Afr. J. Lib, Arch. and Inf. Sc. Vol. 35, No. 1 (July 2025) 1-19
https://doi.org/10.5281/zenodo.17387818

Artificial Intelligence–Driven Knowledge Management in Academic Libraries: A Framework for Enhancing Information Retrieval Efficiency and User-Centred Service Delivery

Hongyin Jin*
Faculty of Education, Chulalongkorn University, Bangkok, Thailand, 10330.
hyjin193@gmail.com

Abstract
The expansion of digital resources in academic libraries (AL) has intensified the need for effective information retrieval systems and services that align with users' needs. This study introduces an Artificial Intelligence-Driven Knowledge Management Framework for AL, designed to optimise retrieval processes while supporting user-focused service delivery. The framework combines Named Entity Recognition (NER) with Word Sense Disambiguation (WSD) to interpret both the semantic content and underlying intent of user queries. A voting-based NER method, built on Intelligent Monarch Butterfly Optimised Residual Bidirectional Gated Recurrent Unit (IMB-Res-BiGRU) classifiers, ensures reliable entity identification in English and Chinese texts, while an example-driven WSD technique addresses semantic ambiguities to enhance query interpretation. The framework was tested on samples drawn from English and Chinese collections, following text cleaning and normalisation procedures. For document retrieval, Bidirectional Encoder Representations from Transformers (BERT) was applied to achieve high relevance and efficiency. User-oriented service provision was accomplished by tailoring search outcomes according to contextual factors, user preferences, and information-seeking patterns, enabling students, researchers, and faculty to access precise and meaningful resources. Implementation in Python produced experimental outcomes with high accuracy, achieving 97.8% precision and 98.2% recall, thereby confirming significant advances in retrieval effectiveness alongside improved user-focused service delivery. This AI-driven framework demonstrates a scalable model for contemporary AL, uniting advanced knowledge management with personalised services to strengthen access to information and enhance user satisfaction.

Keywords: Artificial Intelligence, Knowledge Management, Academic Libraries (AL), Information Retrieval, User-Centred Services, Named Entity Recognition, Intelligent Monarch Butterfly Optimized Residual Bidirectional Gated Recurrent Unit (IMB-Res-BiGRU).

Introduction

Knowledge management in AL has become a critical approach for ensuring effective access, organisation, and utilisation of information in the digital era (Alzahrani et al., 2025). With the rapid increase in both physical and electronic resources, ALs face the challenge of managing extensive volumes of information while ensuring seamless accessibility for diverse user groups (Chukwujindu et al., 2024). Through a structured process of capturing, organising, sharing, and applying knowledge, Knowledge management enables libraries to progress from traditional resource handling to more service-oriented functions (Mosha, 2025). This approach leverages collective expertise, advanced technologies, and streamlined processes to improve information retrieval, optimise workflows, and enhance the overall user experience (Boateng, 2025).

Serving as intellectual hubs within academic settings, libraries play a vital role in supporting teaching, learning, and research innovation (Veerakannan, 2025). By applying knowledge management strategies, libraries can create an environment where information is not only stored but also transformed into practical knowledge that benefits students, faculty, and researchers (Yusof and Sulaiman, 2025). This involves adopting intelligent systems, user-focused service models, and collaborative practices that adapt to the dynamic demands of academic communities (Ademola et al., 2025). Moreover, knowledge management contributes to more informed decision-making, improved resource discovery, and greater collaboration across departments and disciplines (Meesad and Mingkhwan, 2024a). Figure 1 presents the core strategies that support effective retrieval and user-centred services within AL.

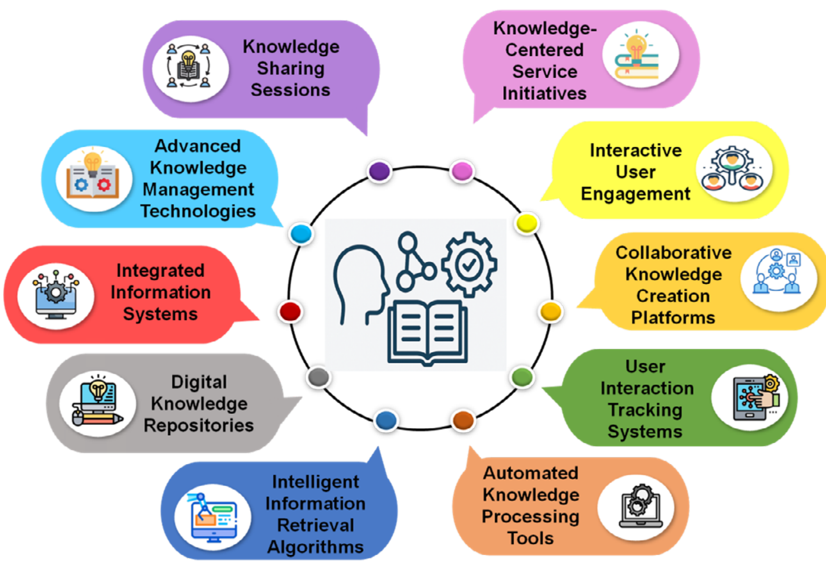


Figure 1: Framework Components for Enhancing Knowledge Management in AL.

Establishing a well-structured knowledge management system within academic institutions ensures that information is delivered with accuracy, efficiency, and in alignment with user expectations (Saadati et al., 2021). Such a system facilitates improved accessibility, minimises redundancy, and enables more personalised responses to information requests (Meesad and Mingkhwan, 2024b). As libraries continue to evolve under the influence of technological advancements and shifting user behaviours, knowledge management emerges as a cornerstone for maintaining relevance and enhancing the quality of services provided within educational contexts (Mbua, 2025). Earlier approaches to knowledge management in AL were largely dependent on rule-based processes and manual indexing. These methods have increasingly been supplemented by classification and recommendation models such as Decision Trees, Random Forest (RF), Support Vector Machines (SVM), Naive Bayes, and K-Nearest Neighbours (KNN), supported by the adoption of machine learning (ML) and deep learning (DL).

Within DL, architectures including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) contribute to improved retrieval functions, semantic searching, and customised service delivery. Despite these advances, challenges persist in applying ML and DL to knowledge management in AL, such as high computational demands, dependence on extensive labelled datasets, difficulties in processing ambiguous and multilingual queries, risks of algorithmic bias, and compatibility issues with legacy systems, all of which can hinder efficient adoption and service quality. To address these limitations, this research employs a voting-based NER method in conjunction with IMB-Res-BiGRU classifiers for multilingual entity extraction, alongside an example-driven WSD for semantic interpretation. This integrated approach ensures accurate information retrieval, reduces ambiguity, limits computational inefficiencies, and enhances user-centred, personalised service delivery within knowledge management systems for AL.

Key Contributions

- Data Collection: Bilingual academic queries, together with user preferences, retrieved documents, and feedback, were utilised to assess the effectiveness of information retrieval in AL.
- Data Pre-Processing: Text cleaning and min–max normalisation were applied to English and Chinese queries, ensuring consistency and enhancing retrieval accuracy.
- Entity Extraction: NER combined with WSD strengthened bilingual entity identification and semantic clarification, enabling more accurate interpretation of user queries.
- Information Retrieval: BERT was employed to retrieve contextually relevant documents from refined queries, thereby improving semantic accuracy and retrieval precision.
- Proposed Method: The framework applied IMB-Res-BiGRU with voting-based NER and example-driven WSD, supporting accurate bilingual entity extraction and effective query disambiguation.

The structural design of this research is outlined as follows: Section 1 presents the study background, Section 2 provides the literature review, Section 3 explains the methodology, Section 4 discusses the results, and Section 5 concludes the study.

Related Works

Muslim (2024) assessed the influence of digital transformation on access, user experience, and knowledge management in AL. The findings indicated improved accessibility of resources, higher user satisfaction, and greater retrieval efficiency through AI-driven systems, although persistent challenges in digital literacy, interface usability, and financial constraints were also identified. Similarly, Oyedokun (2025) examined strategies, challenges, and emerging technologies such as AI, Virtual Reality (VR), Augmented Reality (AR), blockchain, Internet of Things (IoT), and data analytics in transforming AL. The results underscored the significance of user-centred design, collaborative practices, promotion of open access, and strengthened roles of libraries as active centres of learning and research. Jain and Behera (2023) analysed collection development, spatial planning, futuristic technologies, and information services in AL through content analysis, highlighting their evolution into hybrid knowledge hubs and proposing a conceptual framework that projects future directions and challenges.

A comprehensive framework for integrating automation technologies, including Radio Frequency Identification (RFID), AI, and ML, into AL was developed by Ikwuanusi et al. (2024). The results demonstrated more efficient cataloguing, accelerated information retrieval, optimised resource allocation, and enhanced user-focused services. In a related study, Ikwuanusi et al. (2023) investigated AI-driven applications, including ML, Natural Language Processing (NLP), and recommender systems, for improving knowledge delivery and inclusivity in libraries. Akanbiemu (2024) examined the application of Big Data Analytics (BDA) in knowledge management practices within libraries, showing how predictive modelling and data-driven insights optimise organisation and retrieval while addressing concerns over privacy, security, and governance.

Further, Zhang et al. (2025) proposed a methodology for evaluating the quality of discipline-specific data services in AL by incorporating aspects of digital intelligence. Hamad et al. (2023b) studied the relationship between the implementation of smart information services in AL and librarians' digital competencies, while Hamad et al. (2023a) analysed the level of adoption of such services in Jordan. Their findings indicated a moderate implementation level (Mean=3.12) and moderate challenges (Mean=3.57), which included resistance to change, data privacy concerns, limited financial support, infrastructural weaknesses, and inadequate professional training. Ogungbenro et al. (2025) explored the contribution of AI to enhancing cataloguing and information access in Nigerian university libraries. Awogbami (2024) identified the essential professional skills required for managing Digital Reference Services (DRS) in AL, emphasising digital literacy, technical competence, and virtual user support. Obuh and Ogbomo (2024) investigated librarians' proficiency in web technologies for service delivery in Delta State university libraries, reporting that librarians had intermediate web technology skills, frequently used tools such as cloud computing, and enhanced both service delivery and user satisfaction.

Ademilua and Yacob (2024) examined the role of Competitive Intelligence (CI) in rural libraries for improving sustainability and service effectiveness. Dhage and Verma (2025) studied how AI reshapes modern libraries by improving user information-seeking behaviour and satisfaction. Kim (2025) explored the application of Generative AI in AL, highlighting its potential to enhance service personalisation, increase user engagement, and strengthen research productivity. Findings revealed that AI-enabled chatbots significantly improve user experience, streamline services, and redefine librarianship in digital academic environments. An IoT-powered adaptive framework integrating sensor data, image-based occupancy monitoring, and user feedback for optimising study environments in university libraries was developed by Mammadov and Kucukkulahli (2025). Results showed that KNN achieved the highest F1 score, confirming the framework's reliability. Faizan and Munshi (2025) assessed ICT-based library services at the Indian Institute of Technology (IIT) from the perspective of users. Their findings demonstrated that ICT infrastructure had a strong positive impact on academic performance and user satisfaction, though 28.5% of participants reported difficulties in accessing specific information.

Research Gaps

Although several AI-based systems enhanced resource accessibility and retrieval efficiency, their wider adoption was constrained by continuing challenges such as limited digital literacy, interface design weaknesses, and financial restrictions, which reduced long-term user engagement (Muslim, 2024). Similarly, shortcomings in staff training, infrastructural limitations, privacy risks, and resistance to change led to only partial implementation of smart data services in AL, thereby restricting their effectiveness and scalability (Hamad et al., 2023a). The application of BDA supported improvements in knowledge organisation and retrieval; however, issues concerning data security, privacy, and governance hindered its full integration within library contexts (Akanbiemu, 2024). To address these constraints, the IMB-Res-BiGRU approach strengthens entity recognition, streamlines retrieval processes, and enhances adaptability by tackling concerns related to privacy, infrastructure, and staff preparedness, ultimately facilitating efficient, scalable, and user-focused services in AL.

Methodology

The methodology employed bilingual academic queries from the AL dataset, incorporating user inputs, preferences, and relevance-based feedback. During data pre-processing, text cleaning and min–max normalisation were applied to ensure consistency and improve retrieval performance. Entity extraction was carried out through a voting-based NER combined with WSD, which enabled accurate identification and clarification of entities. For information retrieval, BERT was utilised to generate precise and contextually relevant document recommendations. The proposed framework integrates IMB-Res-BiGRU with voting-based entity recognition and example-driven WSD, thereby improving retrieval efficiency and supporting user-focused services. Figure 2 illustrates the overall methodological process.

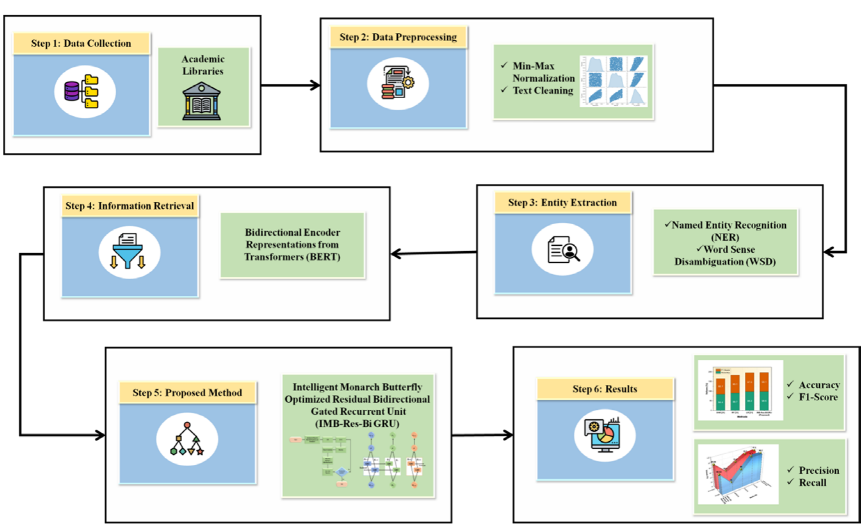


Figure 2: Overall Knowledge Management Flow in AL Framework.

Data Collection

The AL dataset comprises bilingual academic search queries in English and Chinese, developed to facilitate research on knowledge management and information retrieval within AL. It incorporates user queries, pre-processed text, identified entities, retrieved and relevant documents, user preferences, and feedback. This dataset provides a foundation for evaluating retrieval effectiveness and supports investigations into digital libraries, personalised recommendations, and user-centred knowledge management services (Kaggle, n.d.).

Data Pre-Processing using Text Cleaning

Text cleaning represents a vital stage in preparing textual data to ensure both quality and uniformity prior to further analysis. This process eliminates redundant components such as punctuation, special symbols, numerical values, excess spaces, and stop words that do not add analytical value. It also standardises the dataset by applying a consistent case format and correcting spelling errors or textual inconsistencies. Through these steps, the data is refined, structured, and rendered suitable for accurate and efficient analysis, modelling, and information retrieval tasks.

Min-Max Normalization

Min–max normalisation adjusts values to fall within a specified range, thereby ensuring consistency across datasets. Within the context of knowledge management in AL, this technique enhances the efficiency of information retrieval and supports user-centred service delivery by improving both query interpretation and the relevance of retrieved results. The procedure is widely applied to transform data into a normalised state. In this process, all values are rescaled into a decimal range between 0 and 1, with the absolute minimum and maximum values of the input data component adjusted accordingly. Equation (1) expresses the calculation of the normalised value for each input element.

$$X_j' = NeX_{min} + (NeX - NeX_{min}) * \left(\frac{X_j - X_{min}}{X_{max} - X_{min}}\right) \quad (1)$$

Here, X_{min}, X_{max} *to* NeX_{min} - NeX_{max} statistics were proportionately modified. Where X_j as is the input data and X_j' as is the normalized value. This approach retains the proportional relationships among data values, thereby ensuring reliability and making it a preferable method for processing. Figure 3 illustrates the normalisation process applied in evaluating knowledge management within AL.

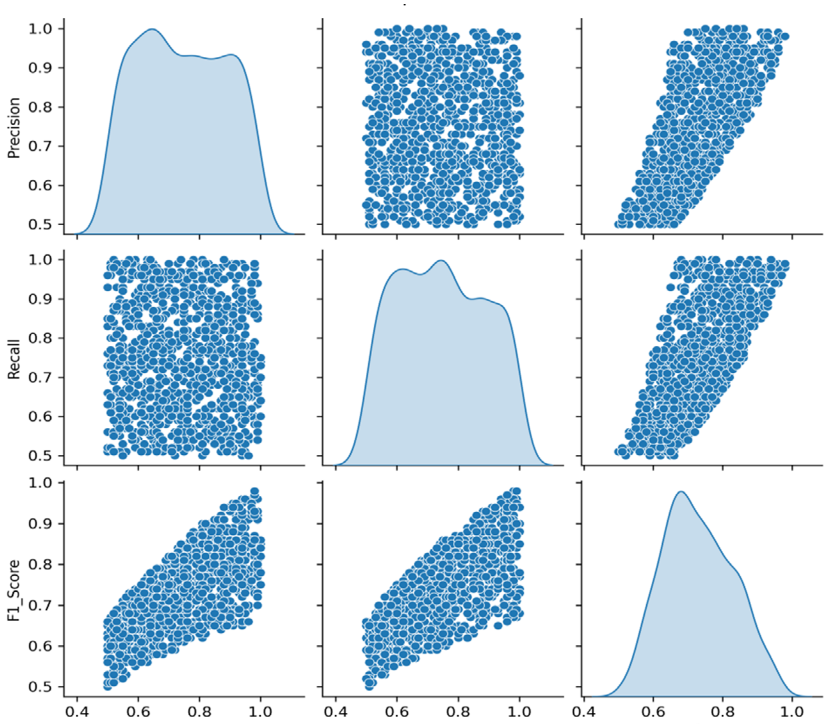


Figure 3: Normalization of Workflow for Optimizing Knowledge Management in Libraries.

Entity Extraction using Voting-based Named Entity Recognition (NER)

Entity recognition involves detecting significant components within user queries, including individuals, subjects, institutions, or concepts. By accurately extracting these elements, the system is able to concentrate on the user's actual search intent. This process ensures correct interpretation of both English and Chinese queries, thereby minimising ambiguity and enhancing the precision of information retrieval.

Word Sense Disambiguation (WSD)

WSD resolves semantic ambiguity by identifying the appropriate meaning of words that carry multiple interpretations. Through an example-driven approach, the system evaluates query terms against context-specific instances to determine the intended sense. This process sharpens query interpretation, ensuring alignment with the user's actual intent and thereby improving retrieval relevance while reducing mismatches in the results.

Information Retrieval using BERT

BERT generates rich contextual representations of academic library texts, facilitating a more comprehensive interpretation of user queries. This contributes to enhanced information retrieval accuracy in knowledge management and strengthens the delivery of personalised, user-focused services within the proposed framework. The model's upper layers can be fine-tuned to perform tasks such as entity recognition and sentiment analysis. Through its bilateral self-attention mechanism, BERT effectively captures the contextual relationships within the text. During its pre-training stage, the model acquires linguistic patterns from a vast corpus using self-supervised learning, and the learned parameters are subsequently applied to text classification tasks to optimise performance.

By modelling word-to-word relationships, the self-attention mechanisms embedded in each Transformer encoder layer enable the construction of long-range dependencies. The sequence of input text was expressed as $W = [w_1, w_2, ..., w_n]$. The output representation was expressed as $G = [G_1, G_2, ..., G_n]$. The BERT model dynamically assigns contextual weights to individual words, enabling the generation of more accurate text representations. For text classification, the model processes input through its final layer to produce a consolidated representation of the entire sequence. To achieve the classification objective, the output from BERT is passed through a fully connected layer, which maps the representation into the appropriate classification space. The process of the BERT model is mathematically represented in Equation (2).

$$Z = softmax(X.d_{cls} + a) \quad (2)$$

Here, d_{cls} represents the last layer of the BERT model, a and indicate the weight biases, and Z as the predicted output value. BERT enhances model robustness by efficiently transforming complex textual information into meaningful features for classification tasks. Through the optimisation of model parameters by minimising the cross-entropy loss function, the training process achieves improved classification performance. Consequently, BERT demonstrates strong potential for advancing text categorisation outcomes and increasing overall classification accuracy.

IMB-Res-BiGRU Classifies to Ensure Accurate Entity Extraction in English and Chinese Texts

The hybrid IMB-Res-BiGRU offers an advanced solution for strengthening knowledge management in AL by integrating deep learning with optimisation techniques. This model utilises the Monarch Butterfly algorithm to optimise parameter selection, while the residual structure addresses vanishing gradient challenges, thereby ensuring stable and efficient learning across extended sequences. The bidirectional gated recurrent unit processes contextual information in both forward and backward directions, making it particularly effective in multilingual contexts such as English and Chinese. Within the proposed framework, the model is applied to NER, allowing precise entity extraction and a refined interpretation of user queries. Through this integration, the system significantly improves information retrieval efficiency and delivers user-centred services that are accurate, adaptive, and contextually relevant for students, researchers, and academic staff. The operational procedure of the proposed IMB-Res-BiGRU is outlined in Algorithm 1.

Algorithm 1: IMB-Res-BiGRU

```
BEGIN IMB_Res_BiGRU
1. Data Preprocessing
  - Load dataset: 10,000 samples (train = 70%, val = 15%, test = 15%)
  - Normalize features into [0,1]
  - Sequence length = 100 timesteps, feature dimension = 20
2. Initialize IMB Parameters
  - Population size (pop_size) = 30
  - Maximum iterations (max_iters) = 50
  - Elite size (elite_k) = 3
  - Search space:
    learning rate ∈ [0.0001, 0.01]
    hidden units ∈ {64, 128, 256}
    residual blocks ∈ {1, 2, 3}
    dropout ∈ [0.1, 0.5]
    batch size ∈ {32, 64}
    dense units ∈ {64, 128}
  - Set best_fitness = +∞
3. Initial Fitness Evaluation
  FOR each candidate in 30 population:
    - Build Residual BiGRU model (1–3 blocks, 64–256 units)
    - Train for E_inner = 5 epochs, batch = candidate batch size
    - Compute validation loss
  END FOR
    - Select best candidate
4. IMB Optimization Loop
  FOR gen = 1 to 50:
    - Retain top 3 elite candidates
    - FOR i = 1 to (pop_size - elite_k):
      Select parent1, parent2 from current population
      Apply migration operator:
      child_param = 0.7*parent1 + 0.3*parent2 + noise(σ=0.05)
      Apply mutation with prob = 0.2 decreasing to 0.05 by gen 50
    - END FOR
    - Evaluate all new candidates for 5 epochs
    - Update best_fitness if improved
  END FOR
5. Final Training with Best Hyperparameters
  - Best HP found: {lr=0.001, units=128, blocks=2, dropout=0.3, batch=32, dense=128}
  - Build Residual BiGRU:
    Block 1: BiGRU (128 units) → Dropout (0.3) → Residual Add
    Block 2: BiGRU (128 units) → Dropout (0.3) → Residual Add
  - GlobalAveragePooling → Dense (128, ReLU) → Dropout (0.3)
  - Train model: epochs = 100, batch = 32, learning rate = 0.001
6. Testing and Evaluation
  - Evaluate on 1,500 test samples
7. Output
  - Best hyperparameters, final model, and evaluation metrics
  END IMB_Res_BiGRU
```

Residual Bidirectional Gated Recurrent Unit

Res-BiGRU strengthens the proposed knowledge management framework by enhancing sequential data processing for information retrieval in AL. The residual connections facilitate stable gradient propagation, preventing performance degradation during training, while the bidirectional structure captures contextual relationships in both directions. This design ensures more accurate entity extraction and query interpretation, thereby advancing the delivery of precise and user-focused services.

BiGRU

GRU, a widely used variant of recurrent neural networks, represents a simplified adaptation of LSTM. By reducing the number of network parameters and simplifying the gating mechanism, GRU achieves comparable outcomes within the same number of iterations while exhibiting a lower risk of overfitting. This enables GRU to retain the performance benefits of LSTM while streamlining the overall network architecture. The structure consists of two principal gates: the update gate, which governs information retention, and the reset gate, which determines the removal of irrelevant data. Their interactions are mathematically represented in Equations (3)–(6). Figure 4 illustrates the architecture of the GRU hidden layer unit.

$$Y_s = \sigma\left(X_y[G_s - 1, W_s]\right) \quad (3)$$

$$Q_s = \sigma(X_Q[G_s - 1, W_s]) \quad (4)$$

$$\widehat{G}_s = \tanh\left(X_G[Q_s \odot G_s - 1, W_s]\right) \quad (5)$$

$$G_s = (1 - Y_s) \odot G_s - 1 + Y_s \odot \tilde{G}_s \quad (6)$$

At time step *s*, Y_s is the update gates and Q_s is the reset gates. At time step *s*, $\tilde{G}_s$ represents the status of the hidden layer's unit; it also acts as an input steps; G_s - *1* represents the hidden layer unit's state at the previous instant, whereas W_s is the input at the present step . BiGRU is constructed by combining a forward GRU with its reverse counterpart, thereby processing both forward and backward time series data. Within the hidden layer, two units operate with the same inputs and outputs, where one captures the forward sequence while the other models the backward sequence. This structure, expressed in Equation (7), enhances the representation of sequential features by enabling the model to consider bidirectional dependencies. As a result, it improves learning efficiency and accuracy, particularly in tasks involving longer sequences. The overall architecture of the bidirectional gated recurrent unit is illustrated in Figure 4.

$$U_s = [\vec{G} : \overleftarrow{G}] \quad (7)$$

The gated recurrent unit's forward state is denoted by $G^{\rightarrow}$,, and its backward state by $G^{\leftarrow}$.

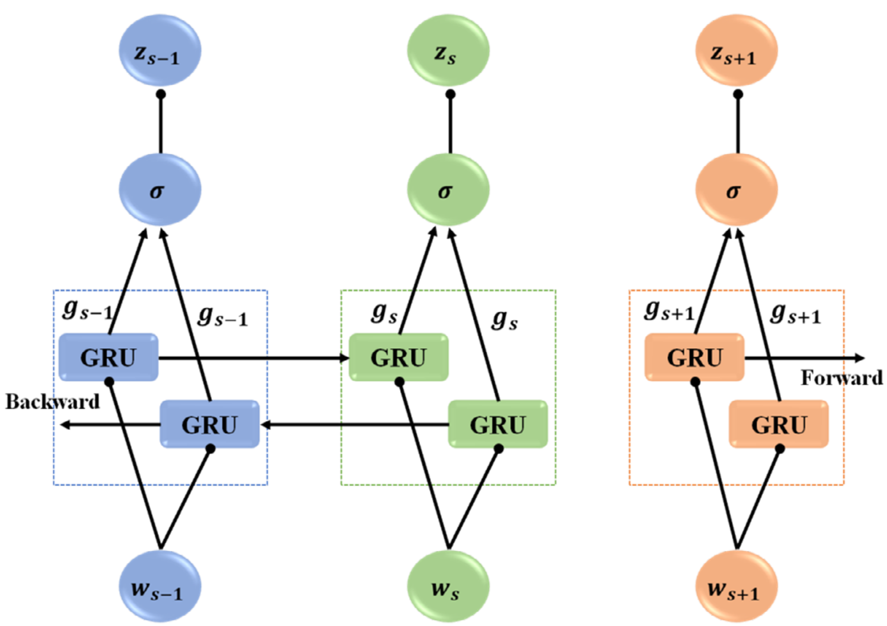


Figure 4: Structure of BiGRU.

Res-BiGRU

Convolutional neural networks generally consist of convolutional layers, pooling layers, and fully connected layers. The convolutional layer is responsible for extracting local features, where different kernels serve as distinct feature detectors. The pooling layer reduces dimensionality by selecting representative features, thereby lowering the number of parameters and improving computational efficiency. However, as the number of convolutional layers increases, challenges such as gradient vanishing and gradient explosion may arise, making it necessary to determine an appropriate network depth. To address this, deeper neural architectures employ skip connections, as expressed in Equation (8), which allow data to bypass intermediate layers and directly activate subsequent ones. The structure of the residual convolutional block is illustrated in Figure 5.

$$G(Z_0) = E(Z_0) + Z_0 \quad (8)$$

Z_0 is the residual network's input, $E(Z_0)$ is its output, and $G(Z_0)$ is the outcome of the residual network block's skip connections.

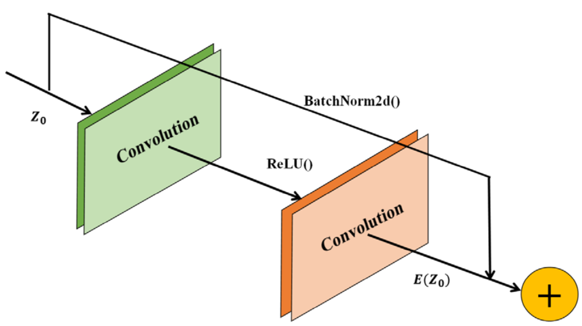


Figure 5: Residual Convolutional Block.

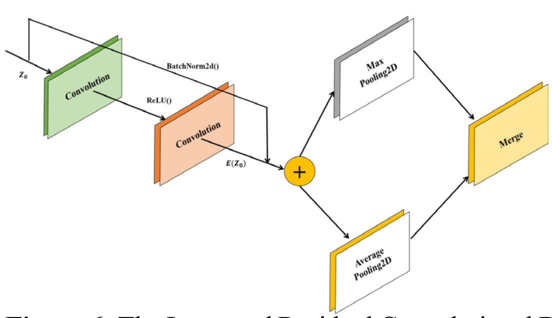


Figure 6: The Improved Residual Convolutional Block.

Although convolutional neural networks are highly effective in extracting spatial features from data, their limited receptive field often leads to increased variance in estimated values. Moreover, parameter inaccuracies within convolutional layers may result in deviations in the estimated mean. To overcome these limitations and strengthen the capacity of residual network blocks in capturing spatial information, enhancements have been introduced to their structural design. The topology of the improved residual block is illustrated in Figure 6.

An improved residual network structure was developed by incorporating the combined benefits of maximum pooling and average pooling, as outlined in Equation (9). This refined design enhances the model's capacity for feature extraction and optimises information representation. Figure 7 illustrates the activations across the layers, demonstrating their role in improving the efficiency of information retrieval within academic libraries.

$$\hat{G}(Z_0) = Max\_pooling(G(Z_0)) + Average\_pooling(G(Z_0)) \quad (9)$$

The outcome of double pooling is $\hat{G}(Z_0)$.

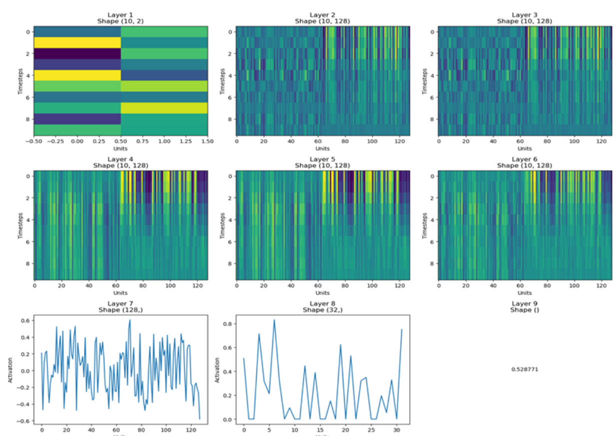


Figure 7: Activation Visualization of IMB-Res-BiGRU Layers in Knowledge Management.

Intelligent Monarch Butterfly Optimization

The IMB-Optimised framework strengthens knowledge management in AL by refining the named entity recognition process, thereby enhancing information retrieval accuracy. It enables precise extraction of essential entities, minimises semantic ambiguities, and facilitates personalised, user-focused service delivery. This optimisation ultimately improves retrieval efficiency and accessibility of academic resources across multilingual datasets. The initial stage of modelling the Monarch Butterfly (MB) Optimisation involves simulating the natural migration process of MBs, guided by specific rules. The MB population is divided into two distinct groups, each inhabiting separate subpopulations referred to as Land 1 and Land 2. Collectively, the individuals from Land 1 and Land 2 represent the complete MB population. The population distribution within Land 1 is expressed mathematically in Equation (10).

$$Land1 = ceil(o \times NP) \times (NP_1) \quad (10)$$

Similarly, the distribution of individuals in the subpopulation of Land 2 is represented mathematically, as shown in Equation (11).

$$Land2 = NP - NP_1 \times (NP_2) \quad (11)$$

If *NP* stands for the total population and *o* is the MB ratio in *Land* 1, *ceil*(*w*) rounds to the closest number greater than or equal to *w*.

The reproduction process within the MB optimisation framework generates offspring by selecting individuals from either Land 1 or Land 2. During this process, older MB individuals are replaced if the newly produced offspring demonstrates superior fitness; otherwise, the offspring is discarded, thereby preserving the stability of the population size. To maintain consistency, the parent individual remains unchanged under these conditions. The most fit MB individuals transmit their genetic characteristics directly to subsequent generations, ensuring the continuity of advantageous traits. This mechanism supports population efficiency and stability throughout generational evolution. The migration process is mathematically represented in Equation (12).

$$w_{i,l}^{s+1} = w_{q_1,l}^{s} \quad (12)$$

$w_{j,l}^{s+1}$ indicates the *lth* component of *wj* at generation $s + 1$, which establishes the location of MB*j*, whereas $w_{q_1,l}^{s}$ denotes the th component of wq_1, which displays the new position of the individual *q*1. Here, *s* is the number of current generations. From *Land* 1 individual *q*1 is chosen at random. When the condition o≤q is satisfied, Equation (13) is applied to derive the component l in the newly generated individuals. The value of o is obtained using the following formulation.

$$q = \delta \times \tau \quad (13)$$

In this context, δ represents a uniformly distributed random integer, while τ denotes the migration duration. For cases where o>q, the component *l* in the newly generated individuals is determined according to Equation (14).

$$w_{j,l}^{s+1} = w_{q_2,l}^{s} \quad (14)$$

Where $w_{q_2,l}^{s}$ is the individual $q_2 l$ th new component location of wq_2. From *Land* 2, individual q_2 was chosen at random. According to the equations, more elements will come from *Land* 1 if is large, while more have been chosen from *Land* 2 if is small. The position of the MB individuals is to be updated by an additional operator in addition to the migratory operator. If a book's number of *rand* for an MB*i* is less than or equal to *o*, it will be modified as follows in Equation (15).

$$w_{j,l}^{s+1} = w_{Best,l}^{s} \quad (15)$$

Where $w_{Best,l}^{s}$ denotes the *lth* individuals of w that indicates the best MB in *Land* 1 and2, and $w_{j,l}^{s+1}$ indicates the *lth* individual of w_i at generation $s + 1$ to show the position of the MB*i*. If δ exceeds *o*, the situation will be as follows in Equation (16).

$$w_{i,l}^{s+1} = w_{q_3,l}^{s}, q_3 \in [1,2, \ldots .NP_2] \quad (16)$$

Where $w_{q_3,l}^{s}$ denotes the *lth* randomly chosen member of wq_3 in Land 2. If δ exceeds Q_{ba} under these circumstances, it can be modified using the subsequent in Equation (17).

$$w_{i,l}^{s+1} = w_{i,l}^{s+1} + \alpha \times (dw_l - 0.5) \quad (17)$$

Where *dw* is the MB's walk step *i*, which may be obtained by using Levy flying in the manner described below, and Q_{ba} is the rate of butterfly adjustment in Equation (18).

$$dw = Levy(w_i^s) \quad (18)$$

The weighting coefficient, represented as α in Equation (24), is derived from the formulation presented in Equation (19).

$$\alpha = T_{max}/s^2 \quad (19)$$

Where T_{max} is the highest walk step that an MB individual can complete in a single stride. When *α* is larger, the search phase is longer, increasing the impact of on $w_{i,l}^{s+1}$, which reinforces the algorithm's exploration term. Conversely, a smaller value of results in a shorter search phase, lessens the impact of *dw* on $w_{i,l}^{s+1}$, and reinforces the exploitation term.

Among the recent optimisation techniques, the MB algorithm generally achieves effective outcomes but can occasionally become trapped in local optima. This limitation arises from premature convergence, which is caused by the random generation of solutions. Although MB starts with a population characterised by high diversity, the updating phase gradually reduces the variation among individuals, lowering algorithmic diversity and ultimately producing solutions that converge to local optima. To mitigate this issue and strengthen evolutionary performance, the first enhancement mechanism integrates mutation with the Improved Monarch Butterfly (IMB) algorithm. In this approach, the diversity within IMB is incorporated as outlined in Equation (20).

$$C = \frac{1}{m \times K} \times \sum_{i=1}^{m} \sqrt{\left(E_j - \bar{E}_j\right)^2} \quad (20)$$

Where K is the length of the largest diagonal lines in the solution space, m is the total number of populations, and Ej - $E^{-}j$ is within the interval [0, 1]. $E^{-}j$ represents the mean value for the E_j as the cost value of the *jth* person.

If $C < C_{low}$, where $C_{low} \in [0,4,1]$, then the population is very diverse under these conditions. Here, the recently modified branch individuals may be regarded as follows by using the mutation mechanism in Equation (21).

$$w_{i,l}^{s+1} = w_{i,l}^{s+1} + \alpha \times (dw_l - 0.5) + \varphi \times \sigma \times \tau \quad (21)$$

Where σ is a specified parameter, $\tau \sim M(0,1)$ $\alpha \in 1 - C_{low}$, $\varphi \geq 10 \times C_{low}$. Additionally, the following is how the anticosine system has been used to build δ as a self-learning mechanism in Equation (22).

$$\delta = Land_{low} + (1 + Land_{high} - Land_{low}) \times \left(1 - \arccos\left(\frac{\left(\frac{-2 \times M_{it}}{It_{max}} + 1\right)}{\pi}\right)\right) \times \alpha \quad (22)$$

Where $Land = [Land\ 1, Land\ 2]$, $Land_{low}$, and $Land_{high}$ stand for the population's lower and higher values, respectively. It_{max} specifies the value of the greatest iteration, whereas M_{it} establishes the number of iterations. The schematic flow of the IMB process is illustrated in Figure 8, while Table 1 presents the hyperparameter configuration for the proposed method.

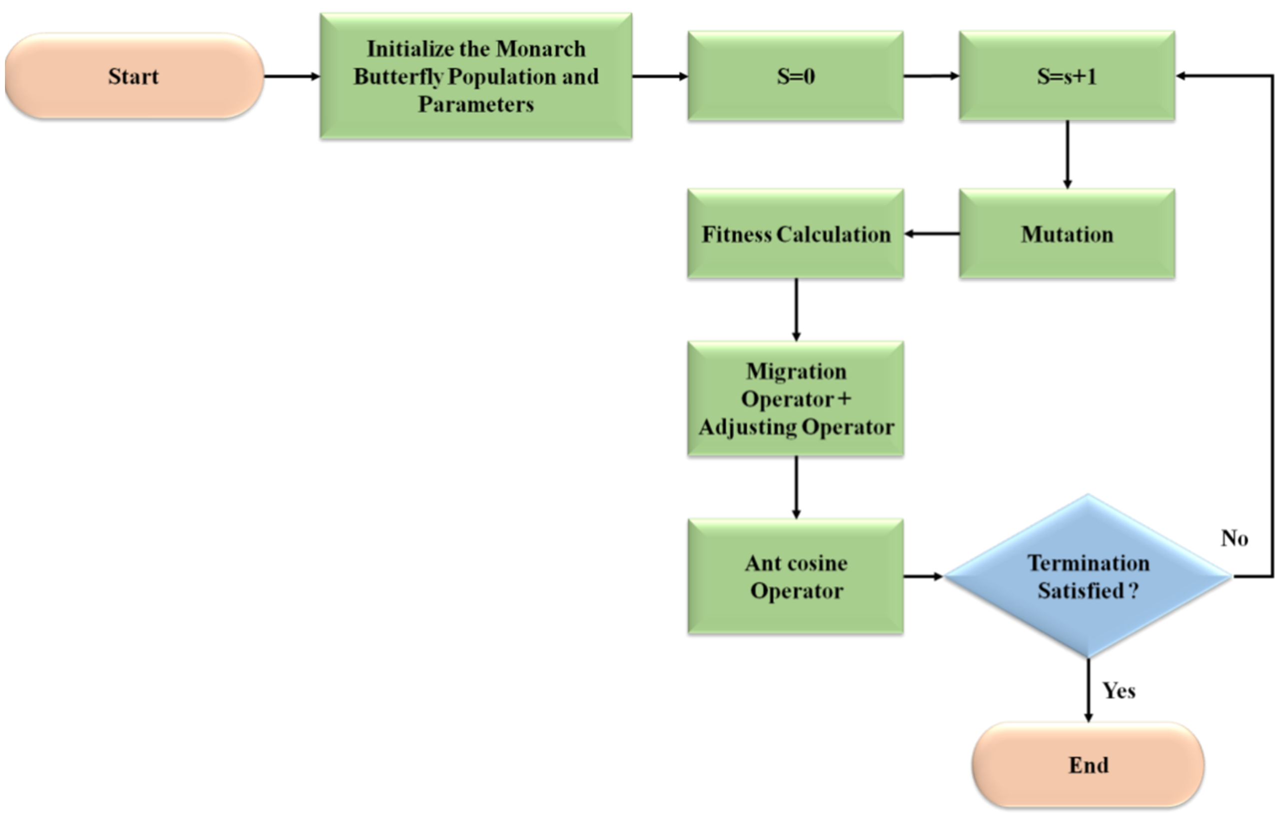


Figure 8: Flowchart of IMB.

Table 1: Hyperparameters of IMB-Res-BiGRU.

Hyperparameter	Search Range / Options	Best Value (Example)
Learning Rate (*lr*)	[0.0001,0.01]	0.001
GRU Hidden Units	{64, 128, 256}	128
Residual Blocks	{1, 2, 3}	2
Dropout Rate	[0.1,0.5]	0.3
Batch Size	{32,64}	32
Dense Layer Units	{64, 128}	128
Population Size	30	30
Maximum Iterations	50	50
Elite Size (Top-k)	3	3
Inner Epochs (E_inner)	3 – 5	5
Final Epochs (E_final)	50 – 150	100
Early Stopping Patience	5 – 15	10

Result

The experimental implementation of the proposed method was carried out in Python, and the outcomes are presented in this section. The evaluation of the model's performance employed four key metrics: Accuracy, Recall, Precision, and F1-Score. To assess its effectiveness, a comparative analysis was conducted against established approaches, including the Semantic Retrieval Model based on BERT and Knowledge Graph (Semantic Retrieval BERT-KG) (Li, 2025), SVM (Zhou and Huang, 2023), RF (Zhou and Huang, 2023), and Linear Regression (LR) (Zhou and Huang, 2023). Figure 9 illustrates the comparative performance of English and Chinese datasets within knowledge management systems using the IMB-Res-BiGRU framework. The results indicate that English queries demonstrated higher retrieval efficiency, while Chinese queries showed similar performance with only marginal differences. These findings confirm the adaptability of the proposed framework across multilingual contexts, ensuring accurate information retrieval and effective service delivery in AL without compromising performance across languages.

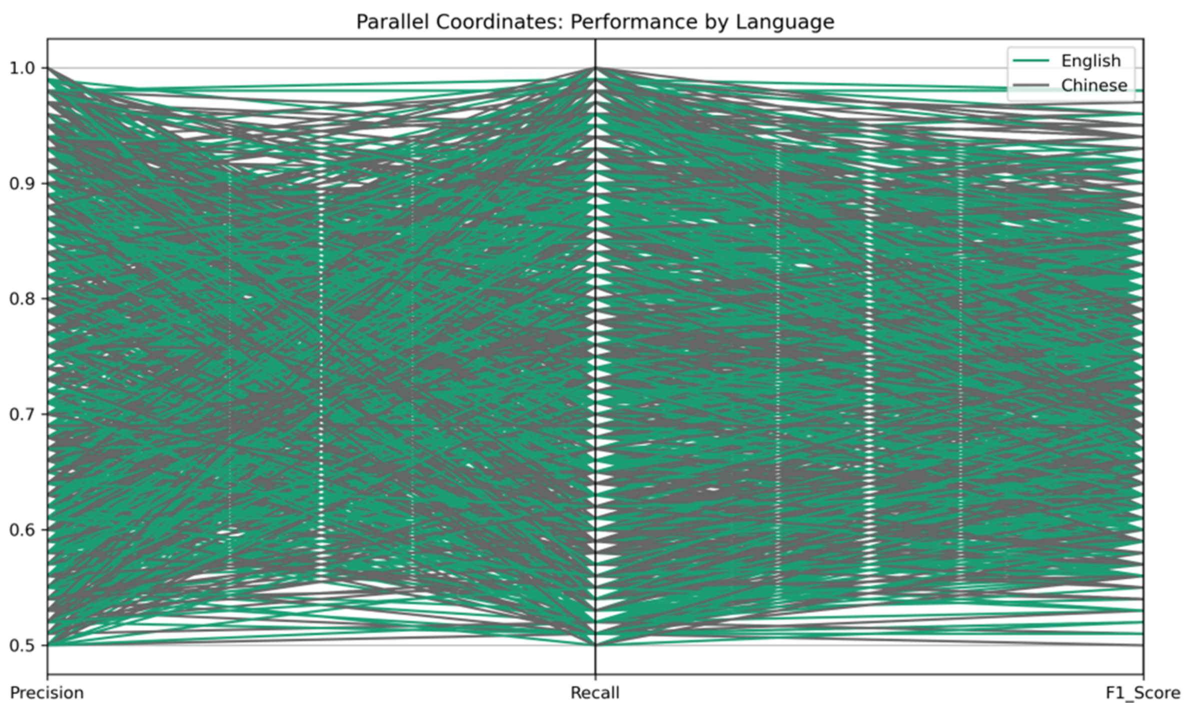


Figure 9: Parallel Coordinates across Languages for Retrieval Efficiency Metrics.

Figure 10 presents the heat density estimation of retrieval performance within knowledge management systems. Areas of high density reflect greater consistency in retrieval outcomes across the evaluated datasets. By employing the IMB-Res-BiGRU method, the framework demonstrates adaptability across diverse academic contexts, thereby strengthening user-centred service delivery. The analysis highlights regions of elevated retrieval performance, supporting efficient access to resources and contributing to improved decision-making in AL management practices.

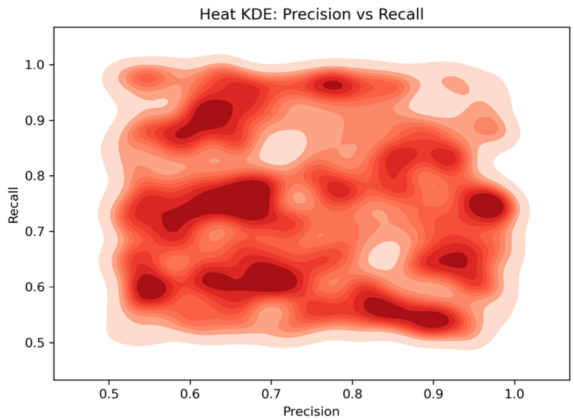


Figure 10: Heat Map of Retrieval Performance Distribution across Metrics.

Figure 11 illustrates clustered retrieval performance, emphasising the variations observed between English and Chinese datasets within knowledge management systems. The clusters represent distinct performance groups that correspond to varying levels of retrieval efficiency achieved by the IMB-Res-BiGRU framework. This clustering analysis demonstrates the model's robustness in managing multilingual contexts, ensuring stable service delivery, enhanced resource accessibility, and optimised user-centred information retrieval in AL across diverse usage environments.

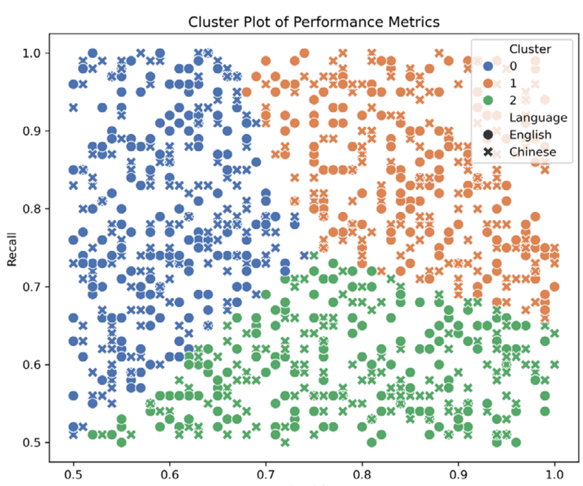


Figure 11: Cluster Distribution of Retrieval Performance across Language Metrics.

The comparative performance evaluation reveals that SVM achieved an accuracy of 84.4% and an F1-score of 80.7%, whereas RF obtained 89.7% accuracy and 93.1% F1-score. LR demonstrated markedly improved outcomes, reaching 98.2% accuracy and 97.5% F1-score. However, the proposed IMB-Res-BiGRU model surpassed all benchmarks, attaining the highest accuracy of 98.5% and an F1-score of 98.1%. These findings highlight the superiority of the proposed framework, confirming its enhanced reliability and precision in the evaluated task. A detailed comparison of performance metrics is presented in Table 2 and visualised in Figure 12, underscoring the model's capacity to improve library retrieval efficiency.

Table 2: Model Performance Comparison for IMB-Res-BiGRU Knowledge Management Framework.

Methods	Accuracy (%)	F1-Score (%)
SVM	84.4	80.7
RF	89.7	93.1
LR	98.2	97.5
IMB-Res-BiGRU [Proposed]	**98.5**	**98.1**

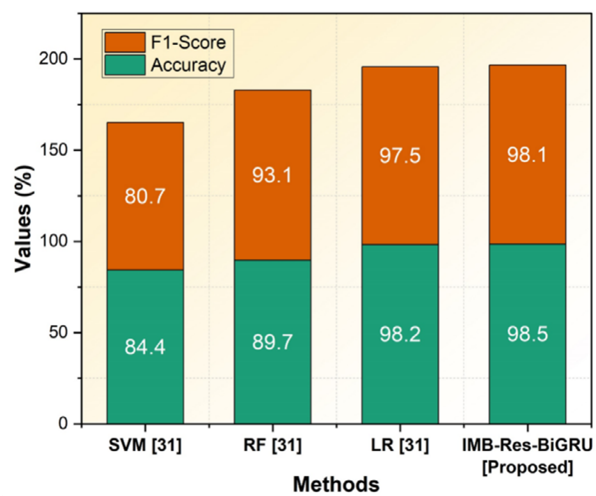


Figure 12: Accuracy and F1-Score Visualization of Library Retrieval Models.

The comparative evaluation of retrieval models demonstrates notable variations in precision and recall performance. The Semantic Retrieval BERT-KG attained 95.08% precision and 90.02% recall, reflecting robust semantic interpretation. In contrast, SVM delivered comparatively lower results, with 85.3% precision and 79.1% recall. RF enhanced performance, achieving 91.4% precision and 88% recall. LR further improved these metrics, recording 96.7% precision and 97.8% recall. The proposed IMB-Res-BiGRU model surpassed all existing approaches, achieving the highest precision of 97.8% and recall of 98.2%. These outcomes underscore the model's superior effectiveness in improving retrieval accuracy and reliability within the framework. A detailed comparison of the performance metrics is provided in Table 3 and illustrated in Figure 13, emphasising the model's contribution to advancing knowledge management in AL.

Table 3: Performance Metrics Comparison for Knowledge Management in AL.

Methods	Precision (%)	Recall (%)
Semantic Retrieval BERT-KG	95.08%	90.02%
SVM	85.3%	79.1
RF	91.4%	88%
LR	96.7%	97.8%
IMB-Res-BiGRU [Proposed]	**97.8%**	**98.2%**

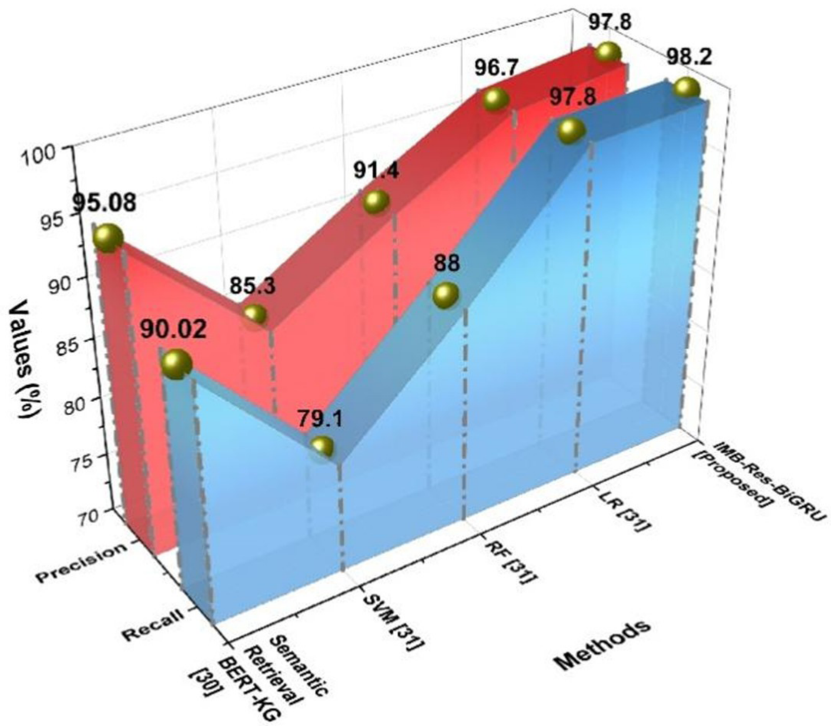
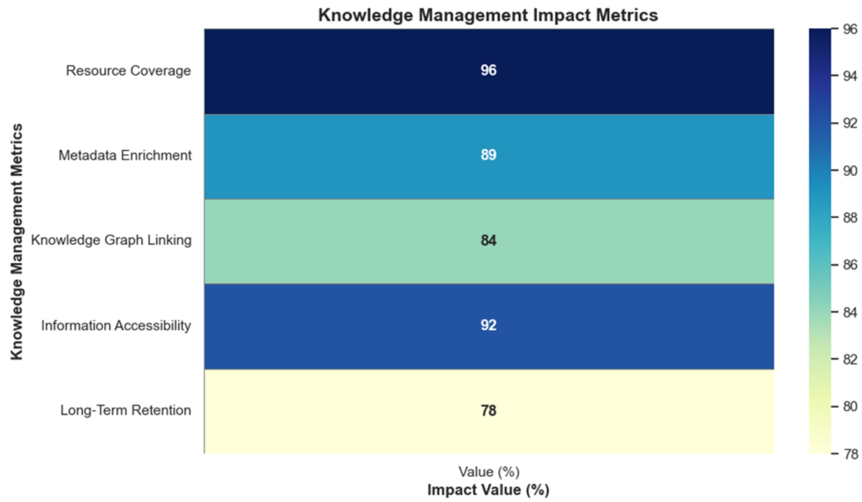


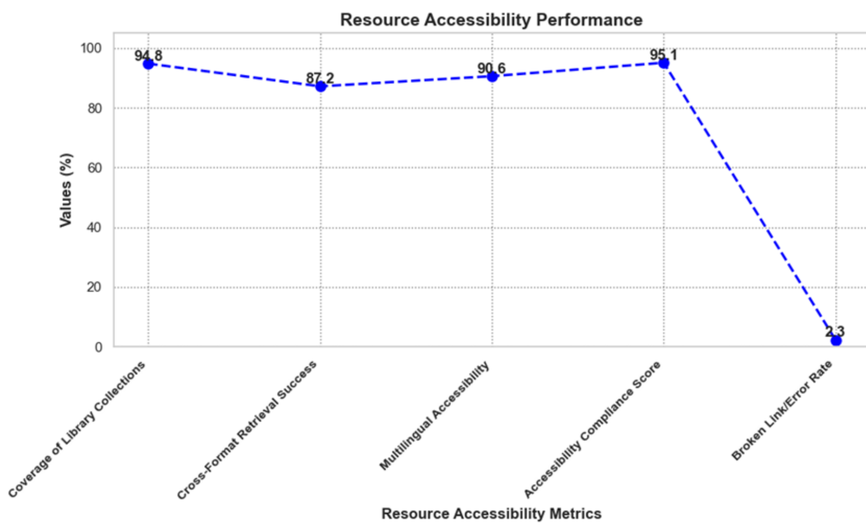
Figure 13: Precision and Recall Comparison for Knowledge Management Framework.

The assessment of the proposed knowledge management framework reveals substantial improvements across four critical dimensions: resource accessibility, retrieval efficiency, user-centred service, and overall knowledge management impact. With respect to resource accessibility, the framework demonstrated extensive coverage of library collections (94.8%), strong cross-format retrieval success (87.2%), and high multilingual accessibility (90.6%). It also achieved a superior accessibility compliance score (95.1%) while maintaining a minimal broken link and error rate (2.3%). Regarding retrieval efficiency, the system exhibited optimised performance with an average query response time of 180 ms, reduced indexing latency of 0.35 seconds per document, and strong scalability of 520 documents per second. Resource utilisation remained balanced, with memory usage at 62% and CPU/GPU consumption at 71%.

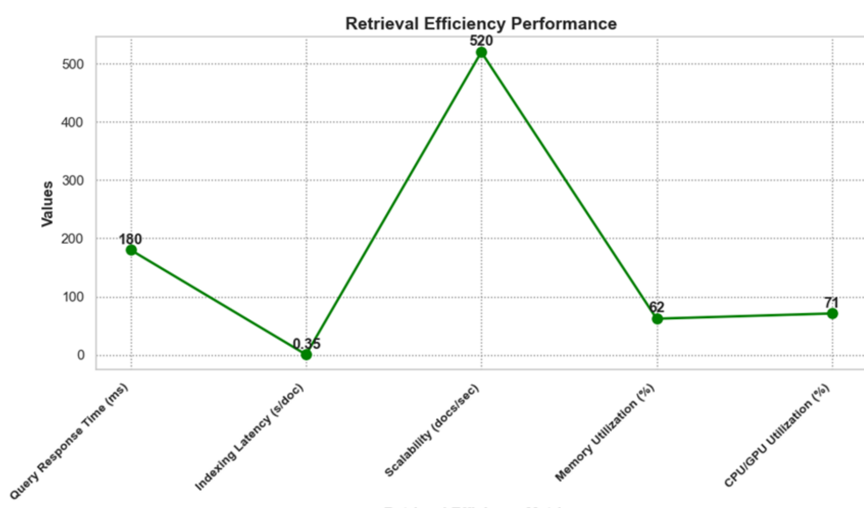
From the user-centred perspective, the framework delivered enhanced personalisation (0.87), strong adaptability to context (88%), valuable recommendation quality (4.3/5), and high diversity in results (79%), which collectively contributed to a user satisfaction rate of 91%. In terms of knowledge management impact, the framework achieved 96% resource coverage, 89% metadata enrichment, 84% integration with knowledge graphs, 92% information accessibility, and a long-term retention rate of 78%. These outcomes demonstrate the framework's capacity to support effective knowledge discovery, expand accessibility, and sustain long-term information utility. Figure 14 illustrates the evaluation results, highlighting performance across the four dimensions of resource accessibility, retrieval efficiency, user service, and overall knowledge management impact.


Knowledge Management Impact Metrics
Knowledge Management Metrics
Resource Coverage
Metadata Enrichment
Knowledge Graph Linking
Information Accessibility
Long-Term Retention
96
89
84
92
78
Value (%)
Impact Value (%)

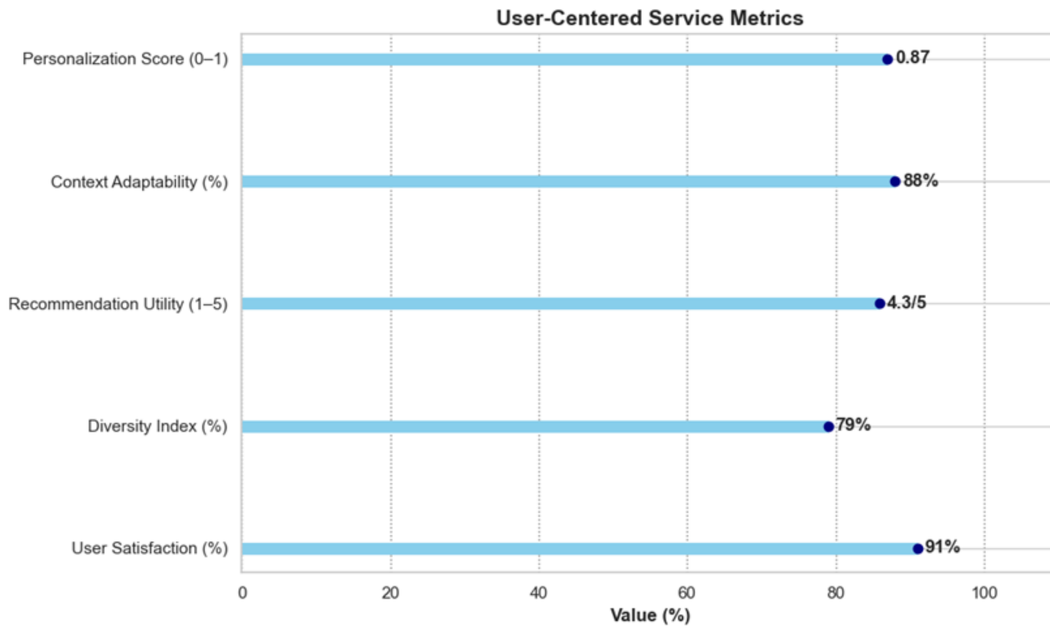
(a)


Resource Accessibility Performance
Values (%)
94.8
87.2
90.6
95.1
2.3
Coverage of Library Collections
Cross-Format Retrieval Success
Multilingual Accessibility
Accessibility Compliance Score
Broken Link/Error Rate
Resource Accessibility Metrics

(b)


Retrieval Efficiency Performance
Values
180
0.35
520
62
71
Query Response Time (ms)
Indexing Latency (s/doc)
Scalability (docs/sec)
Memory Utilization (%)
CPU/GPU Utilization (%)
Retrieval Efficiency Metrics

(c)



(d)

Figure 14: Performance Metrics of Library (a) Knowledge Management System (b) Accessibility (c) Retrieval Efficiency (d) User Service, and Impact.

Discussion

Knowledge management in AL aims to enhance information retrieval efficiency while delivering services tailored to user needs. The proposed framework incorporates advanced retrieval methods, intelligent systems, and user engagement strategies to improve the accessibility, organisation, and utilisation of information resources. This ensures seamless discovery, greater accessibility, and personalised experiences for diverse academic communities. However, existing models applied in AL knowledge management present several limitations that hinder optimal information retrieval and service delivery. For instance, although Semantic Retrieval BERT-KG effectively captures contextual meaning, it is computationally demanding and resource-intensive, which limits its scalability for large datasets.

Similarly, SVM lacks flexibility in handling multilingual, multifaceted, and dynamic library data, resulting in reduced retrieval precision. RF, while more robust, is prone to overfitting when dealing with diverse and imbalanced datasets, which negatively affects recall. LR, being relatively simplistic, struggles to model the nonlinear relationships characteristic of dynamic library retrieval systems. Collectively, these models face challenges in addressing cross-format, multilingual, and real-time retrieval requirements, and they often fail to achieve an effective balance between efficiency, accuracy, and user satisfaction. To address these shortcomings, the IMB-Res-BiGRU approach is adopted. This method offers greater adaptability for multilingual and cross-format data, improves accuracy, recall, and retrieval efficiency, and reduces computational complexity, thereby enabling academic libraries to scale effectively while providing intelligent and user-focused retrieval services.

Conclusion

Knowledge management in AL is centred on the systematic organisation, sharing, and utilisation of information to improve retrieval efficiency and provide user-focused services. It incorporates advanced technologies, optimised processes, and collaborative practices to strengthen accessibility, accuracy, and responsiveness within information services. The AL dataset utilised in this study included bilingual queries and user feedback. Pre-processing involved text cleaning and normalisation, while entity extraction was conducted through a voting-based NER and WSD approach to ensure accurate entity recognition and disambiguation. Information retrieval was carried out using BERT to generate precise and contextually relevant document recommendations. To enhance efficiency, the proposed IMB-Res-BiGRU model was applied, which significantly improved retrieval accuracy and personalisation of services for both English and Chinese academic library resources. The model achieved outstanding results, with 98.5% accuracy, a 98.1% F1-score, 97.8% precision, and 98.2% recall. These findings confirm the framework's effectiveness in enhancing retrieval performance and delivering reliable, user-centred information services across multilingual datasets, surpassing conventional models in both accuracy and retrieval outcomes. Nevertheless, certain limitations remain, including dependency on the quality of input data, challenges in integrating with existing library systems, and restricted adaptability to rapidly evolving technologies. Future research directions include strengthening multilingual support, embedding more advanced AI-driven personalisation, improving scalability, and ensuring seamless interoperability across platforms to maximise retrieval efficiency and deliver more user-centred, accessible, and responsive academic library services.

Reference

Ademilua, O. S. and Yacob, H. (2024). Promoting competitive intelligence strategies to advance library services to rural dwellers. *Indonesian Journal of Multidiciplinary Research*, 4(2): 275-296. https://doi.org/10.17509/ijomr.v4i2.75300

Ademola, R., Kuranga, B. M. and Rasaq, M. O. (2025). Availability and Use of ICT Tools for Effective Library Services Delivery in Academic Libraries in Kwara State. *Journal of Applied Information Science and Technology*, 18(1): 13-22. https://doi.org/10.70118/jaist.202501801.2

Akanbiemu, A. A. (2024). The Integration of Big Data Analytics in Knowledge Management: Implications for Libraries. *Niger Delta Journal of Library and Information Science*, 5(1): 86-108. https://doi.org/10.5281/zenodo.13293311

Alzahrani, S., Shakir, K., Yazeed, A., Salauddin, N. and Ahmed, S. (2025). Navigating the technological frontier: Tracing the journey from library 2.0 to library 5.0 and its impact on knowledge repositories. *Information Development*, 02666669251348907. https://doi.org/10.1177/02666669251348907

Awogbami, M. O. (2024). Professional competencies for digital reference services in academic libraries. *Journal of Library Services and Technologies*, 6(3): 121-132. http://doi.org/10.47524/jlst.v6i3.122

Boateng, F. (2025). The transformative potential of Generative AI in academic library access services: Opportunities and challenges. *Information Services and Use*, 45(1-2): 140-147. https://doi.org/10.1177/18758789251332800

Chukwujindu, N. B., Achugbue, E. I. and Ogbomo, E. F. (2024). Librarians' Attitudes, Work Environment and Service Delivery in Academic Libraries in Delta State, Nigeria. *Library and Information Perspectives and Research*, 6(2): 32-45. https://doi.org/10.47524/lipr.v6i2.33

Dhage, K. J. and Verma, C. B. (2025). AI-Driven Library Systems: Transforming Information Seeking Behaviour and Satisfaction in the Digital Era. *International Journal of Research in Library Science*, 11(2): 1906. https://doi.org/10.26761/IJRLS.11.2.2025.1906

Faizan, M. and Munshi, S. A. (2025). Enhancing academic access: evaluating integrated ICT-based library services at IIT from users' perspectives. *Performance Measurement and Metrics*, 26(1): 49-70. https://doi.org/10.1108/PMM-07-2024-0032

Hamad, F., Al-Fadel, M. and Fakhouri, H. (2023a). The provision of smart service at academic libraries and associated challenges. *Journal of Librarianship and Information Science*, 55(4): 960-971. https://doi.org/10.1177/09610006221114173

Hamad, F., Al-Fadel, M. and Shehata, A. M. K. (2023b). The level of digital competencies for the provision of smart information service at academic libraries in Jordan. *Global Knowledge, Memory and Communication*, 73(4-5): 614-633. https://doi.org/10.1108/GKMC-06-2022-0131

Ikwuanusi, U. F., Adepoju, P. A. and Odionu, C. S. (2023). AI-driven solutions for personalized knowledge dissemination and inclusive library user experiences. *International Journal of Engineering Research Updates*, 4(2): 052-062. https://doi.org/10.53430/ijeru.2023.4.2.0023

Ikwuanusi, U. F., Onunka, O., Owoade, S. J. and Uzoka, A. (2024). Revolutionizing Library Systems with Advanced Automation: A Blueprint for Efficiency in Academic Resource Management. *International Journal of Scholarly Research* *in Multidisciplinary Studies*, 5(2): 019-040. https://doi.org/10.56781/ijsrms.2024.5.2.0045

Jain, S. J. and Behera, P. K. (2023). Visualizing the Academic Library of the Future Based on Collections, Spaces, Technologies, and Services. *International Journal of Information Science & Management*, 21(1): 217-241. https://doi.org/10.22034/ijism.2023.700794

Kaggle. (n.d.). *Academic Libraries Dataset*. https://www.kaggle.com/datasets/ziya07/academic-libraries-dataset/data/data

Kim, J. (2025). Academic Library with Generative AI: From Passive Information Providers to Proactive Knowledge Facilitators. *Publications*, 13(3): 37. https://doi.org/10.3390/publications13030037

Li, C. (2025). Application effect of improved retrieval system based on user understanding on library management system. *Intelligent Decision Technologies*, 18724981251353207. https://doi.org/10.1177/18724981251353207

Mammadov, S. and Kucukkulahli, E. (2025). A User-Centric Smart Library System: IoT-Driven Environmental Monitoring and ML-Based Optimization with Future Fog–Cloud Architecture. *Applied Sciences*, 15(7): 3792. https://doi.org/10.3390/app15073792

Mbua, P. (2025). Applying Service Quality Theoretical Framework for Quality Service Delivery in Academic Libraries from an African Library Perspective. In: D. Ocholla, O. B. Onyancha, & A. O. Adesina (Eds.), *Information, Knowledge, and Technology for Teaching and Research in Africa: Information Behavior in Knowledge and Economy*. Springer Nature Switzerland, pp. 109-130. https://doi.org/10.1007/978-3-031-65745-0_5

Meesad, P. and Mingkhwan, A. (2024a). Knowledge Graphs in Smart Digital Libraries. In: P. Meesad & A. Mingkhwan (Eds.), *Libraries in Transformation: Navigating to AI-Powered Libraries*. Springer Nature Switzerland, pp. 327-389. https://doi.org/10.1007/978-3-031-69216-1_9

Meesad, P. and Mingkhwan, A. (2024b). User Experience and Engagement in Smart Digital Libraries. In: P. Meesad & A. Mingkhwan (Eds.), *Libraries in Transformation: Navigating to AI-Powered Libraries*. Springer Nature Switzerland, pp. 273-326. https://doi.org/10.1007/978-3-031-69216-1_8

Mosha, N. F. (2025). The role of artificial intelligence tools in enhancing accessibility and usability of electronic resources in academic libraries. *Library Management*, 46(1-2): 132-157. https://doi.org/10.1108/LM-08-2024-0088

Muslim, S. A. (2024). Assessing the Impact of Digital Transformation on Access, User Experience, and Knowledge Management in Academic Libraries. *Journal of Asian Multicultural Research for Educational Study*, 5(3): 15-23. https://doi.org/10.47616/jamres.v5i3.547

Obuh, L. and Ogbomo, E. F. (2024). Librarians' skills and application of web technologies for information service delivery in university libraries in Delta State. *Journal of Library Services and Technologies*, 6(2): 39-54. https://doi.org/10.47524/jlst.v6i2.40

Ogungbenro, O. D., Esse, U. C., Olowoporoku, I. and Christopher, A. (2025). Revolutionizing Library Services: The Impact of Artificial Intelligence on Cataloguing and Access to Information in Nigeria Academic Libraries. *Journal of Library Metadata*, 25(2): 99-118. https://doi.org/10.1080/19386389.2025.2475418

Oyedokun, T. T. (2025). Navigating the dynamics of present-day academic libraries: An in-depth analysis of strategies, challenges, and emerging trends. *IFLA Journal*, 51(2): 470-489. https://doi.org/10.1177/03400352241291907

Saadati, P., Pricope, E. and Abdelnour-Nocera, J. (2021). User Engagement and Collaboration in the Next Generation Academic Libraries. In: *34th British HCI Conference*. BCS Learning & Development, pp. 343-347. https://doi.org/10.14236/ewic/HCI2021.37

Veerakannan, S. (2025). AI in Libraries: Revolutionizing Access to Digital Knowledge. *Academic Research Journal of Science and Technology (ARJST)*, 1(05): 1-13. https://publications.ngmc.ac.in/journal/index.php/arjst/article/view/50

Yusof, S. K. and Sulaiman, M. A. H. (2025). A bibliometric analysis of the evolution and impact of reference services and information literacy programs in academic libraries (2014-2024). *Journal of Information and Knowledge Management (JIKM)*, 15(SI2): 77-93. https://doi.org/10.24191/jikm.v15iSI2

Zhang, S., Zhang, T. and Wang, X. (2025). AHP-Based Evaluation of Discipline-Specific Information Services in Academic Libraries Under Digital Intelligence. *Information*, 16(3): 245. https://doi.org/10.3390/info16030245

Zhou, J. and Huang, T. (2023). Application of machine learning algorithm in electronic book database management system. *SN Applied Sciences*, 5(11): 287. https://doi.org/10.1007/s42452-023-05508-3

Jin Hongyin is employed at Chulalongkorn University and holds a Ph.D. in Educational Administration. Her primary research area focuses on the innovative application of artificial intelligence technology in educational management, and she is dedicated to exploring and realizing the intelligent transformation of educational management. She has published several academic papers in international SSCI and Scopus journals.