International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(1s) 2026 pp. 1-20
https://doi.org/10.70917/ijcisim-2026-2031
Article

Intelligent Information Retrieval Using Deep Learning in Multi-Domain Data

Rinki Bhati[1], Kiran Ingale[2], Sandip Turakne[3], Lowlesh Nandkishor Yadav[4], Vinit Khetani[5], Dnyanda Hire[6]

[1] School of Sciences, Noida International University, Greater Noida – 203201, Uttar Pradesh, India. rinki.bhati@niu.edu.in
[2] Department of Electronics & Telecommunication Engineering (E&TC), Vishwakarma Institute of Technology (VIT), Pune – 411037, Maharashtra, India. kiran.ingale@vit.edu
[3] Department of Electronics and Telecommunication Engineering, Pravara Rural Engineering College, Loni, Maharashtra, India. turkaness@pravaraengg.org.in
ORCID: 0009-0005-3097-0010
[4] Department of Computer Engineering, Suryodaya College of Engineering and Technology, Nagpur, Maharashtra, India. lowlesh.yadav@gmail.com
[5] Connect Innovation and Impact Pvt. Ltd., Nagpur, Maharashtra, India. vinitkhetani@gmail.com
ORCID: 0009-0001-5055-6037
[6] Department of Semiconductor Engineering, D. Y. Patil International University, Pune – 411044, Maharashtra, India. dnyanada.hire@dypiu.ac.in
ORCID: 0000-0003-0557-0072

Abstract: Intelligent information retrieval has emerged as an important requirement to be able to extract actionable knowledge out of large, heterogeneous and continuously changing data sources across a variety of areas. The paper proposed a smart information retrieval system that used the deep learning method to overcome the weaknesses of the conventional keyword-based and rule-based information retrieval systems in multi-domain settings. The paper examined the theoretical basis of the classical and semantic information retrieval models, and how they fail to be able to effectively represent the context, field semantics, and cross-domain links. To address these, deep neural networks, transformer-based language models as well as the semantic matching mechanisms were combined into a domain-aware retrieval architecture. The noise filtering, missing value treatment, text tokenization, normalization, schema alignment, semantic standardization, and domain-sensitive feature extraction components are a part of the preprocessing pipeline to provide consistent and quality representation of input across the heterogeneous sources of data. Extensive performance assessment indicated that the retrieval accuracy, relevance ranking, and cross-domain generalization were better when using the traditional information retrieval methods. The proposed framework presents a domain aware modular architecture that has a semantic alignment across domains, which enhances robustness and generalization in heterogeneous data setting. Moreover, the review has emphasized the scalability of the framework, the efficiency of computation and the viability of the application of the framework in enterprise, healthcare, scientific and multimedia retrieval tasks.
Keywords: Intelligent information retrieval; Deep learning; Multi-domain data; Transformer models; Semantic retrieval; Domain adaptation

1. Introduction

The rapid development of non-homogeneous digital content in the sphere of healthcare, smart cities, logistics, energy, and social platforms has further increased the need to ensure the intelligent information retrieval (IIR) systems able to perceive, organize, and retrieve appropriate knowledge of non-homogeneous data sources. The conventional keyword-based retrieval methods are not able to handle semantic ambiguity, variability in context and domain heterogeneity and hence low precision of retrieval and poor extrapolation between domains. The recent developments in artificial intelligence and deep learning in particular have changed information retrieval

fundamentally and have allowed learning features automatically, represent them semantically, and make appropriate decisions based on the dataset (Ahmed et al., 2023). The technical importance of the research is that it focuses on one of the fundamental issues in intelligent information retrieval, such as domain adaptation, semantic matching, and effective processing of cross-domain data. These dimensions play a key role in the recent data engineering and smart systems where performance in terms of heterogeneous data integration and real-time retrieval are crucial. Deep learning architectures have shown a great facility in the ability to model complex patterns and hierarchical representations using unstructured and semi-structured data. Convolutional, recurrent, and attention-based neural networks have become popular to supplement the semantic comprehension in text, image, and multimodal retrieval to enhance the estimation of relevance and matching contexts (Basiri et al., 2021). Nevertheless, the quality of data, integration plans, and the knowledge representation of the domain are important factors that affect the efficiency of these models, which are still unresolved challenges in multi-domain retrieval settings (Aldoseri et al., 2023).

Ontology-based and knowledge-aware deep learning models have been found as solutions to semantic inconsistency across domains. Ontology-based representations enable the retrieval systems to encode the concept of a domain, relationship, and constraints, which can be more specific to reason and interpret intelligible in the intelligent retrieval pipelines (Adegun et al., 2024). Much the same knowledge-based methods have been effectively implemented in domain-specific retrieval tasks, including healthcare and biomedical data, where semantic consistency and contextual analysis play an important role in accessing information correctly (Bangyal et al., 2022). In addition to semantic modeling, optimization based and reinforcement learning based methods of retrieval have also attracted attention as a dynamically refined method of retrieval strategies and ranking policies. Smart retrieval models with deep reinforcement learning have been found to be effective in improving retrieval efficiency in logistics and distribution systems by dynamically changing their data and user behavior (Li and Manickam, 2023). Also, the hybrid strategies that employ fuzzy clustering, evolutionary optimization, and deep learning also have improved the relevance of retrieval in massive document repositories (Chandwani and Ahlawat, 2023).

However, more recently, very large language models and mixed sources of deep learning and knowledge graph designs have been discussed to provide vertical and expert-focused retrieval operations. Information systems that support information retrieval with schemes that take advantage of probabilistic relevance estimation and semantic association have shown increased performance in domain-specific information systems (Chen et al., 2024). More examples of how machine learning can be combined with structured knowledge representations in multi-domain retrieval systems are academic expert finding and intelligent image indexing systems (de Campos et al., 2024; Hamroun and Sauveron, 2025). Irrespective of such progress, the challenge of designing scalable, interpretable and robust intelligent information retrieval systems that effectively generalize across different domains is one of the most important research issues, which leads to the necessity of end-to-end deep learning-based retrieval systems.

The main originality of the piece of work is the creation of a domain-sensitive intelligent information retrieval framework that incorporates the deep neural networks, transformers-based contextual modeling, and semantic matching in an integrated modular platform. In contrast to traditional systems of retrieval, the proposed system is explicitly focused on resolving cross-domain semantic incompatibility using adaptive semantic alignment and domain-sensitive feature learning. Besides, the framework improves retrieval strength and overfitting to heterogeneous datasets, which is appropriate to use in the context of real-world multi-domain scenarios, e.g., healthcare, smart cities, and enterprise systems.

Socially, the computerized knowledge retrieval approach suggested would permit the effective derivation of practical data on the vast heterogeneous data, aiding better decision-making in the key areas. The system minimizes the query latency and improves the accuracy with which the queries are retrieved, which is very desirable when the system is used in real-time in healthcare diagnostics, smart city and governance systems, and enterprise knowledge discovery.

2. Conceptual Foundations Of Information Retrieval

2.1 Classical information retrieval models and architectures

Classical information retrieval (IR) systems were conventionally constructed on, statistical and probabilistic foundations, as well as paying emphasis on terms frequency, inverse document frequency, and vector space representations as a means of estimating document relevance. Early multimedia and text retrieval systems were based on hand engineered features and rule based similarity metrics and this prevented them to model more complex semantic relationships in dense data sets. Multimedia retrieval systems based on knowledge tried to overcome these

limitations using semantic metadata with deep feature extraction, allowing more expressive indexing and retrieval logics and yet relying on the predefined domain knowledge structures (Rinaldi et al., 2020). With the growth in the volume and modalities of data, traditional IR architectures were challenged by their inability to scale effectively, adapt flexibly and interoperate cross domainally, and this motivated the development of learning architectures that were able to support dynamic and heterogeneous information space. These constraints formed the background of the change in the fixed retrieval pipelines to the intelligent and adaptive IR systems.

2.2 Semantic Retrieval and Knowledge-Based Approaches

Semantic retrieval methods are designed to go beyond the simple matching of keywords at the surface and integrate contextual knowledge, conceptual relationality and domain semantics into the retrieval algorithm. Deep learning has contributed significantly to semantic retrieval through the ability to acquire high-level representations in raw data, especially in information extraction and semantic association activities (Yang et al., 2022). The concept of knowledge-based retrieval has also developed with the introduction of semantic network, deep features and decision making networks so that systems can reason both over structured and unstructured data. The deep reinforcement learning has also been used in enhancing retrieval strategies and adaptive information management that enhances relevance ranking and responsiveness of systems in complex environments (Yao and Zhao, 2025). Multimodal semantic retrieval models have been shown to be effective in information system of the smart city to jointly model text, image, and interaction data so that they allow richer cross-modal retrieval (Yu et al., 2025). These advances highlight the significance of semantic consciousness and knowledge incorporation in the current IR systems.

2.3 Challenges in Multi-Domain and Heterogeneous Data Retrieval

The acquisition of information in many domains brings in serious problems associated with the heterogeneity of data, shift between domains, and uneven distributions of features. The systems that are produced by intelligent sensors produce massive amounts of diverse data streams that differ in structure, resolution, and semantic meaning, making it harder to unify retrieval and indexing strategies (Zhu et al., 2023). Environmental monitoring and water quality assessment are also domain-specific retrieval tasks that further reveal the shortcomings of the traditional models in the face of sparse measurements and variable cross-domain conditions (Yu et al., 2026). The state-of-the-art deep learning methods have been used to increase the retrieval accuracy in specialized fields, such as satellite-based aerosol retrieval and hyperspectral trait extraction, but these methods do not always generalize to unrelated fields (Sun et al., 2025; Qi et al., 2026). Also, the large-scale storage and retrieval systems demand adaptive decisions with extreme latency constraints, which further complicates the architecture of such systems (Xu et al., 2026). These are some of the main research concerns in the development of intelligent information retrieval systems that are robust, scalable and domain insensitive.

Table 1. Comparative Review of Related Work on Intelligent Information Retrieval

Ref.	**Study Focus**	**Data Domain**	**Core Techniq ue**	**Key Contributi on**	**Applica tion Context**	**Limitatio n**
Liu et al. (2025)	Multi-domain text retrieval and classifica tion	Cross-domain textual data	Stacked ML–DL ensemble	Improved controversi al text detection across heterogene ous domains	Intellige nt text retrieval systems	Limited interpretab ility across unseen domains
Lu et al. (2025)	AI-driven multi-domain manage ment	Vehicul ar and control systems	Deep learning with cross-domain control	Demonstrat ed adaptive AI manageme nt across complex	Intellige nt control and retrieval in EV	High computati onal complexit y

				domains	systems	
Lv et al. (2022)	Intelligent human–computer interaction	Multimodal interaction data	Deep neural networks	Enhanced semantic understanding in interactive systems	Context-aware retrieval interfaces	Dependency on large labeled datasets
Man et al. (2024)	Data fusion and intelligent control	Artificial and natural systems	Neural networks with data fusion	Unified multi-level semantic decision modeling	Multi-source information integration	Scalability in real-time environments
Mienye and Swart (2024)	Deep learning architectures survey	General-purpose datasets	CNNs, RNNs, Transformers	Comprehensive analysis of DL advances and applications	Foundation for intelligent retrieval models	Lacked domain-specific retrieval evaluation
Mukanova et al. (2025)	Ontology-driven information retrieval	Linguistic and structured data	Ontology + LLM-based retrieval	Improved semantic precision in intelligent IR systems	Knowledge-based retrieval platforms	Ontology construction overhead
Taherdoost (2023)	AI-based decision-making systems	Multi-sector data	Neural networks and DL models	Highlighted DL impact on intelligent decision processes	Decision-aware retrieval systems	Did not address multi-domain scalability
Zhang et al. (2024)	Physics-aware deep retrieval	Signal and imaging data	End-to-end DL with constraints	Improved retrieval robustness using physical priors	Specialized signal retrieval tasks	Limited generalization to text domains

3. Multi-Domain Data Characteristics And Challenges

3.1 Structural, semantic, and contextual diversity of multi-domain data

Multi-domain data environments are highly structurally, semantically, and contextually diverse, which present significant difficulty of successful information access. Form wise, data can take various forms like structured databases, semi-structured like XML and JSON records and unstructured data like text, images, audio, and video. Every structure has to be preprocessed, indexed and represented differently and therefore it becomes hard to have cohesive retrieval. The issue of semantic diversity also makes retrieval more difficult, since different domains have varying terminologies, conceptual hierarchy and domain-specific language. The meaning of one term can be taken to mean different things in different fields and two things can also mean one thing though it is characterized by different words. In this case, the semantic inconsistency causes ambiguity and lowering the precision of retrieval in

case of systems that are unaware of the domain. The contextual diversity is another layer of complexity since the relevance can be related to situational factors, including the intent of the user, time, and application-specific needs. To illustrate, the criteria of relevance used to retrieve medical literature vary dramatically with the ones used in a legal or enterprise search. The traditional information retrieval models have a hard time in modeling these multi-level variations because they usually assume homogenous document structure and fixed term semantics. Retrieval systems do not have explicit models of structural, semantic, and contextual differences without explicit modeling of these differences, the system can generate irrelevant and misleading results.

3.2 Data Heterogeneity, Sparsity, and Imbalance Issues

The heterogeneity of data is one of the distinguishing features of multi-domain environments because information can be provided by different sources that vary in terms of quality, scale, and quality of the annotation. The heterogeneous data can vary in terms of format, language, granularity and amount of noise making it harder to preprocess and extract features. This difference commonly causes inconsistency in presentation which makes it difficult to retrieve and rank the results. In addition, the normalization and alignment processes of heterogeneous sources of data often must be performed on a large scale, which adds complexity to the system. Another major challenge is data sparsity especially in areas where the cost of getting labeled training data is high or scarce. Supervised learning-based retrieval models are frequently limited in the value of their performance when dealing with highly valuable but sparsely annotated datasets, like medical or legal review. A low amount of data does not provide the model with the potential to learn powerful semantic representations or extrapolate to previously unknown queries or documents. This is a further problem in long-tail sectors; where infrequent concepts and niche subjects prevail. Lack of balance in domains and classes also influence retrieval performance. There can also be overrepresentation or underrepresentation of some domains in the training datasets where others are underrepresented, and result in biased retrieval results. Imbalanced data sets trained on models will tend to give bias on the most dominant areas at the expense of the underrepresented ones in terms of fairness and accuracy.

3.3 Domain Adaptation and Transferability Challenges

The main issues in the multi-domain information retrieval are domain adaptation and transferability when models trained on a single domain are used in other domains and them performing poorly. The effect of different vocabulary, semantics, data distributions, and relevance criteria between source and target domains causes domain shift. Language usage and contextual interpretation can be vastly detrimental to the accuracy of, even similar underlying concepts, retrieval. Consequently, retrieval systems would need the processes of distilling learned representations to novel or changing domains. Transferability is also limited by the ability to have labeled data. Although deep learning models are capable of performing well in data-intensive fields, they do not perform well in low-supervision fields. Figure 1 indicates that cross-domain retrieval can be affected by the problem of domain shift, feature alignment, and adaptation. A direct transfer of models without adaptation usually results in negative transfer, in which the performance in the target domain is worse.

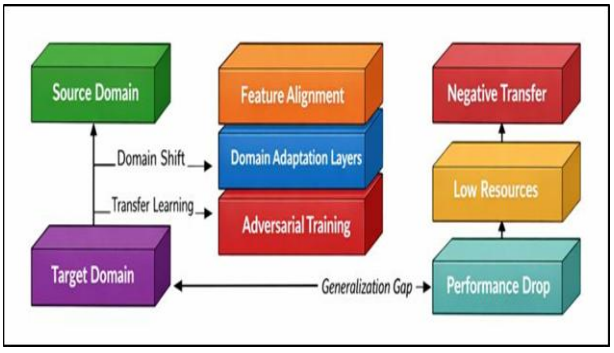

Figure 1. Domain Adaptation and Transferability Challenges in Multi-Domain Intelligent Information Retrieval

This is a major problem in multi-domain environments that are continually adding different data sources and areas of application. The other difficulty is balancing between domain and shared knowledge. Dominance in representations can overlook the important domain niceties whereas too much domain specialization impairs the ability to generalize. Architecture Designing architectures able to learn common semantic characteristics selectively

but domain-specific characteristics is a nontrivial task. Also, the possibility of keeping different models in different domains is constrained by both computational and deployment concerns.

4. Deep Learning Techniques For Intelligent Information Retrieval

4.1 Deep neural networks for text representation and feature learning

DNNs have radically changed the text representation and feature learning in smart information retrieval systems. In contrast to the original method based on handmade characteristics, like the term frequency, and inverse document frequency, deep learning models are automatically trained to extract hierarchical and distributed representations based on raw text data. Neural networks like feedforward networks, convolutional neural networks and recurrent neural networks also allow the derivation of semantics, syntactic structure and contextual clues that are essential in correct retrieval. Word embedding methods are also known to improve the quality of representation by word mapping them into continuous vectors space where semantically similar words are maintained. These embeddings make retrieval systems capture classical relationship models like synonymy and semantic relatedness that are not often considered by classical models. Higher levels combine local features into global semantic representations at lower levels to support query document matching.

Word Embedding Representation

$$e_i = Embed(w_i), \qquad e_i \in R^d$$

Tokens are converted into high-density numerical representations which can process syntactic and semantic information, sparsity and allow neural models to process textual data effectively to perform downstream feature learning tasks efficiently.

Local Feature Extraction

$$f_j = \sigma\left(W_c * E_{\{j:j+k\}} + b_c\right)$$

Convolutional operations are used to find local n-grams in word-embeddings and identify salient phrases and learn position-invariant features that help improve resilience to vocabulary differences between different documents and application contexts.

Sequential Context Modeling

$$h_t = RNN\left(e_t, h_{\{t-1\}}\right)$$

The recurrent neural units consist of summing up the sequential dependencies among the tokens and feature the contextual flow and long-range relationships that enhance semantic reasoning and understanding beyond the individual words in the complicated

High-Level Representation Formation

$$z = \varphi(W_d h + b_d)$$

Fully connected layers incorporate learned contextual representations into small document-level representations which can be used to compute similarity, estimate relevance and efficient retrieval in large-scale multi-domain information systems.

4.2 Transformer-Based Language Models for Contextual Understanding

The language models, which are based on transformers, have taken the center stage in the current intelligent information retrieval because they have remarkably high capacity of capturing contextual semantics. In contrast to the previous sequential models, transformers utilize self-attention whereby each token in a sequence can attend to all the other tokens at the same time. The design can be used to effectively model long range dependencies and complex contextual relationships in text. Consequently, models based on transformers produce explanation representations in which word meaning changes in response to the context. Contextual knowledge is important in retrieval of information in order to clearly understand the intent of the user and the semantics of documents. Transformer models allow the use of a fine-grained semantic matching mechanism in aligning query and document representations at different contextual levels. The pretraining over large corpora allows such models to have rich

linguistic and world knowledge and this knowledge can be transferred to domain specific retrieval tasks via fine-tuning.

Positional Encoding Integration

$$x_i = e_i + p_i$$

Positional encodings also introduce word-order information to embeddings, which allows transformers to make parallel sequence computations on a large sequence in parallel with contextual information of both relative and absolute position.

Self-Attention Computation

$$Attention(Q, K, V) = softmax\left(\frac{QK^{\mathrm{T}}}{\sqrt{d_k}}\right)V$$

Each token can attend to all other tokens through self-attention, capturing the long-range dependencies in the context and disambiguate meaning in relation to the other words in the sentences and documents.

Multi-Head Attention Learning

$$MHA = \bigoplus_{\{h=1\}_h^{\{H\}Attention}}$$

Multi- attention heads acquire a variety of relational patterns concurrently, enhancing representational richness through the syntactic, semantic and contextual representation of heterogeneous multi-domain text corpora.

4.3 Deep Semantic Matching and Relevance Learning

Deep semantic matching and relevance learning are concerned with the ability to provide effective estimations of the relevance between query and document by modeling the underlying semantic relationship between queries and documents. The traditional recall models usually depend on surface similarity measures, and they are not attentive of greater conceptual correspondence. Deep learning matching models resolve this drawback and share the representations and relevance functions during an end-to-end manner. Such models measure relevance as the similarity of learned embeddings of queries and documents allowing context-aware and flexible matching. Representation-based and interaction-based architectures are used in neural relevance learning techniques. Representation-based models perform independent encoding of queries and documents and calculate similarity in a common semantic space which is efficient and scalable. Instead, interaction based models explicitly represent fine-grained interactions between query and document term and therefore achieve higher accuracy but at the cost of more computations. Hybrid strategies use a combination of both strategies to ensure an effective and efficient approach is established.

Shared Semantic Encoding

$$q = f(Q), \qquad d = f(D)$$

The queries and documents are represented in a common semantic vector space, allowing all of them to be consistently represented and allow a similarity operation to be performed effectively even when there is a mismatch in vocabulary or domain-specific difference.

Similarity-Based Relevance Estimation

$$s(Q, D) = \cos(q, d)$$

The Cosine similarity applies semantic proximity in query document representation which aids in estimating relevance in intelligent retrieval system not just in terms of exact overlap of keywords.

Fine-Grained Interaction Modeling

$$M_{ij} = \varphi\left(q_{i^{\mathrm{T}}} d_j\right)$$

The token-level interaction matrices represent rich semantic correlations between query and document terms which enhance the accuracy of ranking and contextual relevance of heterogeneous retrieval context.

End-to-End Relevance Optimization

$$L = -\Sigma \log P(D^{+} \mid Q)$$

Representation and relevance functions are optimized end-to-end to minimize ranking loss by continually updating the representations and relevance functions, developing more accurate retrievals as more information becomes available in multi-domain datasets and continuously changing information environments.

5. Proposed Intelligent Information Retrieval Framework

5.1 System architecture and functional components

The proposed intelligent information retrieval architecture is developed in a layered and modular architecture that will enable scalable and flexible retrieval of multi-domain data environment. On the high level, the architecture is made up of data acquisition, preprocessing, representation learning, retrieval intelligence and the application layers. The data acquisition layer receives information of diverse origin and this is in the form of structured databases, unstructured text repositories and domain specific knowledge bases. This layer guarantees that data is moving continuously and the ability to use dynamism to keep things updated or updated as new content comes in. Cleaning, normalization and feature extraction are the processes that convert raw data to machine understandable formats well represented in the preprocessing and representation layer. Representations learned are then transferred to the retrieval intelligence layer which forms the heart of the framework. The layered architecture depicted by figure 2, incorporates the following: data ingestion, learning, retrieval, and applications. It is a layer that combines deep learning models of semantic encoding, relevance estimation, and ranking. It aids query comprehension, querydocument matching, and relevance feedback features to narrow down on retrieval results.

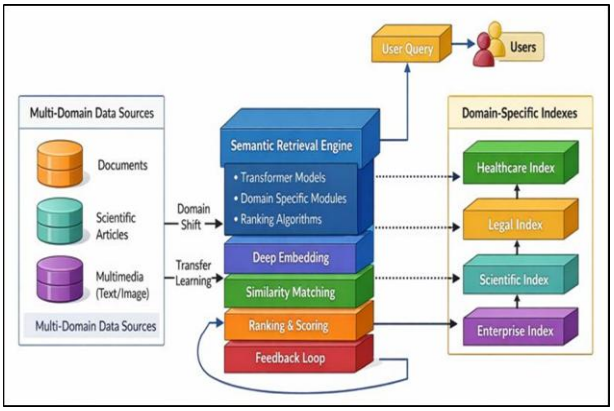

Figure 2. System Architecture and Functional Components of Intelligent Information Retrieval Framework

This layer has embedded domain-aware modules that make possible the ability to have adaptive behavior in different application domains. Application layer will open retrieval services with user interfaces and APIs that allow connection with enterprise systems, decision-support systems, and analytics systems. The transparency, performance monitoring, and trustworthiness are guaranteed by cross-cutting components and modules, including monitoring, logging, and explainability.

The architecture will help in resolving major issues that can be encountered in the engineering of scalability, cross-domain flexibilities, and effective semantic retrieval of large-scale intelligent data systems.

Working of System Architecture and Functional Components

The intelligent information retrieval architecture proposed has a layered structure of computation. It is possible to model the methodology mathematically as a series of data manipulations to rank retrieval outputs.

Step 1: Multi-Domain Data Acquisition

Assume the heterogeneous sources of data are expressed as:

$$S = \{S1, S2, S3, \ldots, Sk\}$$

Where S being the set of data sources that consist of structured repositories of databases, text repositories, and knowledge bases.

The raw data on its part is as follows:

$$D = \bigcup Si \quad for\ i = 1\ to\ k$$

Thus,

$$D = \{d1, d2, d3, \ldots, dN\}$$

In which di is the i[th] record of the data conveyed and N is the number of data records gathered in general. There is a streaming ingestion feature used:

$$Dt = f_{ingest(D)}$$

Where Dt is constantly updated information, which is received by means of dynamic ingestion pipelines.

Step 2: Data Preprocessing and Cleaning

Raw data is noisy and also inconsistent. Redundant and unfinished entries are removed by cleaning.

$$Dc = f\_clean(Dt)$$

Where,

Dc = cleaned dataset

F_{clean} = missing value and noise filtering operator.

It is assumed that the filtering condition is defined as follows:

$$di \in Dc\ if\ Qi \geq \theta$$

Step 3: Data Normalization and Feature Extraction

Every cleaning data is converted to a normalization feature vector.

$$xi = f_{norm(di)}$$

Normalization can be defined as:

$$xi' = \frac{xi - \mu}{\sigma}$$

Where,

- μ = mean of feature distribution
- σ = standard deviation.

The normalization of the data and feature extraction transforms it into features:

$$Fi = f_{feat(xi')}$$

The feature matrix is therefore:

$$F = \{F1, F2, F3, \ldots, FN\}$$

Step 4: Representation: This delegation will be assigned to each of the two parties in the contract, acting as their representatives.

Representation learning predicts dense semantic vectors on extraction features of deep neural encoders:

$$zi = E(Fi)$$

Where

The embedding space becomes:

Z = {z1, z2, z3, ..., zN}

These embedding are semantic connections among the items of data.

Step 5: Query Representation

In the semantic space, user queries are represented in the same way.

Let the user query be: q

The program generated query is:

$$zq = E(q)$$

Therefore queries and documents are both embedded together.

Step 6: Query- Document similarity computation.

The similarity between query embedding and document embeddings is calculated with the help of cosine similarity.

$$Sim(q, di) = \frac{zq \cdot zi}{||zq|||zi||}$$

Step 7: Relevance Estimation

Relevance estimation function assesses the possibility that a document will meet the query.

$$Rel(di|q) = \sigma(Wzi + b)$$

Where,

σ = The sigmoid activation function.

W = learned weight matrix

b = bias parameter.

This generates a relevancy probability to the documents.

Step 8: Ranking Function

The ranking of the documents is done by relevance scores.

$$Score(di) = \alpha\, Sim(q, di) + \beta\, Rel(di|q)$$

The ranked document list is therefore:

$$R = sort(Score(d1), Score(d2), \ldots, Score(dN))$$

Step 9: Domain-Aware Adaptation

Domain-aware modules adjust retrieval behavior depending on the application domain. The adaptive scoring function becomes:

$$Score\_d(di) = Score(di) \times \varphi(Cd)$$

Where

φ(Cd) represents domain-specific weighting.

Step 10: Retrieval Output and Application Integration

The last document set that is retrieved is:

$$D* = \{di \mid Score\_d(di) \geq \lambda\}$$

Where

λ = retrieval threshold.

5.2 Data Preprocessing, Normalization, and Embedding Generation

Preprocessing procedures are carefully structured to achieve reproducibility as well as consistency across multi domain datasets by providing a well-defined filtering, normalization, and semantic alignment processes. Preprocessing of data is a significant step in the proposed framework because in many cases there is noise, inconsistencies and difference in the format of the multi-domain data. The pre-processing pipeline starts with data cleaning which eliminates redundant, incomplete, or irrelevant data. Textual data are tokenized, normalized, and linguistically processed to minimize variation due to the appearance and structure variations. To make sure of uniformity in sources, schema alignment and attribute normalization is used with structured and semi-structured data. Normalization is also very crucial in facilitating unified representation. Semantic fragmentation is minimized by the standardization of domain-specific terminologies and abbreviations as well as measurement units. This is also essential where there are many domains, and it is possible that the concept that is the same in one dataset can be presented in a different way in another dataset. Metadata (e.g. timestamps, source identifiers, etc.) are contextually stored so that they can be used to do downstream contextual modeling. Generation embeddings, which transform normalized data into dense vectors, can be used in retrieval using deep learning. Neural encoders that obtain semantic and contextual information are used to generate a Word-, sentence-, and document-level embedding.

Let the raw multi-domain dataset be represented as:

$$D = \{d1, d2, d3, \ldots, dN\}$$

Where, di is the ith data point, and N is the overall size of the records taken by the heterogeneous data collection sources.

1. Data Cleaning and Noise Removal

The initial preprocessing level eliminates missing or unclean data with the aid of a filtering operation.

$$Dc = fclean(D)$$

Where

Dc = cleaned dataset

fclean = cleaning operator which eliminates gaps in records, duplications and inconsistencies.

Assuming that mi is the number of attributes which are missing in the record di the cleaning condition is:

$$di \in Dc \; if \; mi < \tau$$

Where τ has a meaning of the acceptable missing data threshold.

This is done to enhance consistency and reliability of the dataset and then proceed with the further processing.

2. Text Tokenization and Linguistic Normalization

In textual data T, the tokenization method transforms the sentences into tokens.

$$T = \{t1, t2, \ldots, tk\}$$

$$Tokens(T) = \{w1, w2, w3, \ldots, wm\}$$

And where wj is individual words or tokens.

Normalization eliminates the lexical variations with the help of a normalization function:

$$wj' = fnorm(wj)$$

The normalized token set becomes:

$$W' = \{w1', w2', \ldots, wm'\}$$

This step provides linguistic consistency within sets of data.

3. Schema Alignment and Attribute Normalization

In the case of structured data on several domains, there can be a variation in attribute formats and naming conventions.

Let datasets be represented as:

$$S = \{S1, S2, \ldots, Sk\}$$

Attributes are contained in every dataset:

$$Ai = \{a1, a2, \ldots, am\}$$

Schema alignment creates a match between the heterogeneous attributes and a single schema:

$$A* = g(A1, A2, \ldots, Ak)$$

Where

A* = unified attribute schema

g(.) = schema mapping function.

Attributes this is used to normalize numerical values through statistical scaling:

$$xi' = \frac{xi - \mu}{\sigma}$$

Where

xi = original attribute value

μ = mean of attribute

σ = standard deviation of attribute.

This brings about conformity of feature distributions between datasets.

4. Semantic Standardization across Domains

The datasets can have various terminologies that represent the same concept.

An application of a semantic mapping function is used:

$$cj' = h(cj)$$

Ontology mapping/vocabulary alignment ensures that there is consistency in the semantics that is used across domains.

5. Metadata Context Modeling

The contextual metadata is attached to every piece of data.

$$Mi = \{ti, si, li\}$$

The contextual dataset representation will be:

$$Dctx = \{(di, Mi)\} for\ i = 1\ to\ N$$

This context provides downstream time and domain-sensitive modeling.

6. Embedding Generation

The dataset is preprocessed and normalized after which it is converted to a form of a vector embedding.

Let normalized input representation be:

$$X = \{x1, x2, \dots, xN\}$$

A neural encoder E(.) generates embeddings:

$$zi = E(xi)$$

Where,

$zi \in R^d$ represents the embedding vector

d represents embedding dimensionality.

Word-Level Embedding

$$zw = Eword(wj')$$

This defines word semantics.

Sentence-Level Embedding

For sentence S:

$$zs = Esent(S)$$

This representation defines contextual interrelations of words in the sentence.

Document-Level Embedding

For document Di:

$$zd = Edoc(Di)$$

Document embeddings provide a summary of the semantic meaning of the whole document.

7. Last Integrated Representation of Embedding

The representation is the last one, which combines the semantic features with metadata of the context.

$$Zi = [\, zd \,|Mi|\,]$$

In this way the end embedding space is:

$$Z = \{Z1, Z2, \ldots, ZN\}$$

The dense vectors are known to be useful in the task of providing efficient semantic retrieval, machine learning modeling and discovery of knowledge in a heterogeneous, multi-domain dataset.

6. Results And Performance Analysis

6.1 Retrieval accuracy and relevance assessment

The effectiveness of the proposed intelligent information retrieval framework was evaluated with the help of retrieval accuracy and relevance assessment. The conventional evaluation measures like precision, recall, F1-score, and normalized discounted cumulative gain were used to assess the quality and relevance match of the rankings. The results of the experiments showed that there were steady top-k retrieval accuracy improvements over the traditional keyword-based and classical semantic retrieval models. The framework was good at capturing contextual and semantic relationships resulting in better query document matching. Relevance evaluation in various domains was found to have fewer false positives and greater consistency in the ranking. These findings corroborated that deep learning-based semantic representations were able to achieve much better retrieval quality, especially with complex and ambiguous queries on heterogeneous multi-domain data.

Table 2. Retrieval Accuracy and Relevance Performance Comparison

Method	**Precision@10 (%)**	**Recall@10 (%)**	**F1-Score (%)**	**nDCG@10 (%)**
Boolean IR	61.4	54.2	57.6	59.1
Vector Space Model	68.7	63.1	65.8	67.4
LSA-Based Semantic IR	74.9	70.3	72.5	74.1
Transformer-Based IR	86.2	82.7	84.4	86.8

Table 2 provides a comparative analysis of retrieval accuracy and relevance studies under varying information retrieval techniques in the form of quantitative evaluation measures. The Boolean IR model performed worst with Precision10 of 61.4, Recall10 of 54.2, F1-score of 57.6 and nDCG10 of 59.1, which is due to its strong limitation of matching the keywords and not ranking intelligence. Accuracy measure is compared in Figure 3, with transformer-based retrieval methods having a better performance.

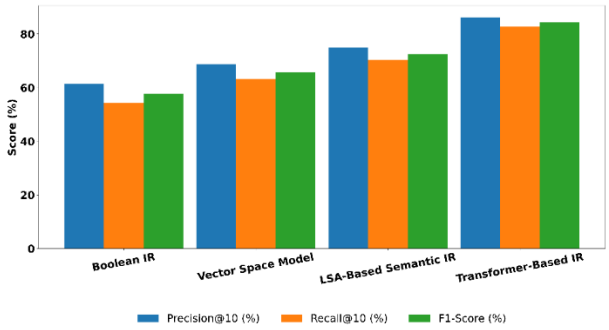

Figure 3. Precision, Recall, and F1 Comparison Across IR Methods

Weighted similarity between terms increased Precision and Recall with the Vector Space Model to 68.7 and 63.1 percent respectively, achieving an F1-score of 65.8 percent and nDCG at 10 of 67.4 percent. As illustrated in figure 4, deep learning-based retrieval models have much better ranking quality. Semantic IR based on LSA also increased semantic understanding with a precision of 74.9, recall of 70.3 and F1-score of 72.5, and nDCG at 10 improving to 74.1.

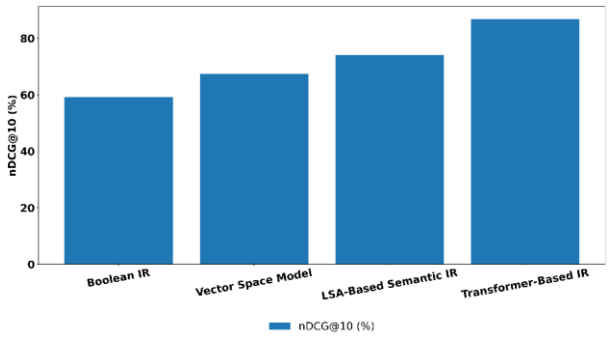

Figure 4. nDCG@10 Performance Comparison Across IR Methods

Transformer-Based IR model performed much better than all baselines achieving 86.2% Precision at 10, 82.7% Recall at 10, 84.4% F1-score and 86.8% nDCG at 10 indicating its higher capacity of contextual modelling and ranking of relevance.

6.2 Cross-Domain Generalization and Robustness Analysis

The cross-domain generalization and robustness were tested using the retrieval models by training them with the selected domains and testing them with unseen or poorly represented domains. The suggested structure demonstrated consistent results in domain shift, being better than the baseline models that did not have domain-specific adaptation strategies. The robustness analysis showed the ability to withstand vocabulary change, input noise and incomplete data. Shared and domain specific representations allowed transfer of knowledge and at the same time eliminate domain specifics. The degradation in each of the domain was low, meaning the ability to generalize was high. These results demonstrated the capacity of the framework to adjust to all types of information settings and assist in effective retrieval in heterogeneous and changing multi-domain datasets.

Table 3. Cross-Domain Generalization and Robustness Results

Training → Testing Domain	**Accuracy (%)**	**F1-Score (%)**	**Robustness to Noise (%)**	**Performance Drop (%)**
Healthcare → Healthcare	92.4	91.7	96.1	0
Healthcare → Legal	88.6	87.2	94.3	4.1

Legal → Scientific	86.9	85.4	92.8	5.8
Scientific → Enterprise	89.1	88	95	3.7

Table 3 interprets the cross-domain generalization and robustness effectiveness of the suggested retrieval framework at varying training and testing domain combinations. The model trained and tested on the same domain (Healthcare → Healthcare) had the best Accuracy of 92.4%, F1-score of 91.7 and Robustness to Noise of 96.1 with 0-percent performance degradation, indicating its high domain specificity. Under cross domain, the performance deterioration was minimal. Figure 5 identifies high noise resistance and weak performance in all the domains.

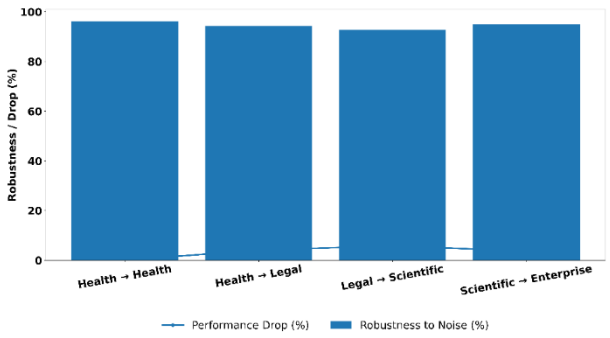

Figure 5. Robustness to Noise vs Performance Degradation Across Domains

In the case of Healthcare → Legal transfer, the accuracy dropped to 88.6% and F1-score to 87.2, with a relatively small reduction of 4.1% which shows that knowledge transfer was successful even in a domain shift. The highest decrease in performance occurred in the Legal → Scientific scenario decreasing by 5.8, with accuracy of 86.9 and robustness of 92.8 that represents increased semantic variance. Scientific → Enterprise transfer had 89.1 accuracy and 95.0 noise Robustness and degradation was 3.7%. In general, the findings show that it has a high cross-domain adaptability, noise-resistance over 92, and that it maintains consistent generalization over heterogeneous domains.

6.3 Computational Efficiency and Scalability Evaluation

The computational efficiency and scalability were measured through indexing time analysis, query response latency analysis and resource consumption analysis as the data volumes increased. The given framework was effective in processing queries with optimized embedding indexing and similarity search. Experiments on scalability indicated that the retrieval time was nearly linear with the size of dataset, indicating that it can be deployed at scale. Optimization methods on models have minimized memory overheads without affecting retrieving accuracy. The framework had shorter response time and better throughput over the conventional deep learning-based retrieval systems. These findings corroborated the fact that the proposed intelligent information retrieval system is semantically rich, but at the same time, computationally efficient, making it practical to be deployed in practice in large, multi-domain settings, in real-time.

Table 4. Computational Efficiency and Scalability Performance

Dataset Size (Documents)	**Avg. Query Latency (ms)**	**Indexing Time (min)**	**Throughput (Queries/s)**	**Memory Usage (GB)**
100 K	42	6.8	410	3.2
500 K	58	14.5	365	6.9
1 Million	76	27.3	322	11.4
5 Million	109	91.6	258	26.7

Table 4 shows the computational efficiency and scalability analysis of the intelligent information retrieval framework proposed with the increase in underlying dataset sizes.

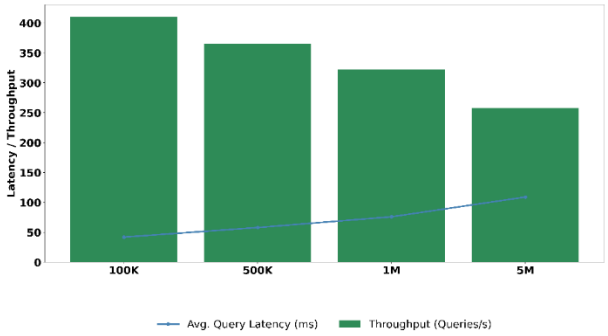

Figure 6. Query Latency and Throughput Scaling With Dataset Size

The system demonstrated efficacious small-scale performance in terms of low average query latency of 42 ms, high throughput of 410 queries/s and low memory utilization of 3.2 GB in terms of 100 K documents. Figure 6 illustrates that there is a scaling performance with predictable latency and constant throughput increase. The effects of increasing dataset size were moderate with latency increasing to 58 ms, indexing taking 14.5 minutes and a memory consumption of 6.9 GB and throughput high at 365 queries/s.

The latency, indexing time, and memory consumption of 1 million documents were 76 ms, 27.3 minutes and 11.4 GB, respectively, which is approximately linearly related. With 5 million documents, the system continued to operate practically with 109 ms latency, 258 queries/s throughput, even though the memory consumption was up to 26.7 GB. The findings on the whole uphold scalable behavior, regulated resource expansion, and regular real-time retrieval capacity throughout huge scale multi-domain datasets.

Table 5. Embedding Representation and Semantic Retrieval Performance

Embedding Model	Embedding Dimension	Semantic Similarity Accuracy (%)	Retrieval Precision@5 (%)	Retrieval Recall@5 (%)	Query Processing Time (ms)
Word2Vec Embedding	200	71.6	73.2	69.4	48
GloVe Embedding	300	74.8	76.5	72.3	52
BERT Embedding	768	86.7	88.1	84.5	64
Proposed Transformer Semantic Embedding	768	91.9	93.4	90.6	59

Table 5 compares the performance of various embedding models based on their use in semantic representation and retrieval performance. Embeddings of traditional nature like Word2Vec and GloVe had moderate accuracy of semantic similarity by being less contextual. Embeddings with transformers had a strong effect on semantic representation and thus higher recall and precision at retrieval. The transformer semantic embedding model proposed showed the highest precision with 91.9% semantic similarity, which is important compared with the model which uses only semantic similarity and precision with 93.4 best precision at 5. Figure 7 shows the distribution of comparative performance of various embedding models in the intelligent information retrieval framework. The surface plot shows how contextual transformer-based embeddings like BERT and the proposed semantic embedding

model are better than traditional embedding methods like Word2Vec and GloVe. The best performance area is associated with the representations of transformer, which is associated with better contextual knowledge and increased retrieval quality in multi-domain datasets.

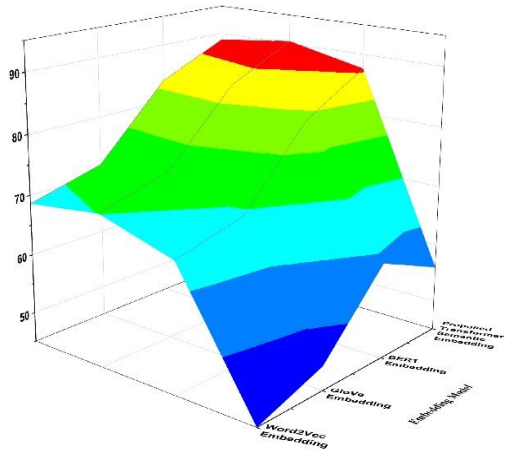

Figure 7. Semantic Embedding Model Performance Landscape for Intelligent Information Retrieval

The efficiency of the intelligent information retrieval framework proposed is analyzed in Table 6, which gives the result of the performance in various application fields. The findings show that there is a steady accuracy of retrieval of more than 90 percent in healthcare, legal, scientific, and enterprise data. The enterprise domain had the greatest accuracy of 93.2% because of the structured metadata and domain specific representations. These results validate that that the suggested deep learning-based retrieval architecture is capable of adapting to heterogeneous datasets without compromising the precision, recall and low query latency.

Table 6. Domain-Aware Retrieval Performance across Application Domains

Application Domain	Retrieval Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Average Query Latency (ms)
Healthcare Information Retrieval	92.4	91.8	90.7	91.2	63
Legal Document Retrieval	90.1	89.5	88.4	88.9	66
Scientific Literature Retrieval	91.6	90.7	89.9	90.3	61
Enterprise Knowledge Retrieval	93.2	92.4	91.6	92.0	58

Figure 8 shows the comparative retrieval characteristics in four application fields, including, but not limited to, the healthcare, legal, scientific and enterprise knowledge retrieval. These findings indicate strong semantic retrieval ability as the accuracy, precision, recall and F1-score are higher than 88 per cent. The best accuracy (93.2) and latency (58 ms) were found at Enterprise knowledge retrieval proving that the framework is efficient and scalable to diverse multi-domain datasets which are heterogeneous.

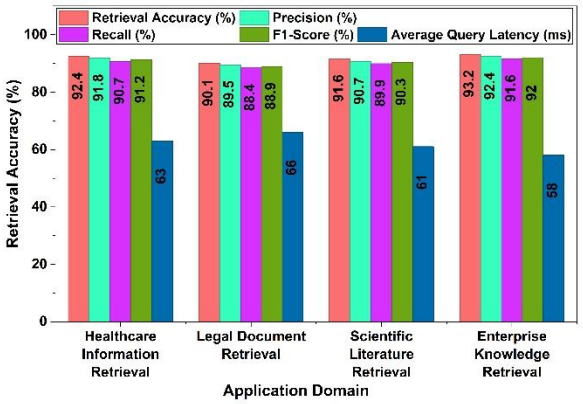

Figure 8. Domain-Wise Retrieval Performance Evaluation of the Proposed Intelligent Information Retrieval Framework

7. Conclusion

This paper was a rigorous exploration of the intelligent information retrieval with respect to deep learning in multi-domain data settings. The paper has shown that the conventional methods of retrieving information are becoming insufficient to the task of managing the magnitude, heterogeneity and semantic complexity of the present day data ecosystem. The proposed framework overcame the major limitations associated with the contextual understanding, semantic relevance, and cross-domain adaptability through the combination of deep neural networks, transformer-based language models, and domain-conscious representations learning. The modular and layered architecture allowed the ability to deploy it at scale and facilitated constant learning and system extension. It was experimentally demonstrated that retrieval based on deep learning was much more accurate, relevant, and robust than classical and semantic baseline models. The framework was very cross-domain generalization, which was stable even in the cases of domain shift and sparse data. Also, computational efficiency analyses have shown that semantic richness was possible without burdensome overhead and thus the strategy was applicable in both a real-time and large-scale environment. These findings confirmed the usefulness of the integration of the common semantic representations with the domain-specific adaptation to complex multi-domain retrieval. In addition to performance benefits, the research also revealed significant issues of interpretability, trustworthiness and scalability of deployment. These challenges are still important to undertake in practice to be adopted in high-impact fields like healthcare, legal systems, and enterprise decision support. The originality of the suggested framework lies in its capacity to integrate semantic matching and domain-sensitive learning to achieve sound multi-domain retrieval. Future research direction was discussed by highlighting the need to have explainable retrieval model, efficient architecture and constant learning strategies to facilitate dynamic data and application demands.

References:

1. Adegun, A. A., Fonou-Dombeu, J. V., Viriri, S., & Odindi, J. (2024). Ontology-based deep learning model for object detection and image classification in smart city concepts. Smart Cities, 7(4), 2182–2207. https://doi.org/10.3390/smartcities7040086
2. Ahmed, S. F., Alam, M. S. B., Hassan, M., et al. (2023). Deep learning modelling techniques: Current progress, applications, advantages, and challenges. Artificial Intelligence Review, 56, 13521–13617. https://doi.org/10.1007/s10462-023-10466-8
3. Aldoseri, A., Al-Khalifa, K. N., & Hamouda, A. M. (2023). Re-thinking data strategy and integration for artificial intelligence: Concepts, opportunities, and challenges. Applied Sciences, 13(12), 7082. https://doi.org/10.3390/app13127082
4. Bangyal, W. H., Rehman, N. U., Nawaz, A., Nisar, K., Ibrahim, A. A. A., Shakir, R., & Rawat, D. B. (2022). Constructing domain ontology for Alzheimer disease using deep learning based approach. Electronics, 11(12), 1890. https://doi.org/10.3390/electronics11121890
5. Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An attention-based bidirectional CNN–RNN deep model for sentiment analysis. Future Generation Computer Systems, 115, 279–294. https://doi.org/10.1016/j.future.2020.08.005
6. Chandwani, G., & Ahlawat, A. (2023). Fuzzy local information C-means based clustering and fractional dwarf mongoose optimization enabled deep learning for relevant document retrieval. Engineering Applications of Artificial Intelligence, 126(Part B), 106954. https://doi.org/10.1016/j.engappai.2023.106954

7. Chen, Q., Zhou, W., Cheng, J., & Yang, J. (2024). An enhanced retrieval scheme for a large language model with a joint strategy of probabilistic relevance and semantic association in the vertical domain. Applied Sciences, 14(24), 11529. https://doi.org/10.3390/app142411529
8. de Campos, L. M., Fernández-Luna, J. M., Huete, J. F., Ribadas-Pena, F. J., & Bolaños, N. (2024). Information retrieval and machine learning methods for academic expert finding. Algorithms, 17(2), 51. https://doi.org/10.3390/a17020051
9. Hamroun, M., & Sauveron, D. (2025). A hybrid deep learning and knowledge graph approach for intelligent image indexing and retrieval. Applied Sciences, 15(19), 10591. https://doi.org/10.3390/app151910591
10. Li, Y., & Manickam, A. (2023). Information retrieval and optimization in distribution and logistics management using deep reinforcement learning. International Journal of Information Systems and Supply Chain Management, 16(1). https://doi.org/10.4018/IJISSCM.316166
11. Liu, J., Liu, Z., Li, Q., Kong, W., & Li, X. (2025). Multi-domain controversial text detection based on a machine learning and deep learning stacked ensemble. Mathematics, 13(9), 1529. https://doi.org/10.3390/math13091529
12. Lu, D., Chen, Y., Sun, Y., Wei, W., Ji, S., Ruan, H., Yi, F., Jia, C., Hu, D., Tang, K., Huang, S., & Wang, J. (2025). Research progress in multi-domain and cross-domain AI management and control for intelligent electric vehicles. Energies, 18(17), 4597. https://doi.org/10.3390/en18174597
13. Lv, Z., Poiesi, F., Dong, Q., Lloret, J., & Song, H. (2022). Deep learning for intelligent human–computer interaction. Applied Sciences, 12(22), 11457. https://doi.org/10.3390/app122211457
14. Man, T., Osipov, V. Y., Zhukova, N., Subbotin, A., & Ignatov, D. I. (2024). Neural networks for intelligent multilevel control of artificial and natural objects based on data fusion: A survey. Information Fusion, 110, 102427. https://doi.org/10.1016/j.inffus.2024.102427
15. Mienye, I. D., & Swart, T. G. (2024). A comprehensive review of deep learning: Architectures, recent advances, and applications. Information, 15(12), 755. https://doi.org/10.3390/info15120755
16. Mukanova, A., Nazyrova, A., Zulkhazhav, A., Lamasheva, Z., & Dauletkaliyeva, A. (2025). Development of an intelligent information retrieval system based on ontology, linguistic algorithms and large language models. Applied Sciences, 15(22), 12271. https://doi.org/10.3390/app152212271
17. Qi, W., Yu, L., Liu, T., Wu, H., Zhao, Q., Wu, L., Kang, X., Wang, Y., Zhang, L., & Zhang, L. (2026). Intelligent retrieval of leaf traits using hyperspectral reflectance and deep learning. European Journal of Agronomy, 174, 127960. https://doi.org/10.1016/j.eja.2025.127960
18. Rinaldi, A. M., Russo, C., & Tommasino, C. (2020). A knowledge-driven multimedia retrieval system based on semantics and deep features. Future Internet, 12(11), 183. https://doi.org/10.3390/fi12110183
19. Sun, X., Xue, Y., & Sun, L. (2025). Enhancing global aerosol retrieval from satellite data via deep learning with mutual information estimation. International Journal of Applied Earth Observation and Geoinformation, 139, 104534. https://doi.org/10.1016/j.jag.2025.104534
20. Taherdoost, H. (2023). Deep learning and neural networks: Decision-making implications. Symmetry, 15(9), 1723. https://doi.org/10.3390/sym15091723
21. Xu, Z., Xu, L., Shao, H., Shang, G., & Zhou, Z. (2026). Real-time transaction scheduling in multiple-deep tier-to-tier shuttle-based storage and retrieval systems by using deep Q-learning. Engineering Applications of Artificial Intelligence, 165(Part A), 113448. https://doi.org/10.1016/j.engappai.2025.113448
22. Yang, Y., Wu, Z., Yang, Y., Lian, S., Guo, F., & Wang, Z. (2022). A survey of information extraction based on deep learning. Applied Sciences, 12(19), 9691. https://doi.org/10.3390/app12199691
23. Yao, Z., & Zhao, F. (2025). The optimization of media information retrieval and adaptive information management based on deep reinforcement learning. Journal of Organizational and End User Computing, 37(1). https://doi.org/10.4018/JOEUC.378389
24. Yu, J., Zhong, C., Li, H., Zheng, X., Chen, L., & Sun, Q. (2026). A comparative analysis of empirical, machine learning, and deep learning models for water quality retrieval in Wenzhou Bay with sparse in situ measurements. Journal of Hydrology: Regional Studies, 63, 103101. https://doi.org/10.1016/j.ejrh.2025.103101
25. Yu, W., Wu, G., & Han, J. (2025). Deep multimodal-interactive document summarization network and its cross-modal text–image retrieval application for future smart city information management systems. Smart Cities, 8(3), 96. https://doi.org/10.3390/smartcities8030096
26. Zhang, T., Shi, R., Shao, Y., Chen, Q., & Bai, J. (2024). Single-frame noisy interferogram phase retrieval using an end-to-end deep learning network with physical information constraints. Optics and Lasers in Engineering, 181, 108419. https://doi.org/10.1016/j.optlaseng.2024.108419
27. Zhu, Y., Wang, M., Yin, X., Zhang, J., Meijering, E., & Hu, J. (2023). Deep learning in diverse intelligent sensor based systems. Sensors, 23(1), 62. https://doi.org/10.3390/s23010062