Satyam Kumar Tripathi, 2026, 14:2
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

RecallX: An Intelligent AI-Based Multimodal Memory Recording and Retrieval System

Satyam Kumar Tripathi[1], Raj Singh[2], Jyoti Maddheshiya[3], Ritika Singh[4], Mr. Sanjeev Pathak[5]

[1,2,3,4]CSE AIML, [5]Lecturer of AIML, Bansal Institute of Engineering and Technology, Lucknow, India

Abstract- People today face an overwhelming volume of daily information, which often causes cognitive overload and makes it hard to manage memories effectively. Standard tools like basic note-taking apps or digital reminders usually fall short because they lack contextual awareness and have limited search capabilities. To address this, we introduce RecallX, an AI-driven memory assistant built to capture, organize, and retrieve various types of data—including text, audio, images, and video. By combining Natural Language Processing (NLP), computer vision, and speech recognition, RecallX moves past basic file storage. Instead, it structures memories with contextual awareness and creates associative links, similar to how human memory works. The system pulls out key details, tags them with metadata like timeframes and related entities, and allows users to search their memories naturally using everyday language. By relying on semantic search rather than strict keyword matching, RecallX delivers much more accurate and relevant results. Ultimately, this scalable architecture is built to lighten the user's mental workload and boost daily productivity.

Keywords: Artificial Intelligence, NLP, Memory Retrieval, Multimodal Systems, Semantic Search

I. INTRODUCTION

We are constantly bombarded with information from meetings, casual conversations, daily tasks, and fleeting ideas. While human memory is remarkable, it naturally struggles to hold onto detailed, context-heavy information over long periods. As a result, we easily forget important details or waste time trying to recall past knowledge.

Current digital aids like task managers, reminders, and note apps only offer a partial fix. They usually demand manual data entry, completely miss the context of the information, and force users to rely on exact keyword searches.

If a user forgets the specific word they used in a note, finding it becomes incredibly frustrating. On top of this, our data is often scattered across half a dozen different apps, creating a fragmented digital footprint that only adds to our mental clutter.

We urgently need a smarter system: one that can automatically capture our data, organize it logically, and let us retrieve it naturally with full context awareness.

Objectives

The main goal of RecallX is to create a smart memory assistant that effortlessly records and retrieves multi-modal data. We want the system to replicate how human memory works by grasping the context and the connections between different pieces of stored information. Specifically, our objectives are to:

- Build a framework that seamlessly captures varied inputs like text, audio, images, and video.
- Apply Natural Language Processing (NLP) techniques to pull meaningful insights and contextual clues from whatever the user inputs.
- Store memories dynamically by tagging them with useful metadata, such as timestamps, involved entities, and relationships.
- Develop a semantic search mechanism that allows users to retrieve information using natural language queries.
- Design an architecture that scales smoothly as the user's data grows.
- Reduce cognitive load and improve productivity by pro- viding intelligent memory assistance.

II. RELATED WORK

Many platforms and studies have tackled the challenge of storing and finding information. Standard apps like Notion or Google Keep do a fine job of holding text, but they lean heavily on users typing everything out manually and don't genuinely understand what the text means. On the other hand, voice assistants like Google Assistant and Amazon Alexa let us speak naturally, yet they are built mostly to execute immediate commands rather than serving as long-term, context-aware memory banks.

Meanwhile, recent breakthroughs in NLP have transformed how machines read text. Models like BERT and Sentence- BERT are excellent at grasping the underlying meaning and similarities in language, paving the way for smarter search engines. Likewise, multi-modal models like CLIP are successfully bridging the gap between images and text. Despite these incredible tools, there is still a lack of a unified system that brings everything together to capture and link multi-modal memories. RecallX fills this gap. By blending computer vision, speech recognition, and NLP into one cohesive platform, it acts as a "digital brain." It doesn't just archive data; it interconnects it, making it easily searchable through casual, natural language.

III. PROPOSED SYSTEM

The proposed system, RecallX, is an intelligent AI-based memory assistant designed to capture, process, and retrieve multimodal data efficiently. The system mimics human memory by understanding the context and relationships between the stored information.

System Overview

The system consists of multiple modules:

- Input Module
- Processing Module
- AI Analysis Module
- Storage Module
- Retrieval Module

System Architecture

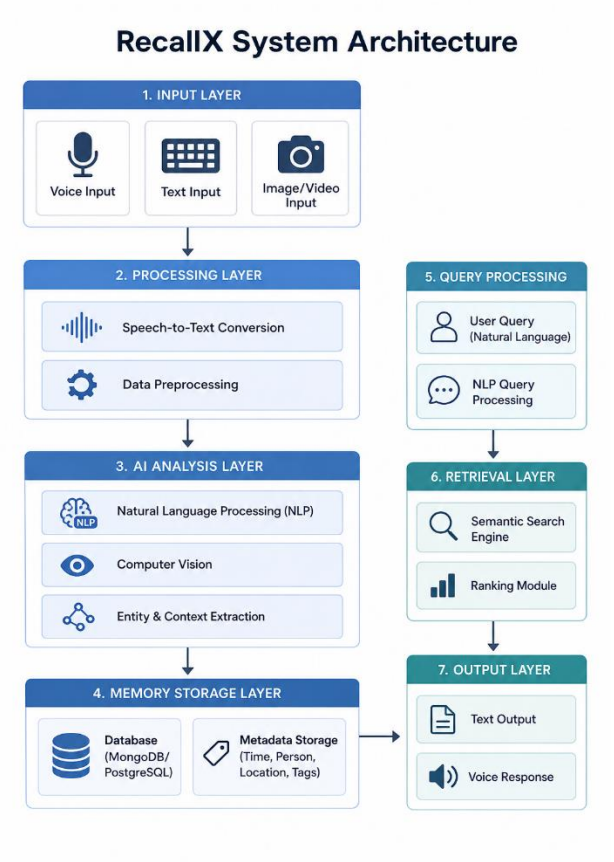

Fig. 1. RecallX System Architecture

The architecture shows how input data flows through various modules, where it is processed, analyzed, stored, and retrieved.

Working Principle

The system works in two phases:

- Memory Creation Phase
- Capture user input
- Convert speech to text
- Extract context using NLP
- Store data with metadata

Memory Retrieval Phase

- The user enters the query
- NLP processes query
- The semantic search retrieves relevant data

IV. NATURAL LANGUAGE PROCESSING (NLP)

NLP is the engine that allows RecallX to truly comprehend and interact with everyday human language.

NLP Pipeline

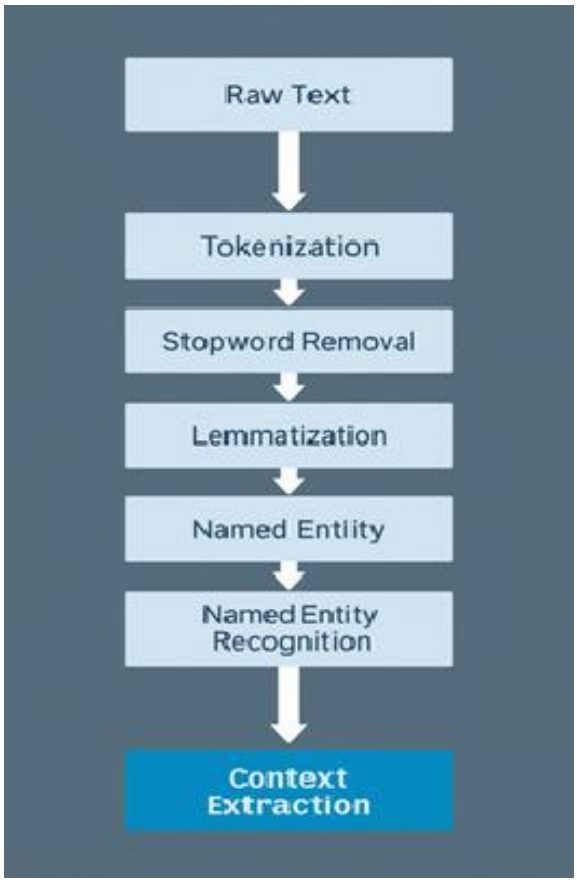

Fig. 2 NLP Processing Pipeline

Detailed NLP Working

Tokenization: Breaking down sentences into individual words.

Stop word Removal: Common words are removed.

Lemmatization: Words are converted to base form.

Named Entity Recognition: Identifies names, dates, and locations.

Context Extraction: Mapping out the relationships and underlying meanings connecting the words.

Semantic Search

$$Similarity = \frac{Q \cdot M}{\|Q\|\|M\|}$$

Importance of NLP

- Improves search accuracy
- Understands user intent
- Enables natural interaction

V. WORKING METHODOLOGY

The RecallX system follows a structured workflow to capture, process, store, and retrieve user data efficiently. The methodology is designed to simulate human memory by understanding the context and relationships between different pieces of information.

The overall functioning of the system can be divided into multiple stages as described below.

Input Acquisition

The first stage involves capturing user input in various forms, such as text, voice, images, or video. Voice input is converted into text using speech recognition techniques, while image and video data are processed using computer vision methods.

Data Preprocessing

Once the input is captured, it undergoes preprocessing. This includes cleaning the data, removing noise, and converting them into a structured format suitable for further processing. For text data, preprocessing includes tokenization, stop word removal, and normalization.

Feature Extraction and Context Tagging

In this stage, Natural Language Processing techniques are applied to extract meaningful features such as keywords, entities, and relationships. Important contextual information such as time, location, and participants is identified and tagged. This step is crucial to enable intelligent memory retrieval.

Data Storage

The processed data are stored in a database along with meta- data. Each memory entry is indexed using tags and timestamps to allow efficient retrieval. The system uses structured and flexible storage mechanisms to handle different types of data.

Query Processing
When a user submits a query, it is processed using NLP techniques to understand the intent behind the query. The query is converted into a vector representation for semantic comparison.

Semantic Retrieval
The system performs a semantic search by comparing the query vector with stored memory vectors. Instead of relying on exact keyword matching, the system identifies the most relevant results based on meaning and context.

Output Generation
Finally, the retrieved information is presented to the user in a readable format. The output can be displayed as text or delivered through voice response.

VI. ALGORITHMS

The RecallX system uses two primary algorithms: the Memory Processing Algorithm and the Semantic Search Algorithm.

Memory Processing Algorithm
The memory processing algorithm is responsible for capturing and storing user data efficiently.
Algorithm Steps:
1. Capture input from user (text, voice, image, video)
2. If input is voice, convert it into text
3. Perform data preprocessing (cleaning, normalization)
4. Apply NLP techniques for feature extraction
5. Identify entities such as names, dates, and locations
6. Generate contextual metadata (time, tags, relationships)
7. Store processed data in database

Mathematical Representation

$$M = f(I, C)$$

Where I represents the input data and C represents the extracted context.

Semantic Search Algorithm
The semantic search algorithm is used to retrieve relevant information based on user queries.
Algorithm Steps
1. Accept user query in natural language
2. Preprocess query using NLP techniques
3. Convert query into vector representation
4. Compare the query vector with the stored memory vectors
5. Compute a similarity score using cosine similarity
6. Rank results based on similarity score
7. Return relevant top results

Similarity Calculation

$$\text{Similarity} = \frac{Q \cdot M}{\|Q\|\|M\|}$$

Where Q is the query vector and M represents the stored memory vectors

Algorithm Efficiency
The use of semantic search significantly improves retrieval accuracy compared to traditional keyword-based systems. It allows the system to understand the intent of the query and retrieve contextually relevant information, even when exact keywords are not present.

In Addition, indexing and metadata tagging improve the speed and efficiency of the retrieval process, making the system suitable for real-time applications.

VII. EXPERIMENTAL RESULTS AND ANALYSIS

The performance of the proposed RecallX system is evaluated on the basis of its ability to retrieve relevant information accurately and efficiently. The system is compared with traditional keyword-based search methods.

Evaluation Metrics
The following metrics are used to evaluate system performance:
- **Accuracy:** Measures the accuracy of retrieved results.

- **Precision:** Indicates how many retrieved results are relevant.
- **Recall:** Measures how many relevant results are success- fully retrieved.
- **F1-score:** Harmonic mean of precision and recall.
- **Response Time:** Time taken to retrieve results.

Performance Comparison

TABLE I: PERFORMANCE COMPARISON BETWEEN TRADITIONAL SEARCH AND RECALLX

Method	Accuracy	Precision	Recall	F1	Time(MS)
Keyword Search	72%	70%	68%	69%	120
RecallX	90%	88%	87%	87.5%	80

VIII. LIMITATIONS

Data Privacy Concerns
Since RecallX stores personal and sensitive information, there are potential privacy risks. Proper security measures such as encryption and access control are required to protect user data.

High Storage Requirements
The system stores large volumes of multimodal data, which can lead to increased storage requirements. Managing and scaling storage efficiently is a significant challenge.

Computational Complexity
The use of advanced AI techniques, such as NLP and semantic search, increases computational cost. Real-time processing may require high-performance hardware and optimized algorithms.

Result Analysis
The results show that RecallX significantly outperforms traditional keyword-based systems. The use of semantic search and contextual understanding enables higher accuracy and better recall.

The reduction in response time is achieved due to efficient indexing and optimized retrieval algorithms. The system is able to understand the intent of the user, which leads to more relevant results even when exact keywords are not provided.

IX. ADVANTAGES

Context-Aware Retrieval
One of the major advantages of RecallX is its ability to understand context. Unlike traditional systems that rely on exact keyword matching, RecallX analyzes relationships between entities such as people, time, and location. This allows the system to retrieve information even when the query is vague or incomplete. As a result, users can interact naturally without remembering exact details.

Reduced Cognitive Load
RecallX significantly reduces the mental effort required to manage and recall information. Users do not need to manually organize or remember large amounts of data. The system automatically captures, processes, and stores information in an organized manner. This improves productivity and helps users focus on more important tasks.

Multimodal Data Processing
The system supports multiple input formats including text, voice, images, and video. This flexibility allows users to interact with the system in a natural and convenient way. By combining different types of data, RecallX provides a more comprehensive and intelligent memory system compared to traditional tools.

X.CONCLUSION

In this paper, RecallX, an intelligent AI-based multimodal memory record and retrieval system, has been presented. The system addresses the limitations of traditional memory management tools by integrating advanced technologies such as Natural Language Processing, Speech Recognition, and Computer Vision. Unlike conventional systems, RecallX pro- vides context-aware memory storage and semantic retrieval, enabling users to access information using natural language queries.

The proposed system successfully demonstrates how multimodal data can be captured, processed,

and organized efficiently. The use of semantic search improves retrieval accuracy and relevance, while contextual tagging enhances the overall user experience. Experimental results indicate that RecallX outperforms traditional keyword-based systems in terms of accuracy, response time, and efficiency.

In general, RecallX acts as a digital brain that bridges the gap between human memory limitations and artificial intelligence capabilities, thus improving productivity and reducing cognitive load.

Future Work

Although RecallX provides an effective solution for intelligent memory management, several enhancements can be made in the future to further improve its capabilities.

- **Emotion Detection:** Future versions of the system can incorporate emotion recognition to understand the emotional context of memories, making retrieval more personalized.
- **Integration with Wearable Devices:** RecallX can be extended to integrate with smart devices such as smart- watches and AR glasses for real-time data capture.
- **Offline AI Models:** Implementing offline capabilities will allow the system to function without internet dependency, improving accessibility and privacy.
- **Real-Time Conversation Tracking:** The system can be enhanced to capture and summarize live conversations automatically.
- **Advanced Knowledge Graphs:** Future improvements can include building knowledge graphs to better represent relationships between stored memories.

REFERENCES

1. J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre- training of Deep Bidirectional Transformers for Language Understand- ing," NAACL, 2019.
2. T. Brown et al., "Language Models are Few-Shot Learners," NeurIPS, 2020.
3. N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," EMNLP, 2019.
4. A. Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," ICML, 2021.
5. T. Baltrus˘aitis, C. Ahuja, and L.-P. Morency, "Multimodal Machine Learning: A Survey and Taxonomy," IEEE TPAMI, 2019.
6. A. Graves, A.-r. Mohamed, and G. Hinton, "Speech Recognition with Deep Recurrent Neural Networks," ICASSP, 2013.
7. A. Vaswani et al., "Attention Is All You Need," NeurIPS, 2017.
8. S. Guo, X. Sun, Y. Wang, and J. Zhang, "Deep Learning for Semantic Matching in Information Retrieval," ACM Computing Surveys, 2020.
9. C. D. Manning, P. Raghavan, and H. Schu¨tze, "Introduction to Informa- tion Retrieval," Cambridge University Press, 2008.
10. R. Szeliski, "Computer Vision: Algorithms and Applications," Springer, 2022.
11. D. Jurafsky and J. H. Martin, "Speech and Language Processing,"Pearson, 2021.
12. S. Russell and P. Norvig, "Artificial Intelligence: A Modern Approach,"Pearson, 2020.
13. P. Lewis et al., "Retrieval-Augmented Generation for Knowledge- Intensive NLP Tasks," NeurIPS, 2020.
14. V. Karpukhin et al., "Dense Passage Retrieval for Open-Domain Ques- tion Answering," EMNLP, 2020.
15. O. Khattab and M. Zaharia, "ColBERT: Efficient and Effective Passage Search via Contextualized Late Interaction over BERT," SIGIR, 2020.
16. J. Li et al., "Align Before Fuse: Vision and Language Representation Learning with Momentum Distillation," NeurIPS, 2021.
17. I. Beltagy, M. Peters, and A. Cohan, "Longformer: The Long-Document Transformer," ACL, 2020.