MemoriEase 4.0: A Lifelog QA System Powered by Reasoning LLMs

Quang-Linh Tran
Dublin City University
Dublin, Dublin, Ireland
quang-linh.tran2@mail.dcu.ie

Binh Nguyen
Ho Chi Minh city University of Science, Vietnam National University
Ho Chi Minh City, Vietnam
ngtbinh@hcmus.edu.vn

Gareth J. F. Jones
Maynooth University
Dublin, Dublin, Ireland
gareth.jones@mu.ie

Cathal Gurrin
Dublin City University
Dublin, Dublin, Ireland
cgurrin@computing.dcu.ie

Abstract

Lifelog Question-Answering (QA) has become an important application of lifelogging, answering questions posed over personal lifelog data to help lifeloggers gain insight into their lives, from lifestyle analytics to past-event reflection and behavioural patterns. Building on our previous work on lifelog retrieval, we present MemoriEase 4.0, our system for the Lifelog Search Challenge 2026 (LSC'26), which targets the QA task using reasoning large language models (LLMs). The system adopts a Retrieval-Augmented Generation (RAG) approach to answer questions. For complex questions that require aggregating information, such as counting or comparison across time, reasoning LLMs analyse the retrieved evidence step by step before producing an answer. At the live competition, MemoriEase 4.0 finished third overall among fifteen teams and recorded the strongest ad-hoc retrieval performance. Our analysis further identifies answer latency and aggregation accuracy as the principal challenges that remain for competitive lifelog QA. Beyond the challenge, the system aims to evolve into a complete assistant for everyday lifelog utilisation.

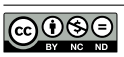
CCS Concepts

• **Information systems → Users and interactive retrieval**.

Keywords

Lifelog Retrieval, Large Language Models, Question-Answering

ACM Reference Format:
Quang-Linh Tran, Binh Nguyen, Gareth J. F. Jones, and Cathal Gurrin. 2026. MemoriEase 4.0: A Lifelog QA System Powered by Reasoning LLMs. In *The 9th Annual ACM Workshop on the Lifelog Search Challenge (LSC '26), June 16–19, 2026, Amsterdam, Netherlands.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3810984.3817080



LSC '26, Amsterdam, Netherlands

ACM ISBN 979-8-4007-2697-2/26/06
https://doi.org/10.1145/3810984.3817080

1 Introduction

Lifelogging is the passive collection and storing of lifelog data. Lifelog data includes Point-of-View (PoV) images, health metrics, location, et cetera [6]. The people who collect the data are called lifeloggers. Wearable cameras are used to capture PoV images to store what lifeloggers see. Smartwatches can be used to collect health metrics and location data. Phones and computers can be used to collect emails, e-commerce shopping, and internet activities. From all the data sources combined, the lifelog data reflect the lives of lifeloggers. Many applications utilise the lifelog data to create value, from health monitoring to activity recommendation, and especially retrieval [3, 4, 10]. Lifelog data can be used to monitor daily activities of patients, from eating and drinking water to exercising and provide comprehensive details to doctors about patients' conditions. In addition, based on recorded past activities, recommender systems can suggest exercise activities like walking or jogging to avoid sitting too long. Some smartwatches already utilise users' health metrics to plan a personalised training program, like running with a specific pace and duration, to optimise for health improvement. These applications show a big potential of lifelog data in various domains, but lifelog retrieval is still one of the most potential to use lifelog in real life.

Lifelog retrieval is one of the most popular tasks in lifelogging, and has significant potential to improve our memory through digital storage [19]. In general, this task aims to find relevant lifelog data from a query. For example, with the query "Find images of my last trip to Vietnam, when I was on the beach.", the retrieval system tries to find all relevant PoV images with the beach scene and ensures that the location is in Vietnam. This task helps lifeloggers to remember their experience, find lost items by tracing back the lifelog, or serve as evidence of doing something at a time. Thanks to its huge potential, there are several challenges for lifelog retrieval, like NTCIR Lifelog [32], ImageCLEFlifelog [15], and Lifelog Search Challenge (LSC) [7, 8], the longest-running interactive benchmark for lifelog retrieval.

The Lifelog Search Challenge (LSC) has become a primary benchmark for evaluating interactive retrieval systems. This challenge provides a large-scale lifelog dataset and participating systems to find the answer as fast as possible to gain a high score. Within this challenge, researchers focus on three main sub-tasks: Known-Item Search (KIS), which targets a specific, unique moment; Ad-hoc

Search (AD), which aims to retrieve various relevant results for a broader query; and Question-Answering (QA), which requires generating precise textual answers from the data. For KIS and QA sub-task, participants are given 300 seconds to find the answer to the query/question. For AD sub-task, participants need to find as many as relevant images in the time frame of 180 seconds. For the KIS sub-task, there are 6 hints in total and released every 30 seconds, while the other two sub-tasks have only 1 completed query. The scoring function of LSC considers both the time to correct submissions and the number of wrong submissions to create a truthful evaluation.

One of the main challenges of lifelog retrieval is the large-scale, multimodal, and highly personal nature of lifelog data. A typical lifelog dataset may contain hundreds of thousands of images along with heterogeneous metadata such as time, location, and semantic concepts, making efficient and accurate retrieval difficult. Moreover, queries are often expressed in natural language, containing implicit temporal, spatial, and semantic constraints, which require the system to understand both visual content and contextual information. The LSC'26 dataset [21] contains approximately 18 months of rich multimodal lifelog data, including wearable camera images, biometrics, activities, and locations, consistent with the dataset used in LSC'25. More information about the dataset will be provided later.

Due to these challenges, lifelog retrieval systems increasingly adopt embedding-based approaches using vision-language models to map both queries and images into a shared representation space, enabling more effective semantic matching. In our previous work, MemoriEase [24, 25, 27], we proposed a conversational lifelog retrieval system that integrates semantic embedding-based retrieval with a Retrieval-Augmented Generation (RAG) framework [12] to support both search and question-answering tasks. The system leverages vision-language models such as CLIP [18] and BLIP2 [13] for effective text-to-image semantic embedding-based retrieval. Additionally, MemoriEase 3.0 [24] incorporates a conversational user interface that allows users to interact naturally with their lifelog data. To address complex queries, particularly in the QA task, we introduced a retriever–reranker–reader pipeline, where relevant lifelog moments are first retrieved, then reranked using a cross-encoder model, and finally passed to a large language model to generate answers. This approach demonstrated a top-3 performance at LSC'25 [8], showing the effectiveness of combining multimodal retrieval with LLM-based reasoning for lifelog understanding.

In this year's challenge, we present MemoriEase 4.0, an enhanced version of our system with a particular focus on improving performance for retrieval with new embedding models and handling aggregation questions in the QA sub-task. We introduce SigLIP2 [28] into our embedding-based retrieval to enhance the retrieval performance. We also extend our previous RAG pipeline by incorporating reasoning-capable LLMs to better handle complex queries that require counting, comparison, and multi-event reasoning across time. Despite these enhancements, we retain the core strengths of our system, including the conversational search paradigm and main user interface interaction, allowing users to interact naturally with their lifelog while benefiting from more advanced reasoning and retrieval capabilities under the hood.

2 Related Works

LSC has attracted numerous participants from all over the world in previous challenges. In this section, we review notable participating systems from the last challenge to illustrate work on interactive lifelog retrieval.

Embedding-based retrieval remains a dominant paradigm among LSC'25 participants. lifeXplore 2025 [17] investigates whether search over text descriptions generated by ChatGPT and BLIP2 outperforms direct embedding-based retrieval using OpenCLIP, concluding that embeddings consistently yield superior results, with an OpenCLIP ViT-H-14 model trained on the laion2b dataset achieving the best performance, especially when queries are refined by removing unusable information and applying temporal filters. SnapSeek 3.0 [9] extends the embedding-based approach with enhanced scene understanding, incorporating scene graph generation, object relationships, spatial context, and Activities of Daily Living (ADL) annotations, while also introducing an interactive feedback mechanism and flexible temporal reasoning framework to support more complex and context-aware queries.

Several systems take a structured or knowledge-driven approach to retrieval. LifeGraph 5 [20] presents the fifth iteration of a multimodal knowledge graph for lifelog retrieval, introducing a novel SPARQL-powered user interface built on the MediaGraph store (MeGraS). By elevating multimedia documents to first-class citizens within the knowledge graph, the system enables complex queries with similarity-based search and near-duplicate detection using native SPARQL syntax. MEMORIA [5], which achieved the best overall score at LSC'25, takes a comprehensive annotation pipeline approach, processing images through an ensemble of models including PaddleOCR, YOLOv11, BLIP2, ClipCap, and Places365 to produce rich OCR, object, captioning, and scene annotations stored in PostgreSQL, augmented with an automatic query parser leveraging NLP and LLM techniques.

Temporal and sequential reasoning is another key theme at LSC'25. T-EAGLE [14] focuses on capturing temporal narratives by combining sequence captioning with text matching, enabling the system to reason across ordered sequences of lifelog moments rather than treating each image independently. MyEachtraX [22] is a mobile-first QA-oriented system that centres the question-answering process around early, high-quality event retrieval using semantic and time-weighted fusion, prioritising temporal context for accurate query resolution. LifeSearch [16] proposes a multimodal approach that integrates heterogeneous lifelog modalities to support expressive search across the dataset.

The other systems at LSC'25 explore novel interaction paradigms. LUMINA-1 [11] introduces a virtual reality workspace for lifelog exploration, supporting gaze interaction, hand gesture detection, and speech-to-text input to enable a fully immersive and locomotion-free retrieval experience, with strong results on atomic QA subtasks. VitaChronicle 2.0 [31] focuses on user experience design with a dual-mode interface catering separately to novice and expert users, applying established UX/UI principles to lower the barrier to entry for interactive lifelog retrieval. U-Cker [29] is a prototype system exploring foundational interactive retrieval capabilities within the LSC framework. Across all participating systems, a clear trend emerges: the combination of powerful visual-language embeddings,

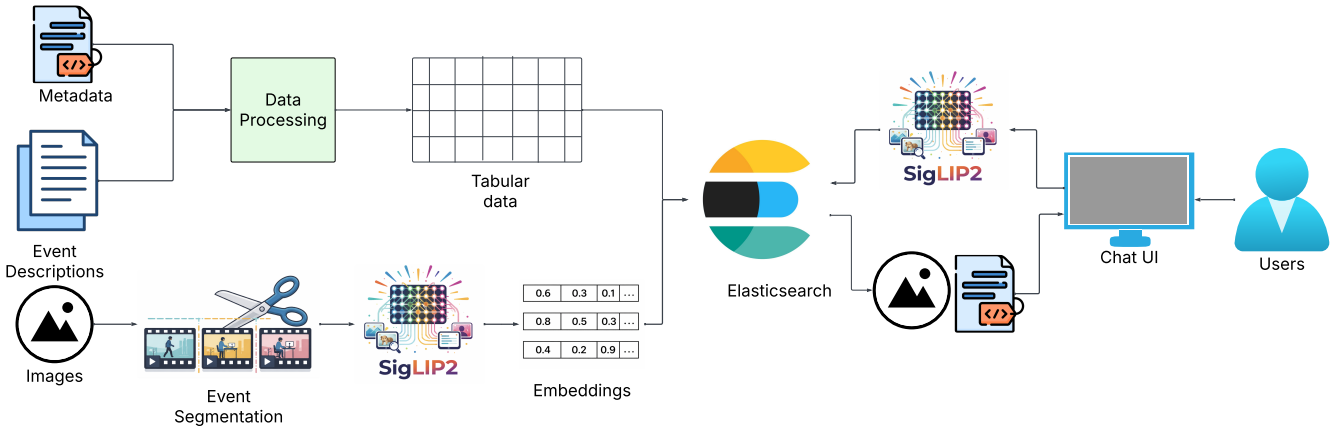


Figure 1: Overall architecture of MemoriEase 4.0

rich multimodal annotations, and LLM-assisted reasoning is becoming the standard architecture for high-performing lifelog retrieval systems, which motivates the improvements we introduce in MemoriEase 4.0.

3 MemoriEase 4.0

In this section, we present the MemoriEase 4.0 interactive lifelog retrieval system. As stated in section 2, we will only present new enhancements and developments of MemoriEase in this fourth version. All other parts remain the same, with details that can be found in the MemoriEase 3.0 paper [24]. Figure 1 illustrates the high-level architecture of the system. Overall, images are grouped into events based on temporal and spatial metadata, and keyframes for each event are embedded into embedding vectors by SigLIP2. The metadata and event descriptions, generated by a captioning model, are processed into tabular data and indexed into Elasticsearch along with the embeddings. More details about the data processing and indexing stage can be found in section 3.1. The retrieval mechanism, with different types of search like conversational search and visual similarity search, uses queries from users to encode to SigLIP2 embeddings before searching on Elasticsearch. The returned results are processed and displayed to users in the interface. Information about the retrieval and interface is provided in section 3.2 and 3.4. The new enhancements in QA for addressing aggregation questions are presented in section 3.3.

3.1 Data Processing and Indexing

The LSC'26 dataset contains approximately 18 months of continuous lifelog data, consistent with the LSC'25 dataset, including wearable camera images, biometrics, activities, and location data.

3.1.1 Description Generation. For each event, we generate a short description using a Vision Language Model. We choose Qwen3-VL 4B Instruct [2] to generate the description, thanks to its SOTA performance in text-image understanding tasks. The description is used for the QA section, which will be discussed later. The images representative of each event are embedded into a vector embedding of 1152 dimensions.

3.1.2 Metadata Extraction. For each image, we extract and process a comprehensive set of metadata fields to serve as filters during retrieval. Specifically, we extract temporal metadata including local time, hour of day, day of the week, weekend flag, and time period (morning, afternoon, evening, night), as well as spatial metadata including city and semantic location name. The semantic location name is from [23]. Biometric signals and activity labels available in the dataset are also indexed as additional metadata fields to support queries involving health or activity context. We also implement an automatic filter extraction from the query, which will automatically extract filters such as hour, day, time period, location, etc. It is a new improvement compared to our last rule-based filter extraction module. Finally, all metadata fields, image embeddings, and textual descriptions are indexed into an Elasticsearch vector database, where metadata fields serve as pre-filters applied before embedding-based similarity computation to narrow the candidate pool and improve both retrieval speed and accuracy. The embedding is stored as a dense vector with cosine similarity as the pre-defined similarity score. The indexed data is ready for retrieval, which is presented in the following section.

3.2 Retrieval

MemoriEase 4.0 supports two retrieval modes: text-to-image retrieval and image-to-image retrieval. These modes can be combined to support the diverse query types encountered in the KIS, AD, and QA sub-tasks. There is not much news in this retrieval section compared to last year's system.

3.2.1 Text-to-Image Retrieval. Text-to-image retrieval is the primary retrieval mode, used for both the KIS and AD tasks. Given a natural language query, the system first applies a query parser to extract structured filters and a filter-free semantic query. The filter-free query is then encoded into an embedding vector using SigLIP2 [28]. The query embedding is compared against pre-indexed image embeddings in the Elasticsearch database using cosine similarity, and a ranked list of results is returned to the user along with the corresponding metadata.

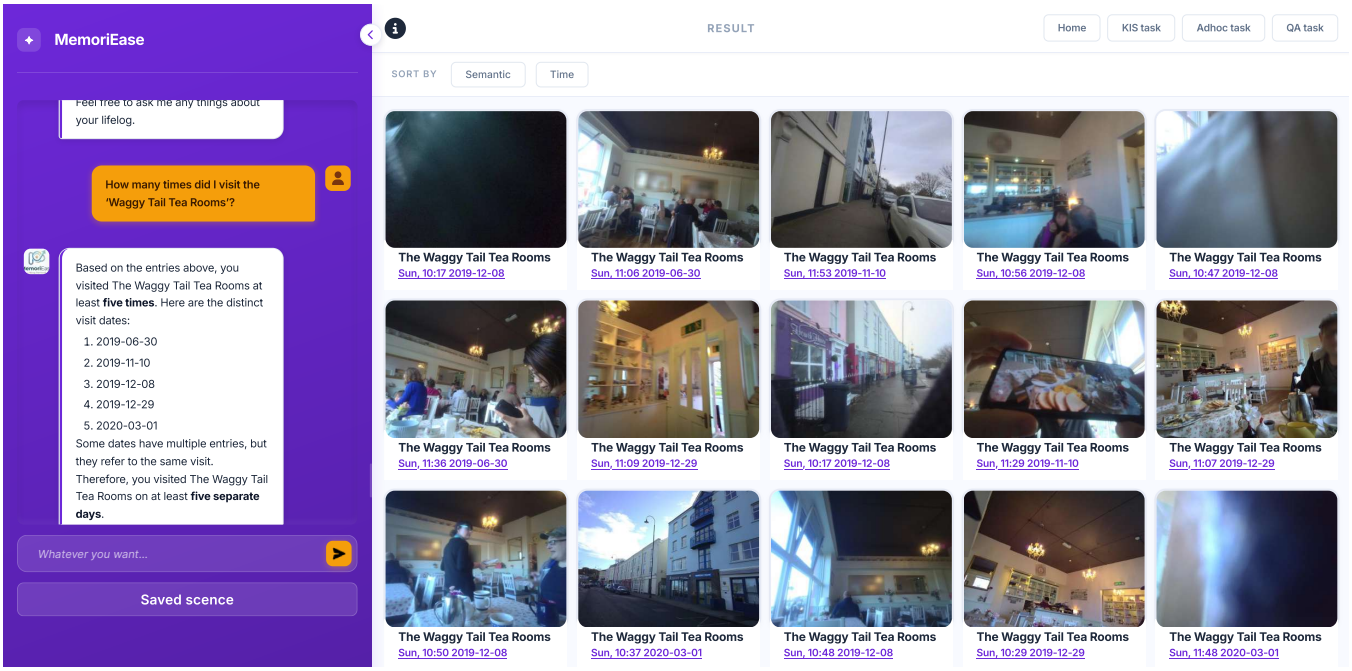


Figure 2: Conversational search interface

3.2.2 Image-to-Image Retrieval. Image-to-image retrieval supports the AD task through a relevance feedback mechanism. After an initial text-to-image search, users can select one or more relevant images from the result list and add them to a pool. The pooled images are encoded into SigLIP2 embeddings, which are then averaged and combined with the original query embedding using a weighted sum to balance the semantic intent of the original query with the visual characteristics of the selected images. This enhanced embedding is used to retrieve a new ranked list of visually and semantically similar images from the database.

3.3 QA

The Question-Answering (QA) sub-task requires the system to generate precise textual answers from lifelog data in response to natural language questions. Questions in the QA sub-task can range from simple factual lookups (*e.g.*, "What did I eat for lunch on Monday?") to complex aggregation queries (*e.g.*, "How many times did I visit a gym last month?") that require counting, comparison, or multi-event reasoning across time. Building on the retriever–reranker–reader pipeline introduced in MemoriEase 3.0 [24] and the RAG pipeline from [26], we retain the same architectural backbone but replace the general-purpose reader LLM with a reasoning LLM. The core pipeline operates as follows. Given a natural language question, the system first retrieves the top-k most relevant events from Elasticsearch using the SigLIP2 query embedding combined with metadata pre-filters. The retrieved events, represented by their textual descriptions generated by Qwen3-VL 4B Instruct [2], are then reranked using a cross-encoder model to ensure the most contextually relevant events are prioritised. Finally, the top reranked events and their associated metadata are passed as context to the reader LLM, which generates a free-text answer grounded in the retrieved lifelog evidence. We also integrate flight data from [1] for the QA.

Simple factual questions can be answered by the LLMs directly from the retrieved event descriptions. However, aggregation questions that require counting distinct events, comparing quantities across time intervals, or reasoning over sequences of activities pose a greater challenge, as they require the model to perform structured logical inference rather than simple extraction. To handle such queries, MemoriEase 4.0 incorporates reasoning-capable LLMs for a dedicated aggregation reasoning pathway. When the system classifies a question as an aggregation query, it retrieves a broader set of candidate events and their structured metadata from Elasticsearch. Rather than passing unstructured text directly to the reader, the system constructs a structured prompt that presents the relevant event records as a tabular or list-based context. The reasoning LLM then performs step-by-step analysis over this structured context, decomposing the question into intermediate reasoning steps before producing the final answer. This chain-of-thought approach [30] significantly improves accuracy on aggregation tasks by reducing hallucination and ensuring that numerical or comparative answers are logically derived from the evidence.

3.4 User Interface

The MemoriEase 4.0 interface is redesigned with a focus on accessibility and ease of use for new users while retaining the full feature set required for competitive retrieval.

The conversational search page, shown in Figure 2, follows a split-panel layout divided into a left conversational panel and a

right result gallery. This layout allows users to simultaneously issue queries and inspect retrieved results without switching views, which is critical under the time constraints of the LSC competition. The left panel presents a chat-style interface built on a deep purple background. Users interact with the system through a natural language input field at the bottom, labelled *"Whatever you want..."*, with an amber send button. The conversation history is displayed as a sequence of message bubbles: system messages appear in soft lavender bubbles, while user messages are rendered in amber bubbles, creating a clear visual distinction between the user and the assistant. The right panel displays the retrieval results as a responsive five-column image grid. Each result card shows the retrieved PoV image alongside a timestamp and, where available, a semantic location label extracted from the metadata. Interactive action controls are overlaid on each card, including a bookmark button to save results, a folder button to group cards, a selection checkmark to submit the image, and a magnifying glass to view the image at a larger size. Results can be sorted either by semantic relevance score or by temporal order using the toggle controls at the top of the panel. A task-switching navigation bar at the top right allows users to switch between the Home, KIS, AD, and QA sub-tasks without leaving the current search context.

4 Results and Analysis

We report the official performance of MemoriEase 4.0 at the LSC'26 competition and analyse the system behaviour using the submission logs exported from the evaluation server. The competition with 15 teams comprised 21 official tasks split across the three sub-tasks: 6 KIS, 6 AD, and 9 QA tasks. To obtain a single overall ranking, each team's average score within a sub-task is normalised against the best-performing team in that sub-task, which is scaled to 1000 points; the three normalised sub-task scores are then summed to give a final score out of a maximum of 3000. Where a team operated with more than one expert (31 experts in total), only the best-performing run is retained for that team.

4.1 Overall Performance

Table 1 presents the final standings of all 15 teams. MemoriEase 4.0 finished in third place with a normalised total of 2545.8 out of 3000, behind MyEachtra (2714.4) and VISIONE (2706.7). The margin separating the top three teams is only 168.6 points (5.6% of the maximum). This result matches the top-3 finish of MemoriEase 3.0 at LSC'25, demonstrating that the system continues to be competitive at interactive lifelog retrieval.

4.2 Performance by Sub-Task

While the overall standing is strong, the per-sub-task breakdown in Figure 3 reveals uneven performance. MemoriEase 4.0 was the best system in the field at Ad-hoc Search, ranked third at KIS, but fell to eleventh at the QA sub-task. However, this is the sub-task targeted by this year's enhancements.

4.2.1 Ad-hoc Search (1st of 15). MemoriEase 4.0 obtained the highest Ad-hoc Search score in the competition. The system returned a correct relevant item for all 6 of 6 AD tasks, and did so at a median time-to-first-correct of only 23 seconds, roughly twice as fast as the overall median of 46 seconds (Figure 4). Because AD rewards as many as correct submissions, this result demonstrates that the SigLIP2-based embedding retrieval introduced in MemoriEase 4.0 produces a dense, high-recall ranking of visually and semantically relevant moments: the system repeatedly surfaced many correct items per task thanks to visual similarity search within the short 180-second window. Embedding-based visual retrieval is the strongest component of the system.

4.2.2 Known-Item Search (3rd of 15). Performance on Known-Item Search was also strong, with the system solving 5 of 6 tasks and ranking third behind VISIONE and MyEachtra. The median time-to-first-correct of 114 seconds was marginally faster than the field median of 122 seconds. The single unsolved KIS task (LSC2622) received three incorrect submissions, suggesting a retrieval miss rather than a submission-strategy error. Overall, the KIS results reinforce the conclusion from AD that the upgraded embedding

Table 1: Official LSC'26 final standings. MemoriEase 4.0 is highlighted.

#	Team	KIS	AD	QA	Total
1	MyEachtra	951.7	831.8	930.9	2714.4
2	VISIONE	1000.0	880.7	826.0	2706.7
3	**MemoriEase**	**922.0**	**1000.0**	**623.8**	**2545.8**
4	ExploringLife	689.6	921.9	901.9	2513.4
5	SnapSeek	813.8	838.9	724.6	2377.3
6	FOCAL	712.2	849.9	661.9	2224.0
7	LifeGraph	903.9	549.1	763.9	2216.9
8	lifeXplore	501.4	603.0	784.7	1889.1
9	EyeSpy	724.4	530.2	609.6	1864.2
10	FirstGen	765.9	695.2	400.8	1861.8
11	Lumina-HT	604.2	338.8	771.5	1714.5
12	MEMORIA	742.7	387.7	491.9	1622.3
13	ChronoSeek	553.6	54.5	1000.0	1608.1
14	U-Cker	722.4	319.5	550.2	1592.1
15	Vita Chronicle	674.0	236.2	633.7	1543.8

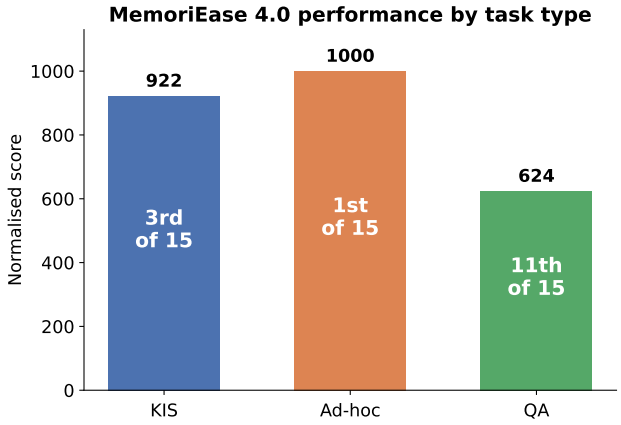


Figure 3: MemoriEase 4.0 normalised score and rank.

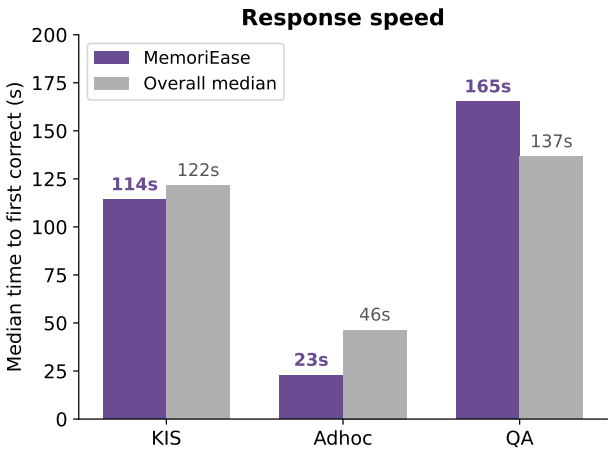


Figure 4: Submission behaviour of MemoriEase 4.0 across sub-tasks.

retrieval and metadata pre-filtering pipeline is highly effective for visually grounded search.

4.2.3 QA (11th of 15). In contrast to the visual-retrieval sub-tasks, QA was the weakest aspect of MemoriEase 4.0, despite being the headline focus of this version. The system solved only 6 of 9 QA tasks and was the slowest where it did succeed, with a median time-to-first-correct of 165 seconds against a field median of 137 seconds (Figure 4). The resulting normalised QA score of 623.8 placed the system eleventh of fifteen, well behind the category leader ChronoSeek and the overall winner MyEachtra (930.9). This is the clearest limitation in our results and merits a candid analysis, which we provide below.

4.3 Discussion: Strengths and Limitations

Taken together, the results tell a coherent story. The strengths of MemoriEase 4.0 lie in embedding-based visual retrieval. The migration to SigLIP2 embeddings, combined with automatic filter extraction and metadata pre-filtering, yields a retrieval stage that is both fast and precise: the system was the quickest in the field on Ad-hoc Search, near the front on Known-Item Search, and consistently maintained a low wrong-submission rate on both. For the search sub-tasks that have historically defined the LSC, MemoriEase 4.0 is a top-tier system, and its third-place overall finish is driven almost entirely by this retrieval quality, with the Ad-hoc Search win contributing a full 1000 points to the total. However, for QA task where the metadata is limited, MemoriEase failed to find the correct answer for questions.

The principal limitation is that the QA enhancements that motivated this version did not translate into competitive QA performance. Inspecting the submission logs against the task ground truth reveals three recurring failure modes. The most damaging is incomplete recall on aggregation questions. On task LSC2603 ("How many intercontinental flights did I take in 2019?", ground truth 12), the system probed answers of 5, 6, 8, and 10 as it accumulated evidence but never reached the full count, because the reasoning model can only count the events that retrieval surfaces; counting accuracy is therefore bounded by retrieval completeness rather than by the reasoning step itself. A second failure mode is temporal disambiguation due to lack of metadata in the QA task. In the unsolved KIS task LSC2622 ("I stopped at the petrol station in the evening"), the system retrieved visually correct petrol-station-at-evening images but submitted instances from 2020 rather than the target date in January 2019. Third, highly specific compound queries caused retrieval misses, as on LSC2605, where a single wrong-event answer was submitted nine months away from the target date. Compounding these accuracy issues, the LSC QA sub-task is scored on time-to-first-correct in the same manner as Known-Item Search, so the latency of the retriever–reranker–reader pipeline and the chain-of-thought aggregation pathway directly depresses the score even when the final answer is eventually correct; our 165-second median QA response time, slower than the field median, reflects this overhead.

These findings directly shape our future work. The retrieval core is already best-in-class and need not be the priority; instead, the QA pipeline should be re-engineered for latency, for example by caching event descriptions, pruning the reranking stage for simple factual questions, and reserving the reasoning pathway only for genuine aggregation queries. Improving the grounding of the reader model to reduce unfaithful answers, and decoupling aggregation accuracy from perfect retrieval recall, are the two most promising directions for closing the QA gap while preserving the strong retrieval performance that already places MemoriEase among the leading lifelog retrieval systems.

5 Conclusion

In this paper, we presented MemoriEase 4.0, an enhanced conversational lifelog retrieval system designed to address the scale and complexity of the LSC'26 dataset. Building on the strong foundation of MemoriEase 3.0, our system introduces several key improvements. We integrate SigLIP2 into our embedding-based retrieval pipeline to improve semantic matching accuracy. We extend the RAG-based QA pipeline with reasoning-capable LLMs to better handle aggregation queries requiring counting, comparison, and multi-event reasoning across time. Despite the significant dataset, MemoriEase 4.0 retains the conversational search paradigm that allows users to interact naturally with their lifelog data, keeping the system accessible while delivering more advanced retrieval and reasoning capabilities under the hood. We achieved the third place in the live competition at LSC'26 with good experience in Ad-hoc search but also revealed the drawbacks of QA modules. We aim to further exploring the potential of combining multimodal retrieval with LLM-based reasoning for lifelog understanding.

Acknowledgments

This research was conducted with the financial support of Research Ireland at ADAPT, the Research Ireland Centre for AI-Driven Digital Content Technology at Dublin City University [13/RC/21-06_P2].

References

[1] Ahmed Alateeq, Mark Roantree, and Cathal Gurrin. 2023. Voxento 4.0: A More Flexible Visualisation and Control for Lifelogs. In *Proceedings of the 6th Annual ACM Lifelog Search Challenge* (Thessaloniki, Greece) *(LSC '23)*. Association for Computing Machinery, New York, NY, USA, 7–12. doi:10.1145/3592573.3593097

[2] Shuai Bai, Yuxuan Cai, Ruizhe Chen, Keqin Chen, Xionghui Chen, Zesen Cheng, Lianghao Deng, Wei Ding, Chang Gao, Chunjiang Ge, Wenbin Ge, Zhifang Guo, Qidong Huang, Jie Huang, Fei Huang, Binyuan Hui, Shutong Jiang, Zhaohai Li, Mingsheng Li, Mei Li, Kaixin Li, Zicheng Lin, Junyang Lin, Xuejing Liu, Jiawei Liu, Chenglong Liu, Yang Liu, Dayiheng Liu, Shixuan Liu, Dunjie Lu, Ruilin Luo, Chenxu Lv, Rui Men, Lingchen Meng, Xuancheng Ren, Xingzhang Ren, Sibo Song, Yuchong Sun, Jun Tang, Jianhong Tu, Jianqiang Wan, Peng Wang, Pengfei Wang, Qiuyue Wang, Yuxuan Wang, Tianbao Xie, Yiheng Xu, Haiyang Xu, Jin Xu, Zhibo Yang, Mingkun Yang, Jianxin Yang, An Yang, Bowen Yu, Fei Zhang, Hang Zhang, Xi Zhang, Bo Zheng, Humen Zhong, Jingren Zhou, Fan Zhou, Jing Zhou, Yuanzhi Zhu, and Ke Zhu. 2025. Qwen3-VL Technical Report. arXiv:2511.21631 [cs.CV] https://arxiv.org/abs/2511.21631

[3] Junho Choi, Chang Choi, Hoon Ko, and Pankoo Kim. 2016. Intelligent healthcare service using health lifelog analysis. *Journal of medical systems* 40 (2016), 1–10.

[4] Tilman Dingler, Passant El Agroudy, Rufat Rzayev, Lars Lischke, Tonja Machulla, and Albrecht Schmidt. 2021. Memory augmentation through lifelogging: opportunities and challenges. *Technology-augmented perception and cognition* (2021), 47–69.

[5] Alexandre Gago, Bernardo Kaluza, and António J. R. Neves. 2025. MEMORIA: A Memory Enhancement and MOment RetrIeval Application at the LSC2025. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge (LSC '25)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3729459.3748693

[6] Cathal Gurrin, Alan F Smeaton, Aiden R Doherty, et al. 2014. Lifelogging: Personal big data. *Foundations and Trends® in information retrieval* 8, 1 (2014), 1–125.

[7] Cathal Gurrin, Liting Zhou, Graham Healy, Werner Bailer, Duc-Tien Dang Nguyen, Steve Hodges, Björn Þór Jónsson, Jakub Lokoč, Luca Rossetto, Minh-Triet Tran, et al. 2024. Introduction to the seventh annual lifelog search challenge, lsc'24. In *Proceedings of the 2024 International Conference on Multimedia Retrieval*. 1334–1335.

[8] Cathal Gurrin, Liting Zhou, Graham Healy, Allie Tran, Luca Rossetto, Werner Bailer, Duc-Tien Dang-Nguyen, Steve Hodges, Björn Þór Jónsson, Minh-Triet Tran, et al. 2025. Introduction to the 8th Annual Lifelog Search Challenge, LSC'25. In *Proceedings of the 2025 International Conference on Multimedia Retrieval*. 2143–2144.

[9] Minh-Quan Ho-Le, Duy-Khang Ho, Van-Tu Ninh, Trong-Thuan Nguyen, Cathal Gurrin, and Minh-Triet Tran. 2025. SnapSeek 3.0: Interactive Lifelog Search System with Enhanced Scene Understanding. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge*. 7–14.

[10] Gunjan Kumar, Houssem Jerbi, Cathal Gurrin, and Michael P O'Mahony. 2014. Towards activity recommendation from lifelogs. In *Proceedings of the 16th international conference on information integration and web-based applications & services*. 87–96.

[11] Nhut-Thanh Le-Hinh, Cong-Triet Huynh, Minh-Quan Ho-Le, Duy-Khang Ho, Minh-Triet Tran, and Viet-Tham Huynh. 2025. LUMINA-1: Learning and Understanding Multimedia in Immersive Navigable Archives for Lifelog Retrieval. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge (LSC '25)*. Association for Computing Machinery, New York, NY, USA, 15–22. doi:10.1145/3729459.3748698

[12] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2021. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv:2005.11401 [cs.CL] https://arxiv.org/abs/2005.11401

[13] Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. 2023. BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models. arXiv:2301.12597 [cs.CV] https://arxiv.org/abs/2301.12597

[14] Thang-Long Nguyen-Ho, Allie Tran, Minh-Triet Tran, Cathal Gurrin, and Graham Healy. 2025. T-EAGLE: Capturing Temporal Narratives via Sequence Captioning and Text Matching. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge (LSC '25)*. Association for Computing Machinery, New York, NY, USA, 23–27. doi:10.1145/3729459.3748691

[15] Van-Tu Ninh, Tu-Khiem Le, Liting Zhou, Luca Piras, Michael Alexander Riegler, Pål Halvorsen, Mathias Lux, Minh-Triet Tran, Cathal Gurrin, and Duc Tien Dang Nguyen. 2020. Overview of imageclef lifelog 2020: lifelog moment retrieval and sport performance lifelog. CEUR Workshop Proceedings.

[16] Shriram Piramanayagam, Enting Chen, Andre Melo, Pavlos Vougiouklis, Chenxin Diao, Pascual Merita, Zeren Jiang, Ruofei Lai, and Jeff Z. Pan. 2025. LifeSearch: Multimodal Lifelog Search System for LSC'25. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge (LSC '25)*. Association for Computing Machinery, New York, NY, USA, 28–33. doi:10.1145/3729459.3748696

[17] Martin Rader and Klaus Schoeffmann. 2025. lifeXplore 2025: Is Search in Text Generated by ChatGPT and BLIP2 better than in Embeddings from OpenCLIP?. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge*. 34–38.

[18] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning Transferable Visual Models From Natural Language Supervision. arXiv:2103.00020 [cs.CV] https://arxiv.org/abs/2103.00020

[19] Ricardo Ribeiro, Alina Trifan, and António JR Neves. 2022. Lifelog retrieval from daily digital data: narrative review. *JMIR mHealth and uHealth* 10, 5 (2022), e30517.

[20] Florian Ruosch and Luca Rossetto. 2025. A SPARQL in the Dark: Shining a Light on Multimodal Lifelogs with LifeGraph 5. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge*. 39–43.

[21] Allie Tran, Werner Bailer, Duc-Tien Dang-Nguyen, Graham Healy, Steve Hodges, Björn Þór Jónsson, Wolfgang Hürst, Luca Rossetto, Klaus Schoeffmann, Minh-Triet Tran, Liting Zhou, and Cathal Gurrin. 2026. Introduction to the 9th Annual Lifelog Search Challenge, LSC'26. In *Proceedings of the 2026 International Conference on Multimedia Retrieval (ICMR '26)*. Association for Computing Machinery, New York, NY, USA, 2904–2905. doi:10.1145/3805622.3811234

[22] Allie Tran and Cathal Gurrin. 2025. MyEachtraX: Semantic Event Retrieval with Time-Weighted Fusion. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge (LSC '25)*. Association for Computing Machinery, New York, NY, USA, 44–51. doi:10.1145/3729459.3748690

[23] Ly-Duyen Tran, Dongyun Nie, Liting Zhou, Binh Nguyen, and Cathal Gurrin. 2023. VAISL: Visual-Aware Identification of Semantic Locations in Lifelog. In *MultiMedia Modeling*, Duc-Tien Dang-Nguyen, Cathal Gurrin, Martha Larson, Alan F. Smeaton, Stevan Rudinac, Minh-Son Dao, Christoph Trattner, and Phoebe Chen (Eds.). Springer Nature Switzerland, Cham, 659–670.

[24] Quang-Linh Tran, Binh Nguyen, Gareth JF Jones, and Cathal Gurrin. 2025. MemoriEase 3.0: A RAG-Enhanced Conversational Lifelog Retrieval System at LSC'25. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge*. 52–57.

[25] Quang-Linh Tran, Binh Nguyen, Gareth J. F. Jones, and Cathal Gurrin. 2024. MemoriEase 2.0: A Conversational Lifelog Retrieve System for LSC'24. In *Proceedings of the 7th Annual ACM Workshop on the Lifelog Search Challenge* (Phuket, Thailand) *(LSC '24)*. Association for Computing Machinery, New York, NY, USA, 12–17. doi:10.1145/3643489.3661114

[26] Quang-Linh Tran, Ngo Ngoc Diep Pham, Quoc Trung Truong, Minh Hung Nguyen, Hong Cat Le, Dang Khoi Vu, Van Minh Thien Nguyen, Van Kinh Nguyen, Luu Phuong Ngoc Lam Nguyen, Tan Le, Minh Phuc Dang, Binh Nguyen, Gareth J. F. Jones, and Cathal Gurrin. 2025. A RAG Approach for Multi-Modal Open-ended Lifelog Question-Answering. In *Proceedings of the 2025 International Conference on Multimedia Retrieval* (Chicago, IL, USA) *(ICMR '25)*. Association for Computing Machinery, New York, NY, USA, 1303–1312. doi:10.1145/3731715.3733263

[27] Quang-Linh Tran, Ly-Duyen Tran, Binh Nguyen, and Cathal Gurrin. 2023. MemoriEase: An Interactive Lifelog Retrieval System for LSC'23. In *Proceedings of the 6th Annual ACM Lifelog Search Challenge* (Thessaloniki, Greece) *(LSC '23)*. Association for Computing Machinery, New York, NY, USA, 30–35. doi:10.1145/3592573.3593101

[28] Michael Tschannen, Alexey Gritsenko, Xiao Wang, Muhammad Ferjad Naeem, Ibrahim Alabdulmohsin, Nikhil Parthasarathy, Talfan Evans, Lucas Beyer, Ye Xia, Basil Mustafa, et al. 2025. Siglip 2: Multilingual vision-language encoders with improved semantic understanding, localization, and dense features. *arXiv preprint arXiv:2502.14786* (2025).

[29] Kazuya Ueki, Ryo Muto, Takuya Wada, Ryota Akaba, and Genesis Faith Layag Fernandez. 2025. U-Cker: A Prototype System for the Lifelog Search Challenge 2025. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge (LSC '25)*. Association for Computing Machinery, New York, NY, USA, 58–62. doi:10.1145/3729459.3748688

[30] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc Le, and Denny Zhou. 2023. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. arXiv:2201.11903 [cs.CL] https://arxiv.org/abs/2201.11903

[31] Michela Wilhelmsen, Linnea Sannes Rønning, Duc-Tuan Luu, Minh-Quan Ho-Le, Thanh-Hai Tran, Cathal Gurrin, Minh-Triet Tran, and Duc Tien Dang Nguyen. 2025. VitaChronicle 2.0: A Dual-Mode UX-Centered Approach to Lifelog Image Retrieval for Novice and Expert Users. In *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge (LSC '25)*. Association for Computing Machinery, New York, NY, USA, 63–69. doi:10.1145/3729459.3748692

[32] Liting Zhou, Cathal Gurrin, Hsin-Hung Chen, Hideo Joho, Chenyang Lyu, Longyue Wang, Graham Healy, Ly Duyen Tran, Quang-Linh Tran, Hoang Bao Le, et al. 2025. Overview of the NTCIR-18 lifelog-6 task. In *Proceedings of the 18th NTCIR Conference on Evaluation of Information Access Technologies*.