Delta University Scientific Journal
Journal home page: https://dusj.journals.ekb.eg

Intelligent Document Archiving Using Natural Language Processing Technology

Yehia A. El_ Mashad[1], Hesham A. Ali[1], Rawda Fathy[1].

[1]*Faculty of Artificial Intelligence, Delta University for Science and Technology, Gamasa 35712, Egypt*

Correspondence:[Rawda Fathy] [Faculty of Artificial Intelligence, Delta University for Science and Technology, Gamasa 35712, Egypt]; Tel [+201090827749]; Email : AI4221054@deltauniv.edu.eg

ABSTRACT

The digital transformation era is the source of inspiration for most institutions and companies to rethink how to align their document circulation mechanism with their business objectives. Therefore, document management has become one of the most important areas in which technology works to change the mechanisms of document circulation quickly and accurately by giving institutions the ability to automate their document management processes and reduce costs. The main goal of this paper is to propose a new methodology for an ontology-based archiving system. The proposed methodology can be used to design artificial intelligence systems that can analyze and understand document content using natural language processing. In the proposed methodology, an ontology mapping engine is used as pivotal in converting raw data into a semantically rich format. By mapping raw data elements to pre-defined ontological entities, it enriches the data set with context and meaning, making it suitable for advanced processing and inference at later stages of the system. The overall architecture and its components are presented and discussed, followed by an illustrative example demonstrating the accuracy and integrity of the proposal.

Keywords: *(NLP, Document archiving, smart systems, ontology mapping, tokenization)*

1. Introduction

Businesses in today's cutthroat environment must manage massive amounts of data that must be saved and made available to users so they can make decisions. Nearly one trillion DM (Documents of Message) are produced daily worldwide, many of which may be of interest to people or enterprises. These DMs (Documents of Message) contain a wide variety of multimedia documents, movies, news, and reports of every conceivable kind and from diverse sources as depicted by Gamido, M.V.

E-document management systems, or EDMSs, are traditional information systems with hierarchical hierarchies. They demand that papers be previously categorized using one's standards. The easier it is to categorize and provide the user with better search results for certain information, the more criteria there are. Nonetheless, contemplation of the actual importance of the applicable criteria is necessary as stated by Khare, S and Akella, Janaki. The company's age and structure have an impact on this reflection. Furthermore, the process of manually classifying documents grows expensive and time-consuming as the quantity of documents grows.

Document management and the necessity to store ever-increasing volumes of information, particularly digital information, are constants in any information culture. This is due to the difficulty in determining the legitimacy and dependability of digital documents and the ease with which they can be falsified, altered, or tampered with during transmission. Furthermore, digital papers possess distinct characteristics not seen in electronic copies, and it is frequently difficult to follow the transfer and verify paper records. Consequently, several modifications are made to paper-based business procedures, which slow down document processing and make company data less accessible. In this case, document processing governed by predefined criteria can be supported and information flow accelerated if alternative technologies could convert paper documents into digital documents as discussed by Akella, Janaki.

In recent years, as a result of the continuous development of information technology, artificial intelligence technology has also advanced, and natural language processing technology has also advanced accordingly. It is this technology that directly affects the quality of the construction of the document's features. In addition, as artificial intelligence technology has developed, applications for text extraction, content construction, and feature construction have become increasingly sophisticated. In particular, the rapid development of machine learning has pushed the application of intelligent models to a higher and broader field as discussed by Beaver, William H.

In recent years, the use of AI in enterprise document management has grown in significance. Enterprise data's top goal now is to extract and leverage important information as natural language processing and other AI technologies continue to mature. While structured metadata management and storage are prioritized in traditional enterprise document management, users of search engines must enter a set of keywords to access documents and data. Manual document searches are now ineffective and time-consuming due to the company data's rapid growth in size and dimensions as stated by Carmichael.

While using a search engine for retrieval is necessary for businesses, these businesses should move away from the structured retrieval approach and instead employ commands based on predetermined terms and properties. Since search engines have become more widespread, businesses now need to be able to quickly ascertain the kinds of data that are crucial, the attributes of reliable sources, who owns a document, and how well a document performs over time in response to changes in the general content and nature of the documents that are retrieved. Additionally, to save expenses, increase operational efficiency, and make better judgments based on documentation, businesses must adopt office automation and business platforms as stated by Xiao G.

2. Background and recent work

Classifying a document—is it an email, report, presentation, spreadsheet, or another format—is the first step in most organizing strategies. Content indexing is usually the next step in the analytical process after classification. There could be millions of words or PDF documents that companies like Bloomberg, Thomson Reuters, and Dow Jones need to index. A user can choose documents of interest by providing a series of metadata tags, for example, "show me all capital markets research that has been written on IBM from May 17, 2007, to June 4, 2007, that discusses software with a bullish view in the conclusion” by Casellas, N, and Petre, I.

2.1 Traditional Methods vs AI Solutions

As shown in Table 1, traditional methods are slow, less accurate, and lack scalability compared to AI solutions, which offer high speed, accuracy, and scalability. While traditional methods incur high operational costs and require extensive training, AI systems, though initially costly, reduce ongoing costs, offer advanced search capabilities, and enhance security with minimal user training.as described by Smith, R.N.

2.2 Benefits of AI in Document Archiving

Businesses need to effectively handle an increasing amount of papers in a digital world that is changing quickly as stated by Gamido, M.V; Khare, S.; Akella, Janaki.; Beaver, William H.; Carmichael; Xiao G.; Casellas, N .; Petre, I.; Smith, R.N.; and Munir, K. Conventional filing techniques are frequently laborious and prone to mistakes. Fortunately, automated document archiving is made possible by artificial intelligence (AI). This technology offers substantial advantages in terms of efficiency, accuracy, and accessibility by classifying and storing documents according to content, metadata, and relevance. In this article, we examine the advantages of using AI for automated document preservation. So, we concluded the most important benefit as follows: (I) Improved efficiency: AI can process large volumes of documents quickly and accurately, significantly reducing the time and resources required for manual filing. This allows employees to focus on more valuable tasks. (II) Higher accuracy: AI reduces the risk of human error when categorizing and storing documents. This leads to a more accurate and consistent filing system. (III) Better accessibility: Using AI to index and categorize documents makes it easier to find documents quickly. This improves accessibility and increases productivity. (IV) Cost savings: Automation of filing reduces costs through reduced manual labor and improved operational efficiency.

2.3 Document Archiving Techniques

Document archiving refers to the management of documents to make them available, retrievable, and sustainable in the future when the need arises. Good filing practices are, however, important in this process. Integrating both the conventional and newly developed techniques offers the best approach when it comes to document management in that it makes work easier, effective, and authoritative. Optimal storage differs from choosing the right methods in compliance with the organizational needs and document type on the one hand and considering both offline and online approaches within organization objectives on the other hand.

Aspect	Traditional Methods	AI Solutions (NLP-based)
Processing Speed	Slow, the model takes a long time in terms of central processing.	High speed and efficiency, particularly when dealing with large data sets.
Accuracy	More susceptible to human variance, classification may be tender and not distinct.	Highly accurate and consistent in terms of categorization with the added semantic analysis capabilities.
Scalability	Limited scalability due to manual processes.	Highly scalable, processes more documents without extra resources.
Cost	High operational costs due to manual labor and updates.	High initial costs, and lower ongoing costs due to automation.
Flexibility	Struggles with new document types and criteria changes.	Easily updated with new AI models for handling changes.
Search Capabilities	Limited to basic keyword searches, relies on manual indexing.	Advanced search, understands the context for accurate results.
Error Handling	Errors corrected manually, are time-consuming.	AI reduces errors through machine learning and learns from data.
User Training and Expertise	Requires high levels of training and expertise.	Minimal training needed, user-friendly design.
Security and Compliance	Relies on manual audits, a higher risk of compliance breaches.	Automated compliance, and advanced security measures.

Table 1: comparative analysis of traditional methods VS AI solutions in document archiving

2.4 Fundamentals of Natural Language Processing (NLP) with Ontology Integration

NLP is a sub-discipline of AI that aims at getting machines to interact with natural language by comprehending and processing it. It is important to mention that the use of ontologies as a component of NLP improves the performance and capability of making sense of text since it provides a structured representation of relationships and concepts belonging to a specific domain. Ontologies are key in the enhancement of the performance of NLP applications in situations requiring document storage, classification, and control.

2.4.1. Key Techniques

Ontology-based tokenization enhances the potential of improving on token boundaries with the aid of the domain ontology. Ontology-enhanced POS Tagging adds ontological data to the model for better and more focused specific field tagging. The applications of the Ontology-Driven NER are useful in specializing domains because of the usage of categories and relations of a specific context. Ontology-based parsing is the procedure of ontological frameworks that improve syntactical analyses. Semantic Role Labeling with Ontologies describes the specific role information in a sentence. Ontology-informed method called, Ontology-Infused Sentiment Analysis uses ontologies of the specific domain to make a more effective assessment of sentiments. Ontology-augmented machine Translation makes it possible to obtain contextually meaningful translations due to ontological backing. The approach of Ontology-Guided Text Summarization assists in selecting the appropriate concepts or important components of a text document by using ontologies. Last, Ontology-Enhanced Information Retrieval increases the variants of queries by, relevance scoring, and accuracy in results.

3. Related work

An artificial intelligence-based IDMS suggested as stated by Başıbüyük M, Ergüzen A for processing administrative and tax records, there are issues like data quality and integration where the employment of this system can strengthen classification and data extraction. The work as stated by Di Marzo Serugendo G discusses applications of AI in the Records and Archival Management Systems (RAMS) in Nigerian public organizations focusing on the prospects of incorporating classification and retrieval enhancement with the consideration of obstacles such as infrastructure requirements and personnel development. The IDMS design introduced by Omigie C focuses on the utilization of artificial intelligence to improve the ability to process, store, and secure documents.

The benefits of the system include greater efficiency and user satisfaction and the drawbacks consisting of privacy issues and the need for technological enhancement.

As stated by Pandey M is about the application of AI in archives where advantages such as work optimization, and better search can be mentioned but where there is also advice on ethical concerns and examples of historical context loss. The role of Document Management System (DMS) in digital transformation in organizations is discussed by Jordan S, its key findings include the role of user support, system compatibility, and upgrade of the DMS to respond to emerging trends.

The ontology-based record management system as discussed by Samsudin A can augment decision support since it makes data intake and interpretation more efficient, which results in improved organizational throughput. The enhancement of document classification through the YOUFILE system that employs ontology-based smart indexing. It is for this reason that the system's architecture allows for sophisticated searches as well as enabling the indexing of documents since a list of such concepts as stated by Mainetti L. The proposed ontology introduced by Kyriaki-Manessi D is going to enhance relevance and accessibility to resources in the fields of education and history, this paper aims to contribute to the ontology of university archives particularly in technological education.

We can apply ontologies to increase the quality of the description and the conformity and interoperability of the metadata. An example of how this approach may be applied is illustrated by a practical case of archival management as discussed by Zou Q. The recommendation to use ontologies for both digital and paper archives and how a system based on an ontology can integrate the two mediums thus enhancing compatibility and searchability as stated by Barchetti, U and Llanes-Padrón D. the best way to have semantic interoperability in archives is by adopting semantic technologies and ontologies as discussed by Long, S. The study offers insights into the nature of context in describing archives and supports the development of connectivity through semantic means as described by Cheng, Z. As depicted below in Table 2, provides an overview of key papers focused on ontology-based systems in the document and archival management.

Paper	Focus	Strengths	Weaknesses	Methodology	Key Features
Mainetti, L., et al., (2008)	Ontology-based smart document indexing	Improves document search relevancy through semantic modeling.	Requires significant effort from knowledge engineers to manage large sets of axioms.	System design and implementation with case evaluations	Focuses on ontology evolution to capture new document types
Kyriaki-Manessi, D., & Dendrinos, M., (2014)	Ontology-based university archives management	Designed to enrich university archives, particularly in technological education	It may not be generalizable to other archival contexts.	Domain-specific ontology design with user perspective assessments.	Examines applicability to other educational domains and compatibility with existing archival systems.

Zou, Q. ,(2019)	Ontological approach to archival descriptions	Normalizes archival descriptions to improve interconnectivity and uniformity.	Difficulties in transitioning from traditional classification methods to ontology-based systems.	Theoretical analysis and case study of metadata standardization	. Addresses challenges in persuading staff to adopt new archival management techniques.
Barchetti, U., et al. ,(2009)	Ontology deployment in digital and paper-based records collaboration	Demonstrates how ontologies can integrate different archival formats.	Results are case-specific and may not apply broadly.	Case study approach with system implementation and evaluation	Highlights complexities in transitioning from traditional to ontology-based systems

Table 2: Comparative Analysis of Recent Research in Document Archiving Systems Using Ontology

3.1 Problem formulation and plan of Solution

Among the essential back-office tasks is traditional document processing, especially document archiving. With the assistance of AI, companies can process archiving of documents, in different forms like emails, letters, and reports in large volumes with minimal human assistance or interpretation. The main problem of the ideal AI-enabled document archiving solution when integrated into your existing processing workflow is how to achieve the following goals (i) process structured and unstructured (free text), (ii) developing an easy-to-use with no coding or technical skill (iii) handle various document types. (iv) achieve the highest accuracy and quality in data extraction. (v) be easy to integrate with the third-party systems using API/RPA**.**

4. Material and methods

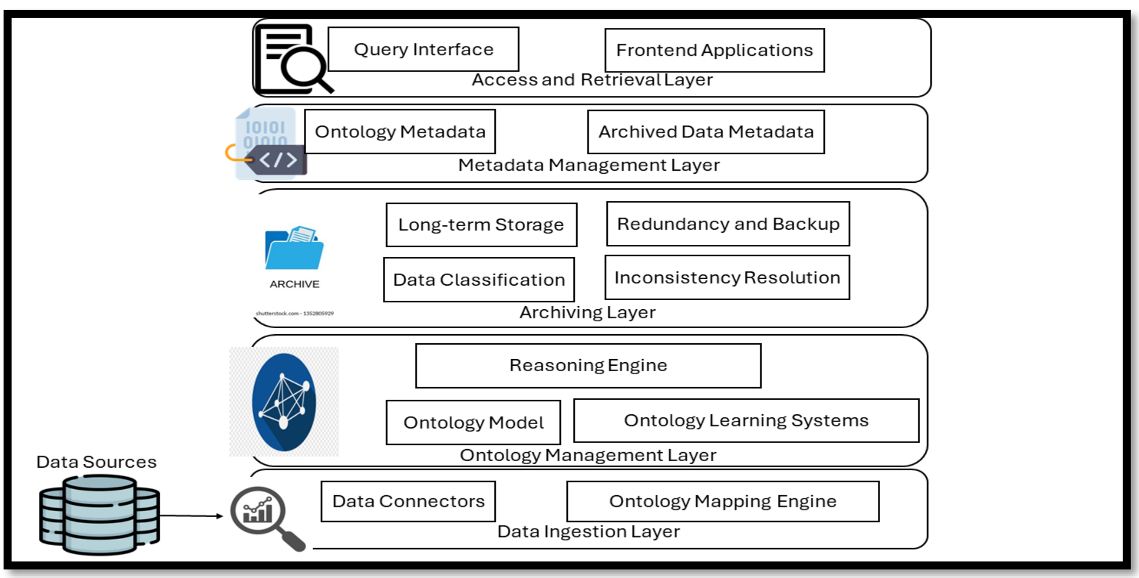

Figure 1: Ontology-Based Archiving System Architecture.

The paper outlines a detailed process for document archiving, using Natural Language Processing (NLP) techniques like Text Extraction and Text Classification. The proposed approach in Figure 1 is divided into five layers, each responsible for specific tasks to ensure proper control, storage, and retrieval of documents:

Data Ingestion Layer: This layer handles the initial loading of raw data into the system. It includes:(i) Data Connectors: These are specialized interfaces designed to pull in data from diverse sources, including databases, file systems, and web services. The connectors ensure that the data is seamlessly integrated into the system, regardless of the format or origin, allowing for smooth transitions between different data environments. (ii) Ontology Mapping Engine: This engine is pivotal in transforming raw data into a semantically enriched format. By mapping raw data elements to predefined ontology entities, it enriches the dataset with context and meaning, making it suitable for advanced processing and reasoning in later stages of the system.

Ontology Management Layer: Responsible for data structuring and semantics. Components include: (i) Ontology Model: Defines the structure of data, organizing it into classes, relationships, and properties. The Ontology Model is the cornerstone of the system, organizing data into a structured framework that enhances both the understanding and retrieval of archived documents. This step involves several critical sub-processes:

1. *Ontology Design and Conceptualization*

Figure 2 depicts the proposed Ontology methodology. The first step in building an ontology is to clearly define the purpose it will serve. This includes identifying the domain of interest, the scope of the ontology, and the types of questions it should help answer. In the context of document archiving, ontology focuses on classifying documents by type, subject, and relevance as in (1)

$$D= \{T, S, R\} \tag{1}$$

, Second Identify the core concepts that are central to the domain. These include entities such as "Document," "Author," "Date of Creation," and "Classification." As in (2)

$$C=\{Document, Author, Date\_of\_Creation, Classification\} \tag{2}$$

Each concept must be carefully defined to avoid overlap and ensure clarity, Third Establish a hierarchical structure for the concepts, starting with the most general categories and breaking them down into more specific subcategories, a hierarchical structure H as a directed acyclic graph G=(V, E), where V represents the set of concepts and E represents the edges indicating hierarchical relationships as in (3).

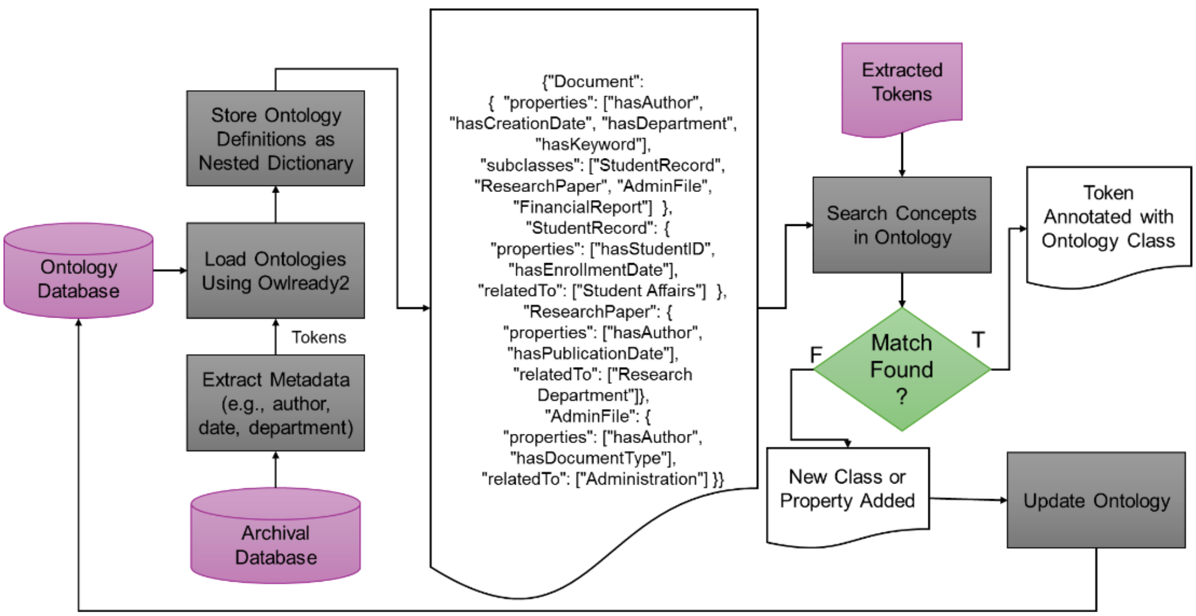

Figure 2: Proposed Ontology methodology

$$H(G)=(V,E) \text{where } V \subseteq C \tag{3}$$

This hierarchy helps in organizing the data logically and facilitates efficient data retrieval.

2. *Ontology Formalization*

- Class Definition: Translate the identified concepts into formal classes within the ontology. Each class represents a set of entities that share common characteristics. For example, a "Document" class may have subclasses like "Research Paper," "Legal Document," and "Financial Report.", classes CL as subsets of the concept set C as in (4),and(5):

$$CL=\{CL1,CL2,...,CLn\} \tag{4}$$

where $CLi \subseteq C$

$$CLDocument=\{Research\_Paper, Legal\_Document, Financial\_Report\} \tag{5}$$

- Property Specification: Define the properties that describe the relationships between different classes. These properties can be attributes (e.g., "title," "author") or relationships (e.g., "isPartOf," "references"). Properties help in defining the characteristics of each class and how they interact with one another.
  - a set of properties P that maps the relationships between classes as in (6):
    $P: CL \times CL \rightarrow Relationships$ (6)
  - Example: P(Document, Author)=has Author

- Constraints and Rules: Implement constraints and logical rules to govern the behavior of the ontology. This includes defining cardinality (how many instances of a class can be related), value restrictions (e.g., a document's creation date must be in the past), and other logical assertions that ensure data integrity.
  - Constraints S can be defined as a set of logical rules that apply to the properties and classes as in (7):
    $S=\{s1,s2,\ldots,sp\}$ (7)
  - Example of a rule: Cardinality constraint s1:

For all x∈Document, there exists exactly one y∈Author:P(x,y)=hasAuthor
Value restriction as in (8):
For all x∈Document, the Date_of_Creation must be <current date (8).

(ii) Ontology Learning Systems: Continuously refine and expand the ontology by analyzing data patterns and user interactions, integrating new concepts and relationships. They ensure the ontology evolves with the data, maintaining consistency and preventing conflicts, (iii) Reasoning Engine: Applies logical rules to structured data, deducing new information, classifying data, and maintaining consistency. Enhances data organization with inference mechanisms and supports complex queries for deeper insights.

Archiving Layer: (i) Long-term Storage: Manages secure and reliable storage for large data volumes.
(ii) Redundancy and Backup: Ensures data safety with multiple copies in different locations,
(iii) Data Classification: Uses NLP to automatically categorize documents, improving organization and retrieval efficiency, (iv) Inconsistency Resolution: Detects and resolves inconsistencies in archived data, ensuring accuracy and reliability.

Metadata Management Layer: (i) Ontology Metadata: Manages information about the ontology's structure, versioning, and elements, aiding its evolution and proper data mapping, (ii) Archived Data Metadata: Stores attributes like classification, creation date, source, and author, facilitating efficient document retrieval.

Access and Retrieval Layer: (i) Query Interface: Allows users to search using natural language or keywords, offering intuitive and efficient retrieval, (ii) Frontend Applications: Provides user-friendly tools for accessing, navigating, and managing documents.

Results

The proposed methodology offers a reliable and efficient system for storing, categorizing, and retrieving documents, using NLP to handle large volumes of data while maintaining data integrity and providing easy access for users.

5.1 Performance Metrics:

- Reliability
  1. Redundancy and Backup: This means that the system is designed in a way that will back up data to more than one location.
- 2. Inconsistency Resolution: It has features for the identification of contradictions as well as for their removal, which ensures the quality of the stored information. This is important especially to ensure that the users can be assured of the integrity of the data they are getting from the system.
  3. System Uptime: Availability is a major aspect of architecture where constant monitoring and maintenance are done to guarantee constant use of the system.
- Efficiency
  1. Data Classification: The automatic data classification through NLP helps the system to classify a large volume of documents in a very short time. This automation cuts down the amount of time and energy that would be needed in cases where the work of the system involves the manual sorting of data and records.
  2. Storage Efficiency: The system also ensures that there is proper use of the space in storage by organizing the data properly and minimizing wastage. This is especially true when handling big data as it puts into use the optimum usage of resources we have.

- 3. Retrieval Speed: In essence, the query interface is created to deliver fast and appropriate responses to a user's question. It enables the user to find the information that he or she is interested in as quickly as possible, thus improving the usability of the site.

Discussion

6.1 Illustrative Example:

In a university document repository, an ontology-based system is used to classify and organize various types of documents such as administrative decisions, government correspondence, and financial reports. The system employs an ontology to define document categories like "ResearchPaper" and their relationships with properties such as author, creation date, and department.

Step-by-Step Process:

1. Ontology Definition: The system uses an ontology to define document types (e.g., "ResearchPaper") and their associated properties like hasAuthor, hasPublicationDate, and related (e.g., ResearchDepartment).
2. Loading Ontology with Owlready2: The ontology is loaded into the system using Owlready2, creating a structured dictionary. For instance, "ResearchPaper" is associated with relevant properties and relations.
3. Metadata Extraction: Metadata such as the author (Rawda Fathy), creation date (July 27, 2024), and department (Artificial Intelligence) are extracted from the document.
4. Token Generation: These metadata fields are converted into tokens, for example, "Rawda Fathy" for the author, "July 27, 2024" for the date, and "Artificial Intelligence" for the department.
5. Ontology Matching: The system checks these tokens against the ontology. The author is linked to the author's property, the date to hasPublicationDate, and the department related in the "ResearchPaper" class.
6. Checking for Matches: Since all tokens are found in the existing ontology, the document is classified as a "ResearchPaper."
7. Ontology Update: The ontology is updated if a new document type is encountered. In this case, no update is necessary as "ResearchPaper" already exists.
8. Annotation: The document is annotated with the ontology properties: hasAuthor (Rawda Fathy), hasPublicationDate (July 27, 2024), and related (Artificial Intelligence department).

The document is now categorized as a "ResearchPaper," enabling efficient retrieval through queries like finding papers by a specific author or documents from a department. This approach enhances document organization, classification, and searchability using semantic methods.

Conclusion

The suggested approach to intelligent document archiving based on NLP and AI can increase the potential of innovative centralized document management (DOCM) systems. As a result, the use of this approach to classify and retrieve documents is far improved compared to traditional methods due to the increase in the volume of data. Thus, the usage of ontology-based systems increases the semantic level of understanding, which results in more pertinent results. They concluded that as more businesses continually produce enormous amounts of data, there will be a need for efficient document storage solutions, and with the continuous development of AI, machine learning, and ontology management, there will be a constant improvement in this area to the advantage of organizations' information asset management.

Disclosure

The authors report no conflicts of interest in this work.

References

Akella, Janaki. (2016), The value promotion of archives management in the era of big data, Journal of Finance, 23(4), pp.589-609.

Barchetti U, Bucciero A, Guido AL, Mainetti L, Paiano R. How can Ontologies Support Digital and Paper Archives? A Case Study. International Journal of Computer Information Systems and Industrial Management Applications. 2009 Jul 1;1:13-.

Başıbüyük M, Ergüzen A. Electronic document management system for Kırıkkale University. Unified Journal of Computer Science Research. 2015;1(2):8-15.

Beaver, William H. (2015), Challenges and opportunities of archives management in data Era, (4),ournal of Documentation pp.71-79.

Carmichael. (2011), Construction and sharing mechanism of enterprise archives information resources under the background of big data, ournal of Documentation, No.2, pp.6-8.

Casellas N. Legal ontology engineering: Methodologies, modelling trends, and the ontology of professional judicial knowledge. Springer Science & Business Media; 2011 Aug 12.

Cheng, Z.; Xu, Y.; Bai, F.; Niu, Y.; Pu, S.; Zhou, S. AON: Towards Arbitrarily-Oriented Text Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 5571–5579.

Di Marzo Serugendo G, Cappelli MA, Falquet G, Métral C, Wade A, Ghadfi S, Cutting-Decelle AF, Caselli A, Cutting G. Streamlining Tax and Administrative Document Management with AI-Powered Intelligent Document Management System. Information. 2024 Aug 2;15(8):461.

Gamido, M.V.; Gamido, H.V.; Macaspac, D.J.P. Electronic document management system for local area network-based organizations. Indones. J. Electr. Eng. Comput. Sci. 2023, 32, 1154–1163.

Jordan S, Zabukovšek SS, Klančnik IŠ. Document Management system–a way to digital transformation. Naše gospodarstvo/Our economy. 2022 Jun 15;68(2):43-54.

Khare, S. Gartner BPM Summit 2014: Digitalization of Business Processes; Micropact: Herndon, VA, USA, 2014.

Kyriaki-Manessi D, Dendrinos M. Developing ontology for the university archives: the domain of technological education. Procedia-Social and Behavioral Sciences. 2014 Aug 25;147:349-59.

Llanes-Padrón D, Pastor-Sánchez JA. Records in contexts: the road of archives to semantic interoperability. Program. 2017 Nov 7;51(4):387-405.

Long, S.; He, X.; Yao, C. Scene Text Detection and Recognition: The Deep Learning Era. Int. J. Comput. Vis. 2021, 129, 161–184.

Mainetti L, Paiano R, Bucciero A, GUIDO AL, SANTO SABATO S, CAPONE L, Barchetti U, PULIMENO E. YOUFILE: Ontology based system for document smart indexing. InThe 12th World Multi-Conference on Systemics, Cybernetics and Informatics: WMSCI 2008 2008 (Vol. 7, pp. 240-245). IIIS-International Institute of Informatics and Systemics.

Munir, K.; Anjum, M.S. The Use of Ontologies for Effective Knowledge Modelling and Information Retrieval. Appl. Comput. Inform. 2018, 14, 116–126.

Omigie C, Krubu D, Solomon A. Exploring artificial intelligence for records and archival management system (ai-rams) for Nigerian public organizations. Artificial intelligence. 2023 Sep.

Pandey M, Arora M, Arora S, Goyal C, Gera VK, Yadav H. AI-based Integrated Approach for the Development of Intelligent Document Management System (IDMS). Procedia Computer Science. 2023 Jan 1;230:725-36.

Petre, I.; Boncea, R.; Radulescu, C.Z.; Zamfiroiu, A.; Sandu, I. A Time-Series Database Analysis Based on a Multi-attribute Maturity Model. Stud. Inform. Control 2019, 28, 177–18.

Samsudin A, Miah S, McGrath G. An ontology-based record management systems approach for enhancing decision support.

Smith, R.N.; Aleksic, J.; Butano, D.; Carr, A.; Contrino, S.; Hu, F.; Lyne, M.; Lyne, R.; Kalderimis, A.; Rutherford, K.; et al. InterMine: A flexible data warehouse system for the integration and analysis of heterogeneous biological data. Bioinformatics 2012, 28, 3163–316.

Xiao G., Calvanese D., Kontchakov R., Lembo D., Poggi A., Rosati R., Zakharyaschev M. Ontology-based data access: A survey Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI, ijcai.org (2018), pp. 5511-5519.

Zou Q. The representation of archival descriptions: An ontological approach. McGill University (Canada); 2019.