibliographic and non-bibliographic information in the semantic web too"* (Bianchini et al., 2021, 23). The authors state that Wikidata relies on VIAF for identifying objects, and hence acknowledges the crucial position of VIAF as the initial system in the data identification scheme for each object (Bianchini et al., 2021, 23).

Gustavo Candela (2023) presented his creation of a structure called the "Data Foundry", which aims to transform metadata sets into the LOD format as it is published by the relevant GLAM institutions. The Moving Image Archive of the National Library of Scotland, the National Catalogue of Scotland and the Catalogue of Translated Literary Works of Scotland were the three datasets used by the framework. *"The evaluation showed that the framework can be useful for other organisations willing to publish datasets as LOD following best practices. Future work to be explored includes the evaluation of additional datasets, the improvement of the framework to include additional types of datasets such as OCR and the exploration of data spaces to include the final datasets"* (Candela, 2023, 11).

Semantic web technologies formed the foundation of the INCEPTION¹² project which was funded by the European Union. Iadanza et al. (2019, 381) highlighted that the development of a cloud platform is the main achievement of the project. The platform aims to achieve the project's objectives of accessing, understanding and enhancing European cultural heritage through enriched 3D models. The integration between Building Information Modelling (BIM) technological solutions and cutting-edge internet-based technologies underpins the project. The INCEPTION platform has been introduced by the authors as a medium for exchanging knowledge amongst professionals, students, scientists, curators, ordinary users, and others. *"The Semantic Web structure interlinks the platform with exter-*

12 INCEPTION – Inclusive Cultural Heritage in Europe through 3D semantic modelling

nal Cultural Heritage available linked data and makes it gradually enhanced by specific flexible data structures provided as project specific ontologies" (Iadanza et al. 2019, 381). *"A methodology of archiving digital data and linking them to the final product is one of the main outcomes. Before and during the creation of H-BIM, the nomenclature (vocabularies, thesaurus, etc.) is critical to maintain a common typology and to support interoperability"* (Iadanza et al. 2019, 387). In conclusion, the focus is on the discovery of solutions that enable the transformation of data in compliance with the technologies of the Semantic Web (Iadanza et al. 2019, 387).

In their 2018 paper, Yongming Wang and Sharon Q. Yang state that libraries have increasingly focused on implementing Semantic Web and linked data in the last decade. Their article offers theoretical foundations of the Semantic Web and Linked Data, along with a global survey of how libraries are transforming existing data into Linked Data formats. *"The road to Linked Data has been bumpy, but there is no way to turn back"* (Wang and Yang, 2018, 18). The authors predict that within the next five years, a significant majority of library data will transform into linked data because of its potential to aid unrestricted and straightforward exploration of online data. Libraries will face the challenge of retrieving bibliographic data from web browser search results. *"What libraries are trying to accomplish will benefit the society"* (Wang and Yang, 2018, 18).

Ranjgar et al. (2022) conducted a study that aimed to design a POI¹³-based data model suitable for cultural heritage organisations. Their work focused on the Sa'dabad complex in Iran, which after the 1979 revolution, underwent a transformation for the display of numerous cultural heritage objects and information relating to events from the monarchy era. The authors selected two museums, namely the Mellat Museum and the Museum of Fine Arts, to exemplify their objectives (Ranjgar et al., 2022, 4). The CIDOC-CRM standard is employed in producing a data model that connects historical data and POIs, while GeoSPARQL is utilised to generate spatial semantics. Ranjgar et al. (2022, 9) state that CIDOC-CRM has a distinct advantage as a top-level ontology, developed without a specific purpose

13 Points of interest

in mind. Ranjgar et al. (2022, 19) propose that additional investigation is necessary to explore the potential of integrating and enhancing the data model with open linked data and global knowledge bases.

4. EXAMPLES OF THE USE OF SEMANTIC WEB TECHNOLOGIES IN THE FIELD OF ARCHIVAL SCIENCE AND SCHOLARSHIP

"The QueryLab portal originates from the need to simultaneously query different archives available both locally and online, to facilitate and speed the process of searching data regardless of the language in which the content is expressed" (Artese and Gagliardi, 2022, 1). The paper outlines an ongoing study by authors Maria Teresa Artese and Isabella Gagliardi, which examines the integration of an ontology with the capacity to generate Linked Open Data for local heritage. The article presents an overview of QueryLab portal functionality, concerning tangible and intangible cultural material and offering tools for users with varying backgrounds and abilities. Artese and Gagliardi (2022, 4) state that QueryLab's design permits exploring various websites that include cultural heritage inventories. Such inventories come in different kinds: those available through web services and those only accessible via dedicated websites, the latter affecting the emergence of data silos. The authors (Artese and Gagliardi, 2022, 6) consider that mapping to appropriate ontologies and converting data into triples for publishing in the LOD cloud is crucial. The usability of the new QueryLab tool has shown to have a positive impact on search performance by visualising the results from various archives in different arrangements. The sharing, preservation and dissemination of cultural heritage is facilitated by the portal. *"We also intend to evaluate if and how the ontology allows to increase QueryLab content at the index level by integrating data exposed as LOD"* (Artese and Gagliardi, 2022, 18).

"In 1987, the French Ministry of Culture (Interdepartmental Service of Archives of France) developed the Thesaurus W. Standardized vocabularies for describing and indexing local administration records, in order to enable the French archival agencies to index descriptions of modern records created by local public services. The use of this thesaurus is mandatory for all territorial archives"

(Grimoüard, 2014, 206). The project aims to develop a machine-readable version of the indexing tool to be integrated into archival institutions' software. Grimoüard (2014, 208) argues that SKOS¹⁴ is a concept-oriented system, placing a significant emphasis on the description of concepts and the connections between them. RDF language is used to describe concepts, allowing them to be combined with properties. The platform, originally designed for archival institutions, has become a significant aspect of the Ministry of Culture's information system on controlled vocabularies. Grimoüard's (2014, 211) research indicates that the benefits for users include easier discovery, evaluation and reuse of vocabularies that can be integrated into websites to enrich the available information.

Goy et al. (2019) propose that ontologies can enhance accessibility to cultural resources, including historical documents. In their study, the authors outline two collaborative projects, Harlock900 and PRISMHA, with archives of cultural institutions that hold rich archival material, which demonstrate the utility of ontologies in this area. The authors contend that sufficient data, which has been suitably prepared, is not yet available to enable the implementation of contemporary technologies, including machine learning and artificial intelligence, within the domain of historical archives. *"When dealing with historical archives, metadata are usually very poor with respect to the needs of such techniques, and textual resources are frequently not obtainable"* (Goy et al., 2019, 288). The article aims to demonstrate that utilising a rich semantic model can lead to noteworthy enhancements in accessing archival resources, as noted by the authors. According to Goy et al. (2019), the two projects studied share an ideology that the proficiency of a system designed to avail historical resources can be elevated by enabling users to formulate queries based on concepts. The core functionality is bolstered by a semantic layer, exemplified by a semantic knowledge base (RDF triples), incorporating formal portrayals of the substance of archival documents. Two studies have *"shown that this approach provides better results if compared with standard access systems"* (Goy et al., 2019, 312).

14 Simple Knowledge Organization System

Javier Cha (2018) provides an introduction to the use of digital technologies in the field of cultural heritage in South Korea. *"In 2018, South Korea boasts large collections of heritage materials captured, archived and curated, using cutting-edge database technology"* (Cha, 2018. 228). The author enumerates various benefits of the archives in South Korea. These include an ontology of archival data that is available for online publishing, high-quality machine readability, the capacity to search through 336,267 court records that can be cited, an up-to-date database, and the use of contemporary technologies.

Semantic web technologies are utilised in the BBC¹⁵ archive to enhance content accessibility. Raimond et al. (2013) centre their paper on the "BBC World Service" archive, which employs a mix of Semantic Web technologies, automated linking, feedback, and data visualisation.

Demidova et al. (2014) describe the ARCOMEM software tool, which is used by the European Parliaments of Greece and Austria to create political archives from web and social network sources. *"Through ARCOMEM, the Greek and Austrian parliaments aspire to transform their flat digital content archives to historical and community memories"* (Demidova et al., 2014, 434). According to the authors, entities, topics, opinions and events are the main logical concepts of the software tool (ETOE¹⁶). The evaluation results indicated the significant semantic data types relevant for retrieving and displaying data. The fundamental logical concepts were topics and their related entities. *"Topic-driven search was very well received and was the preferred starting point for most of the tasks that included exploration of data. Overall, the evaluation verified the ARCOMEM digital preservation approach regarding the provision of many types of semantic data that enables the successful exploration of political web and social media archives"* (Demidova et al., 2014, 452).

5. CONCLUSION

The analysis examines 20 articles on the use of Semantic Web technologies in the GLAM sector. The research reveals that cultural institutions have increas-

15 British Broadcasting Corporation

16 ETOE - Entities, Topics, Opinions, Events

ingly published their catalogues and inventories, and have embraced new technologies, including Artificial Intelligence, over the past decade. *"Not so long ago, the records that were created in all spheres of society and which public archives were accepting for preservation and use, were tangible; those were boxes of paper documents, photographs and maps. Modern digital technologies, as we experience them today, are changing the way of creating records. They are becoming digital, and digital technologies are changing the way of their use and, above all, their quantity, which is growing exponentially"* (Hajtnik, 2019, 49). In order to successfully incorporate new technologies, cultural institutions must ensure that their data is machine-readable. This can be accomplished through the use of Semantic Web technologies, such as Linked Data.

In the field of GLAM, there are already established standards that facilitate the use of such technologies, including the ISO-approved CIDOC-CRM standard. The ICA has also recognised the necessity of utilising Semantic Web technologies and they are currently formulating a RIC conceptual framework for use in archives.

REFERENCES

- Artese, M. T., and Gagliardi, I. (2022). Integrating, Indexing and Querying the Tangible and Intangible Cultural Heritage Available Online: The QueryLab Portal. *Information*, 13(5), 260. Retrieved at: <http://dx.doi.org/10.3390/info13050260> (accessed on 9. 1. 2023)
- Azwari, S. A. (2016). *Updating rdf in the semantic web*. Glasgow: University of Strathclyde. Retrieved at: <https://www.proquest.com/dissertations-theses/updating-rdf-semantic-web/docview/1864754073/se-2> (accessed on: 10. 1. 2023)
- Berners-Lee, T., Connolly, D., Andrea Stein, L., Swick, R. (15. 8. 2000). The Semantic Web. *W3C*. Retrieved at: <https://www.w3.org/2000/Talks/0906-xmlweb-tbl/text.htm> (accessed on 9. 1. 2023)
- Bianchini, C., Bargioni, S., and Pellizzari di San Girolamo, C. C. (2021). Beyond VIAF. *Information Technology and Libraries*, 40(2). Retrieved at: <https://doi.org/10.6017/ital.v40i2.12959> (accessed on 9. 1. 2023)

- Candela, G. (2023). Towards a semantic approach in GLAM Labs: the case of the Data Foundry at the National Library of Scotland. *arXiv.org; Ithaca*. Retrieved at: <https://doi.org/10.48550/arXiv.2301.11182> (accessed on 9. 10. 2023)
- CERN (s.d.). Where the Web was born. Retrieved at: <https://home.cern/science/computing/birth-web/short-history-web> (accessed on 9. 1. 2023)
- Cha, J. (2018). Digital Korean studies: recent advances and new frontiers, *Digital Library Perspectives*, Vol. 34 No. 3, pp. 227-244. Retrieved at: <https://doi.org/10.1108/DLP-04-2018-0013> (accessed on 10. 1. 2023)
- Demidova, E., Barbieri, N., Dietze, S., Funk, A., Holzmann, H., Maynard, D., Papailiou, N., Peters, W., Risse, T. and Spiliotopoulos, D. (2014). Analysing and Enriching Focused Semantic Web Archives for Parliament Applications. *Future Internet*, 6(3), 433-456. Retrieved at: <http://dx.doi.org/10.3390/fi6030433> (accessed on 10. 1. 2023)
- Grimoüard, C. (2014). The Thesaurus for French Local Archives and the Semantic Web, *Procedia - Social and Behavioural Sciences*, Volume 147. Retrieved at: <https://doi.org/10.1016/j.sbspro.2014.07.153> (accessed on 10. 1. 2023)
- Golub, K., Ziolkowski, P.M. and Zlodi, G. (2021). Organizing subject access to cultural heritage in Swedish online museums, *Journal of Documentation*, Vol. 78 No. 7, pp. 211-247. Retrieved at: <https://doi.org/10.1108/JD-05-2021-0094> (accessed on 10. 1. 2023)
- Goy, A., Magro, D. and Baldo, A. (2019). A SemanticWeb Approach to Enable a Smart Route to Historical Archives. *Journal of Web Engineering*, 18(4-6). Retrieved at: <https://journals.riverpublishers.com/index.php/JWE/article/view/3149> (accessed on 9. 1. 2023)
- Hajtnik, T. (2019). Digital age: Time to transform of public archives. *Atlanti+*, 29(2), 49-65. Retrieved at: <https://journal.almamater.si/index.php/atlantipus/article/view/312/287>. (accessed on 12. 12. 2023)
- Hawkins, A. (2021). Archives, linked data and the digital humanities: increasing access to digitised and born-digital archives via the semantic web. *Archives and Museum Informatics* 22(5), 319-344. Retrieved at: <https://doi.org/10.1007/s10502-021-09381-0> (accessed on 9. 1. 2023)

- Iadanza, E., Maietti, F., Ziri, A. E., Di Giulio, R., Medici, M., Ferrari, F., Bonsma, P., and Turillazzi, B. (2019). Semantic Web Technologies Meet Bim For Accessing And Understanding Cultural Heritage, *The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, XLII-2/W9, 381–388, Retrieved at: <https://doi.org/10.5194/isprs-archives-XLII-2-W9-381-2019>. (accessed on 9. 1. 2023)
- Linked Data* (b. d.). Retrieved at: <https://www.w3.org/wiki/LinkedData> (accessed on 20. 11. 2023)
- Liu, Y., and Hong, Z. (2021). Mapping XML to RDF: An algorithm based on element classification and aggregation. *Journal of Physics Conference Series*, 1848(1). Retrieved at: <https://doi.org/10.1088/1742-6596/1848/1/012012> (accessed on 9. 1. 2023)
- Lytras, M.D., Hassan, S.-U. and Aljohani, N.R. (2019). Linked open data of bibliometric networks: analytics research for personalized library services. *Library Hi Tech*, 37(1), 2-7. Retrieved at: <https://doi.org/10.1108/LHT-03-2019-277> (accessed on 10. 1. 2023)
- Meghini, C., Bartalesi, V., Metilli, D. and Benedetti, F. (2019). Introducing narratives in Europeana: A case study. *International Journal of Applied Mathematics and Computer Science*, 29(1) 7-16. Retrieved at: <https://doi.org/10.2478/amcs-2019-0001> (accessed on 10. 1. 2023)
- Melo, D., Rodrigues, I., and Varagnolo, D. (2023). A Strategy for Archives Metadata Representation on CIDOC-CRM and Knowledge Discovery. *Semantic Web*, 14(3), 553-584. Retrieved at: <https://doi.org/10.3233/SW-222798> (accessed on 12. 12. 2023)
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021; 372 :n71. Retrieved at: <https://doi.org/10.1136/bmj.n71> (accessed on: 11. 12. 2023)

- Raimond, Y., Smethurst, M., McParland, A. and Lowis, C. (2013). *Using the Past to Explain the Present: Interlinking Current Affairs with Archives via the Semantic Web*. In: Alani, H., et al. The Semantic Web – ISWC 2013. ISWC 2013. Lecture Notes in Computer Science, vol 8219. Springer, Berlin, Heidelberg. Retrieved at: https://doi.org/10.1007/978-3-642-41338-4_10 (accessed on 11. 1. 2023)
- Ranjgar, B., Sadeghi-Niaraki, A., Shakeri, M. et al. (2022). An ontological data model for points of interest (POI) in a cultural heritage site. *Heritage Science* 10 (1), 1-22. Retrieved at: <https://doi.org/10.1186/s40494-021-00635-9> (accessed on 11. 1. 2023)
- Oxford Reference. (s. d.). Oxford: Oxford University Press. Retrieved at: <https://www.oxfordreference.com/search?q=ontology> (accessed on 20. 11. 2023)
- Song, D. (2014). *Towards a linked semantic web: Precisely, comprehensively and scalably linking heterogeneous data in the semantic web*. Bethlehem: Lehigh University. Retrieved at: <https://www.proquest.com/dissertations-theses/towards-linked-semantic-web-precisely/docview/1494142192/se-2> (accessed on 11. 1. 2023)
- Summers, E. and Salo, D. (2013). Linking Things on the Web: A Pragmatic Examination of Linked Data for Libraries, Archives and Museums. *arXiv*. Retrieved at: <https://doi.org/10.48550/arXiv.1302.4591> (accessed on 11. 1. 2023)
- W3.C (2008). *Extensible Markup Language (XML) 1.0 (Fifth Edition)*. Retrieved at: <https://www.w3.org/TR/xml/> (17. 1. 2023)
- Wang, Y., and Yang, S. Q. (2018). Linked data technologies and what libraries have accomplished so far. *International Journal of Librarianship*, 3(1), 3-20. Retrieved at: <https://doi.org/10.23974/ijol.2018.vol3.1.62> (accessed on 11. 1. 2023)
- Wijesundara, C., and Sugimoto, S. (2018). Metadata model for organizing digital archives of tangible and intangible cultural heritage, and linking cultural heritage information in digital space. *LIBRES*, 28(2), 58-80. Retrieved at: <https://doi.org/10.32655/LIBRES.2018.2.2> (accessed on 11. 1. 2023)

Ziku, M. (2020). Digital Cultural Heritage and Linked Data: Semantically-informed conceptualisations and practices with a focus on intangible cultural heritage. *LIBER Quarterly: The Journal of the Association of European Research Libraries*, 30(1), 1–16. Retrieved at: <https://doi.org/10.18352/lq.10315> (accessed on 11. 1. 2023)

TPOLOGY: 1.01 Original scientific research