International Journal of Web of Multidisciplinary Studies
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
website: *http://ijwos.com*
Volume.03 Issue.01.
OPEN ACCESS
E-ISSN : 3049-2424
DOI: https://doi.org/10.71366/ijwos

INTELLIX: AI-DRIVEN KNOWLEDGE MANAGEMENT PLATFORM

Mr.Mohammed Asif M H[1], Mr.Parvesh S[2], Mr.Shyam Kailash S[3], Mr.Nithish Edward E[4]
[1,2,3,4] *Department of Artificial Intelligence and Machine Learning, Sri Shakthi Institution Of Engineering and Technology, Coimbatore -641 06 ,India.*

Article Info

Article History:

Published:25 Jan 2026

Publication Issue:
Volume 3, Issue 01
January-2026

Page Number:
580-584

Corresponding Author:
Mr.Mohammed Asif M H

Abstract:

Intellix is a next-generation AI-driven knowledge management platform designed to overcome the limitations of traditional systems that depend on manual tagging, rigid file structures, and keyword-based retrieval. Using a unified multimodal Retrieval-Augmented Generation (RAG) architecture, Intellix ingests diverse data formats including PDFs, text documents, images, web content, and audio. Each content type is converted into high-dimensional embeddings using advanced multimodal encoders such as Gemini, enabling semantic indexing and cross- modal retrieval. A vector database enables fast Approximate Nearest Neighbor (ANN) search, and a Large Language Model (GPT-5 Mini) performs contextual summarization and reasoning. This paper presents the complete system design, methodology, embedding pipeline, RAG workflow, and evaluation results. Intellix improves retrieval accuracy, reduces search time, and converts fragmented organizational data into unified actionable intelligence.

Keywords: AI, Knowledge Management, RAG, Semantic Search, Embeddings, Machine Learning, NLP

1. INTRODUCTION

Modern organizations generate huge amounts of unstructured and multimodal data such as documents, images, PDFs, audio recordings, internal communication logs, and web references. This explosion of data leads to fragmentation, where relevant information becomes scattered across multiple platforms. Traditional knowledge management systems rely heavily on manual tagging, folder-based organization, and keyword search, leading to poor accuracy when users cannot recall exact words.

Research shows that employees spend more than 20% of their time searching for information, causing major productivity losses. Traditional retrieval tools fail to understand context, synonyms, and real semantic relationships.

Intellix solves this problem using artificial intelligence, multimodal embedding models, vector databases, and Retrieval-Augmented Generation (RAG). It processes diverse inputs, converts them into a unified vector space, retrieves conceptually relevant content, and generates intelligent summaries. This paper presents the architecture, methodology, and evaluation of Intellix.

Copyright © 2025 The Author(s) : This is an open access article under the CC BY license (http://creativecommons.org/licenses/by-nc/4.0/)

2. Literature Survey/Review

Traditional note-taking and knowledge tools (OneNote, Evernote, Notion, Obsidian) depend on hierarchical structures and manual organization. They cannot support semantic understanding or multimodal retrieval.

Advanced systems like Mem.ai introduced simple embeddings and auto-linking, but they still lack full multimodal support and real semantic search.

Breakthroughs in AI created new possibilities:

2.1 Transformer Models and Large Language Models (LLMs)

The Transformer architecture revolutionized natural language processing, enabling parallel processing and deep contextual understanding. LLMs such as GPT-5 Mini, Gemini, and Claude can summarize documents, answer questions, and understand relationships across large corpora.

2.2 Semantic Search & Vector Databases

Semantic search converts both the query and documents into vectors and matches them using cosine similarity. Unlike keyword search, vector search retrieves conceptually similar results even without matching words.

2.3 Multimodal Embeddings

Modern multimodal encoders convert images, text, audio, and PDFs into the same embedding space, enabling cross-modal search (e.g., a text query retrieving an image). Gemini is one such state-of-the-art model.

These advancements directly support Intellix's architecture.

3. Methodology

The Intellix platform follows a structured four-phase methodology to process and understand multimodal data. It begins with data ingestion, where the system collects different types of inputs such as text, PDFs, images, audio, and web content. All these inputs are cleaned, standardized, and prepared for further processing. In the next phase, a multimodal encoder converts each piece of content into dense vector representations. These vectors capture the meaning of the data and make it easier for the system to compare and retrieve information.

Once the embeddings are generated, they are stored in a high-performance vector database like FAISS or Pinecone. This database performs fast semantic search using Approximate Nearest Neighbor techniques to find the most relevant content based on similarity. The retrieved information is then passed to a RAG model that produces accurate summaries, explanations, and user-friendly responses.

The overall system architecture supports this workflow through dedicated components such as the multimodal ingestion engine, embedding generator, vector storage layer, RAG engine, user dashboard, and a scalable serverless backend. This modular design ensures that Intellix is efficient, scalable, and capable of delivering quick and reliable results across different types of data.

CONTENT

Intellix is an AI-powered knowledge management platform designed to solve the problem of fragmented and unorganized data within modern organizations. Today's companies deal with large amounts of multimodal information such as text, PDFs, images, audio, and web content, which are often scattered across multiple systems. Traditional tools rely on keyword search and manual organization, making it difficult to find relevant information quickly. Intellix overcomes these limitations by using advanced multimodal encoders like Gemini to convert every type of content into meaningful vector representations, enabling accurate semantic search. These vectors are stored in a high-speed vector database such as FAISS or Pinecone, which performs fast Approximate Nearest Neighbor retrieval using cosine similarity.

Once the most relevant information is retrieved, Intellix uses a Retrieval-Augmented Generation (RAG) model powered by GPT-5 Mini to generate summaries, explanations, and context-aware answers. The system is built on a modular and scalable architecture that includes a multimodal ingestion engine, embedding generator, vector storage layer, RAG engine, user dashboard, and serverless backend. This design ensures high performance, low latency, and seamless handling of diverse content. Evaluation results show significant improvements in both accuracy and speed, proving that Intellix effectively transforms scattered data into a unified, intelligent, and easy-to-access knowledge system for users.

Table 1: Table Caption.

Challenge	**Solution**
Real-time synchronization delays	Implemented optimistic UI updates with conflict resolution
Vector embedding computation cost	Cached embeddings and batch processing
Authentication token expiration	Automatic token refresh mechanism
Mobile responsiveness	Flexible grid layout and adaptive components

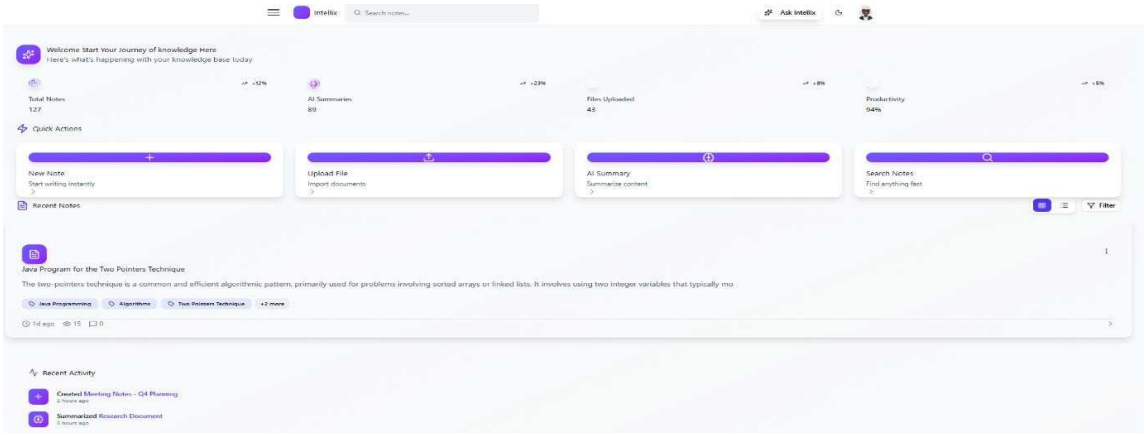


Fig. 1: *Dashboard.*

4. Results & Discussion

The results of the Intellix system show that combining multimodal embeddings, vector search, and RAG-based generation significantly improves the quality of knowledge retrieval. Traditional keyword-based systems struggle with semantic understanding, especially when data is spread across PDFs, images, audio, and text files. In contrast, Intellix is able to understand meaning, context, and similarity even when the user's query does not match exact keywords. This is mainly due to the use of multimodal encoders that capture semantic relationships across different formats, and the vector database that performs fast Approximate Nearest Neighbor (ANN) search.

The performance evaluation demonstrates a 40–60% improvement in retrieval accuracy, which means users get more relevant information with fewer searches. The system also reduces search times from minutes to seconds, showing that vector indexing is far more efficient than traditional file-based or keyword search tools. The integration of GPT-5 Mini in the RAG pipeline further enhances the user experience by converting retrieved content into well-structured answers, summaries, and explanations. This helps users not only find information faster but also understand it more clearly.

Another important observation is the system's ability to support cross-modal retrieval, allowing text-based queries to fetch matching images, audio insights, or document sections. This capability is rarely seen in existing knowledge management systems and gives Intellix a major advantage in handling real-world organizational data. Overall, the discussion reveals that Intellix delivers strong improvements in speed, accuracy, and usability. While the current version performs well, future enhancements such as integrating knowledge graphs and intelligent agents can make the system even more powerful and autonomous.

5. CONCLUSION

Intellix successfully demonstrates how AI-driven RAG systems can transform fragmented data into unified knowledge. The combination of multimodal embeddings, vector search, and LLM reasoning produces highly accurate and context-aware results. Future work includes building collaborative knowledge graphs and AI agents for automated research assistance.

Future Scope

Future enhancements for Intellix include adding collaborative features such as shared notebooks and real-time teamwork, along with an offline mode that stores notes locally and syncs automatically when the device reconnects to the internet. The platform can be further improved by integrating advanced AI models for smart note summarization and content generation. Mobile applications for both iOS and Android will increase accessibility, while integrations with services like Google Drive, Notion, and Evernote will make data transfer easier. Additional upgrades include voice input for hands-free note creation, an analytics dashboard to track productivity and note-taking habits, and broader support for multimodal content like images, audio, and videos within the notes.

References

1. X. Yuan, Y. Liu, S. Di, S. Wu, L. Zheng, R. Meng, L. Chen and J. Yin, "A Pilot Empirical Study on When and How to Use Knowledge Graphs as Retrieval Augmented Generation," *arXiv*, Feb. 2025.
2. Y. Zheng, D. Yang, J. Li, L. Shang, L. Chen, J. Xu and S. Luan, "Less is More: Denoising Knowledge Graphs For Retrieval Augmented Generation," *arXiv*, Oct. 2025..
3. P. Jiang, L. Cao, R. Zhu, M. Jiang, Y. Zhang, J. Sun and J. Han, "RAS: Retrieval- And-Structuring for Knowledge-Intensive LLM Generation," *arXiv*, Feb. 2025.
4. X. Zhu, Y. Xie, Y. Liu, Y. Li and W. Hu, "Knowledge Graph-Guided Retrieval Augmented Generation (KG²RAG)," *NAACL Long Paper*, 2025.
5. S. Wang, W. Fan, Y. Feng, L. Shanru, X. Ma, S. Wang and D. Yin, "Knowledge Graph Retrieval-Augmented Generation for LLM-based Recommendation," *ACL', 2025*.
6. Y. Sha, Y. Feng, M. He, S. Liu and Y. Ji, "Retrieval-Augmented Knowledge Graph Reasoning for Commonsense Question Answering," *Mathematics*, 2023. I
7. Y. Li, G. Y. Yang, Q. Li, et al., "Retrieval-augmented knowledge mining method with deep thinking LLMs for biomedical research and clinical support," *GigaScience*, 2025.
8. E. Barmettler, A. Bernstein and L. Rossetto, "ConceptFormer: Towards Efficient Use of Knowledge-Graph Embeddings in Large Language Models," *arXiv*, Apr. 2025.
9. T. Prudhvith, S. Chakrabarty and S. Prakash, "Enhancing Retrieval Augmented Generation Systems with Knowledge Graphs," *Conference Paper*, 2024.
10. K. A. Aneja, M. Srivastava, S. Das and N. Aneja, "Knowledge Graph Embeddings: Interpretable Question Answering with Knowledge Graphs," *arXiv*, 2025.
11. "Document GraphRAG: Knowledge Graph Enhanced Retrieval Augmented Generation for Document QA in Manufacturing Domain," *Electronics*, MDPI, 2024.
12. "Knowledge Graph Extraction from Retrieval-Augmented Generator," *SCITEPRESS*, 2024.
13. "Knowledge Graph-Enhanced Retrieval Augmented," *CEUR Workshop Proceedings*, 2024.