MemoriEase 3.0: A RAG-Enhanced Conversational Lifelog Retrieval System at LSC'25

Quang-Linh Tran
ADAPT Centre, School of Computing
Dublin City University
Dublin, Ireland
linh.tran3@mail.dcu.ie

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

Gareth J F Jones
ADAPT Centre, School of Computing
Dublin City University
Dublin, Ireland
Gareth.Jones@computing.dcu.ie

Cathal Gurrin
ADAPT Centre, School of Computing
Dublin City University
Dublin, Ireland
cgurrin@computing.dcu.ie

Abstract

We present the third version of the MemoriEase lifelog retrieval system. This system is a conversational lifelog retrieval system built on an embedding-based retrieval method. For LSC'25, we enhance our system with a RAG approach to the question-answering task. In addition, we also incorporate two embedding models, CLIP and BLIP2, for the embedding-based retrieval. We improve our relevance feedback for visual similarity search by adjusting the query embedding. We describe the results of our system at the LSC'25 challenge, where it achieved third place overall and second place in the QA task. The enhancements in this version of the system improve our performance in the LSC'25 challenge.

CCS Concepts

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

Keywords

Lifelog Retrieval, Conversational Search, Retrieval-Augmented Generation

ACM Reference Format:
Quang-Linh Tran, Binh Nguyen, Gareth J F 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 (LSC '25), June 30–July 03, 2025, Chicago, IL, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3729459.3748689

1 Introduction

Lifelogging is the process of passively collecting daily life data through wearable devices [9]. The devices include wearable cameras, smartwatches, etc. Lifelog data is used in many applications, ranging from health monitoring [5] and lifestyle analysis [15] to memory enhancement [7]. With the aging population, the application of memory enhancement is becoming increasingly important. It can potentially provide help to elderly people or Alzheimer's patients enabling them to reminisce about memories of past events. This application area has several downstream tasks for lifelog data, but the most actively investigated is lifelog retrieval. The Lifelog search challenge (LSC) [8, 10, 11] is one of the longest-running benchmarks for the evaluation of interactive lifelog retrieval systems. This year marks the eighth iteration of the challenge, and this is also the third time our system, MemoriEase [30–33], has participated in this challenge. We bring several enhancements to our system for LSC'25 [11], mainly focusing on the question-answering (QA) task, since it remains significantly challenging for participants.

The dataset for LSC'25 remains unchanged since the previous challenge [10]. This dataset contains 18 months of lifelog data with 725K lifelog images and associated metadata, along with visual concepts. There are three tasks in the challenge, including known-item search (KIS), ad-hoc search (AD), and question-answering (QA). The KIS task involves querying a specific moment in lifelog data, where the system is considered correct if it retrieves at least one accurate lifelog image from that moment. In contrast, the AD task focuses on retrieving as many relevant results as possible for a broader, more general life moment, such as "finding times when I read a menu in a café.". Lastly, the QA task requires generating a textual answer to a question based on the lifelog data. For example, "How many times did I have a picnic by the sea?" is a challenging question. This requires the retrieval system to find all the moments of a picnic by the sea and an aggregation module to count the number of times. Our MemoriEase system focuses on this type of task for LSC'25 by incorporating a RAG approach with a retriever-reranked-reader pipeline to solve the QA task.

This is the third time we have participated in this challenge, so we already have a base retrieval system built on the Elasticsearch [1] database and BLIP2 embedding model [18]. However, to improve our system further, especially for the QA task, we have incorporated several enhancements. Firstly, we employ the CLIP embedding model [23], along with the BLIP2 model, to process the image and query to improve the search accuracy. Secondly, we improve our relevance feedback algorithm to incorporate original query embedding at the stage of recalculating embedding after

LSC '25, Chicago, IL, USA

ACM ISBN 979-8-4007-1857-1/25/06
https://doi.org/10.1145/3729459.3748689

[1] https://www.elastic.co/

feedback to retain important information from the query. Finally, we incorporate a new RAG approach [17] with a state-of-the-art GPT-o1 [2] and a high-performance reranking model to improve the accuracy of the system for the QA task. These enhancements are expected to improve our performance in the LSC'25 challenge.

2 Related Work

In this section, we provide details of research in the domain of lifelog retrieval, especially in recent interactive lifelog retrieval systems at the LSC in previous years [8, 10, 28]. The technique of lifelog retrieval has developed significantly, especially from the introduction of CLIP models [23], which bridge the gap between images and texts in zero-shot retrieval. After that, most of the systems now use the embedding-based retrieval approach, which involves using vision-language models such as CLIP [23], BLIP2 [18], OpenCLIP [4], etc, to encode lifelog images and textual queries into a shared-embedding space and calculating the similarity score. However, each team in the LSC'24 [10] had unique features such as data processing, user interface, and retrieval techniques. In total, 21 teams participated in LSC'24, and summary details of their systems are as follows.

LifeInsight 2.0 [36] employed an ensemble approach of CLIP and BLIP2 models. In addition, they incorporated a temporal query mechanism and an automatic query parser. Vitrivr-VR system [26] utilized virtual reality to construct a new search interface with easier and more flexible temporal and spatial query formulation. The EAGLE system [20] enhanced novice users' retrieval performance by incorporating implicit interactions in eye movements and automatic search flow. This is one of the first systems to try to utilize gaze detection in the retrieval process. Snapseek [13] focused on fast and accurate data retrieval by using the Milvus[3] database for indexing and 4 vision-language embedding models for embedding. This system had an innovative UI for lifelog exploration. VitaChronicle [21] applied UX/UI principles and guidelines to enhance the lifelog retrieval system. This constructed a novel user interface based on an analysis of the current lifelog retrieval systems. T@Retrospect [27] also prioritized UX and UI to address the needs of novice users. The LifeLens [34] 2.0 system improved its user interface to be more appealing and intuitive. The Vitrivr system [25] addressed the challenges of semi-structured and heterogeneous metadata by proposing a general-purpose, content-based multimedia retrieval in combination with a traditional Boolean retrieval system. CollaXRSearch [19] was a collaborative virtual reality system designed to facilitate coordination between teammates in terms of inputting and exploration operations in lifelog search.

Lifeseeker 6.0 [16] team has participated for the longest number of consecutive years. In the 6th version, their enhancements include improvements to the user interface and the backend reconstruction by combining the E-LifeSeeker structure with contrastive learning between texts. VISIONE [3] is a strong team in the VBS challenge [35]. They adapted the video to a lifelog retrieval system by enhancing result visualization. LifeExplore [22] employed a free-text search with FAISS and CLIP with other modalities like semantic concepts, contained objects, recognized text, and metadata. It also improved the GUI by integrating a query-building component with temporal search and new query filters and sub-filters. Libro [12] proposed a lifelog search browser that applies video search technologies to lifelog search. They treated lifelog data as continuous video clips and adopted text-to-image and image-to-image search. LifeGraph 4 [24] used multimodal knowledge graphs and vision-language models for lifelog retrieval. Memento 4.0 [1] proposed a similar approach to us by introducing a prototype conversational search system. They leveraged GPT3.5 Turbo and Mistral 7B to enable free-form natural language interaction. MyEachtraX [29] proposed a mobile-based interface for lifelog question answering. This used LLM and Multimodal LLM to enhance query-parsing, post-processing, and question-answering processes in lifelog retrieval. Vovento-Pro [2] employed an advanced voice interaction for lifelog retrieval. It used Whisper technology to capture queries from voice to support quick search. Exquisitor [14] blended conversational search with relevance feedback to address the challenges in descriptive lifelog retrieval.

We observe a trend of using LLM to enhance retrieval and QA performance from all the previous LSC systems. In addition, the embedding-based retrieval technique remains the best approach. From our previous work and related research, we aim to improve our system by incorporating RAG to improve the performance in QA tasks.

3 MemoriEase 3.0 System

This section presents the MemoryEase 3.0 system, from data processing and indexing to retrieval for KIS, relevance feedback for AD, and RAG for QA. Figure 1 illustrates the system's overall architecture. There are two phases in how the lifelog images and data are indexed offline and retrieved online. In the offline stage, images are encoded into embeddings by two embedding models, and along with metadata and descriptions of images, they are indexed into an Elasticsearch database. In the online stage, users interact with the conversational user interface to pose a query. The query is encoded into embeddings and compared to image embeddings in the database before returning the results to users. More details about each stage are provided in the following subsections.

3.1 Data Processing and Indexing

The images in lifelog dataset for LSC'25 are captured at a resolution of 1024x768 pixels using a Narrative Clip device. Additionally, the dataset includes metadata such as time, location, and biometrics. A visual concept dataset derived from the core image dataset provides scene information for each image. To further enrich the dataset, we incorporate descriptions of lifelog images from the BLIP2 model [18]. These descriptions provide narratives about the content of images. Learning from the experience of previous LSCs, we do not group images into events because the retrieval system may find incorrect images if two semantically different images are grouped.

From the lifelog images, we extract image embeddings using the BLIP2 and CLIP models. The embedding size for BLIP2 is 256, while CLIP produces a larger embedding size of 1024. We also extract metadata details, including city, semantic name, local time, hour, day of the week, weekend, and time period (morning, afternoon, etc) to serve as filters. Along with BLIP2-generated image descriptions,

[2]https://openai.com/o1/
[3]https://milvus.io/

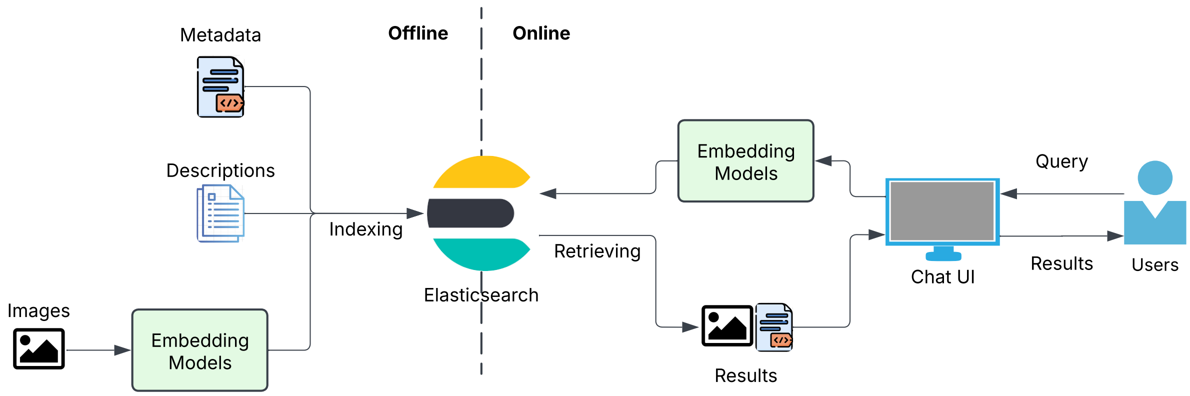

Figure 1: MemoriEase overall architecture for retrieving lifelog data.

all this information is indexed into an index in the Elasticsearch vector database.

3.2 Text to Lifelog Images Retrieval

For the KIS task, we propose a text-to-image retrieval pipeline. Firstly, the query is processed to extract the filters and the filter-free query. The filters include several pre-extracted metadata from the previous subsections and are used in filtering before the similarity calculation, while the processed query is encoded into two embeddings by two models (CLIP and BLIP2). The embeddings are used to calculate the cosine similarity with the image embeddings. We use a weighted average approach to combine the cosine similarity from two embeddings. The weight is chosen by retrieving several queries and tuning the best weight. A ranked list of images is returned to users along with corresponding metadata to display in the user interface.

3.3 User Interface

We have retained the conversational user interface, which is nearly unchanged from the previous version. Figure 2 illustrates an example of retrieving images for an Ad-hoc query. The LLM model generates a response to the user's query and displays the images and metadata on the right side. Users can sort the image by semantic name or by time. Users can immediately submit the result or save it using the button on the image. We also divide the user interface by tasks. In the Ad-hoc task, users can choose an image to place in a pool and search for similar images. In the KIS task, users have more filters and temporal search boxes to put in. These interfaces were introduced in the previous version of this system.

3.4 RAG for QA task

To address the challenge of the QA task, especially in the diverse and lifetime-spanning nature of lifelog data, we propose a RAG approach. This RAG approach first retrieves events related to the question and then passes the retrieved data to an LLM to generate the answer. Depending on the type of question, which can be a visual-related question such as "What is the colour of my carpet?" or a metadata-related question such as "How many times did I picnic by the sea?", we have different strategies in retrieval and prompts for LLM to address them. We use a heuristic-based approach to classify two types of question based on the question words. For visual-related questions, we only retrieve the top 10 images with no descriptions or reranking applied and provide them to LLM, along with the question. We retrieve more images (30 to 50), descriptions, and metadata and apply reranking descriptions for metadata-related questions. Figure 3 illustrates the pipeline of our RAG approach. The question is processed in the same way as the query to extract filters, encode embeddings, and search in an Elasticsearch index. Retrieved results include images, metadata of the images and the textual descriptions of the images.

In the reranking stage, we use a cross-encoder model to rerank the list of descriptions to the question. We fine-tune a BERT [6] model on a set of questions and ground-truth descriptions to create a reranking model. This is a binary classification model that classifies whether the description is relevant to the question or not. The relevance score is the probability that the description is relevant to the question. We use the reranking model to compute the relevance score of descriptions of retrieved results. This relevance score is used to rerank the ranked list of retrieval results.

The reranked descriptions, along with images and metadata, are formulated into a context in the following format: "<image> <description> on <metadata_time>" <metadata_location> ... <question>. For visual-related questions, the prompt is simple as follows: "<image> <image> ... <image>. <question>". We then feed the prompt to the state-of-the-art reasoning GPT-o1 model to generate the answer. Aggregation questions can be solved correctly if the input descriptions are correct. The answer is then passed to the user interface for display.

3.5 Enhanced Relevance Feedback

In the previous version of MemoriEase, relevance feedback [37] was used for the ad-hoc task in searching images using images. Users first input a query and retrieve a list of images. Then, they choose the images to put in a pool and click search again to search for similar images of pooled images. We encode the pooled images into embeddings, get their average, and retrieve new images in

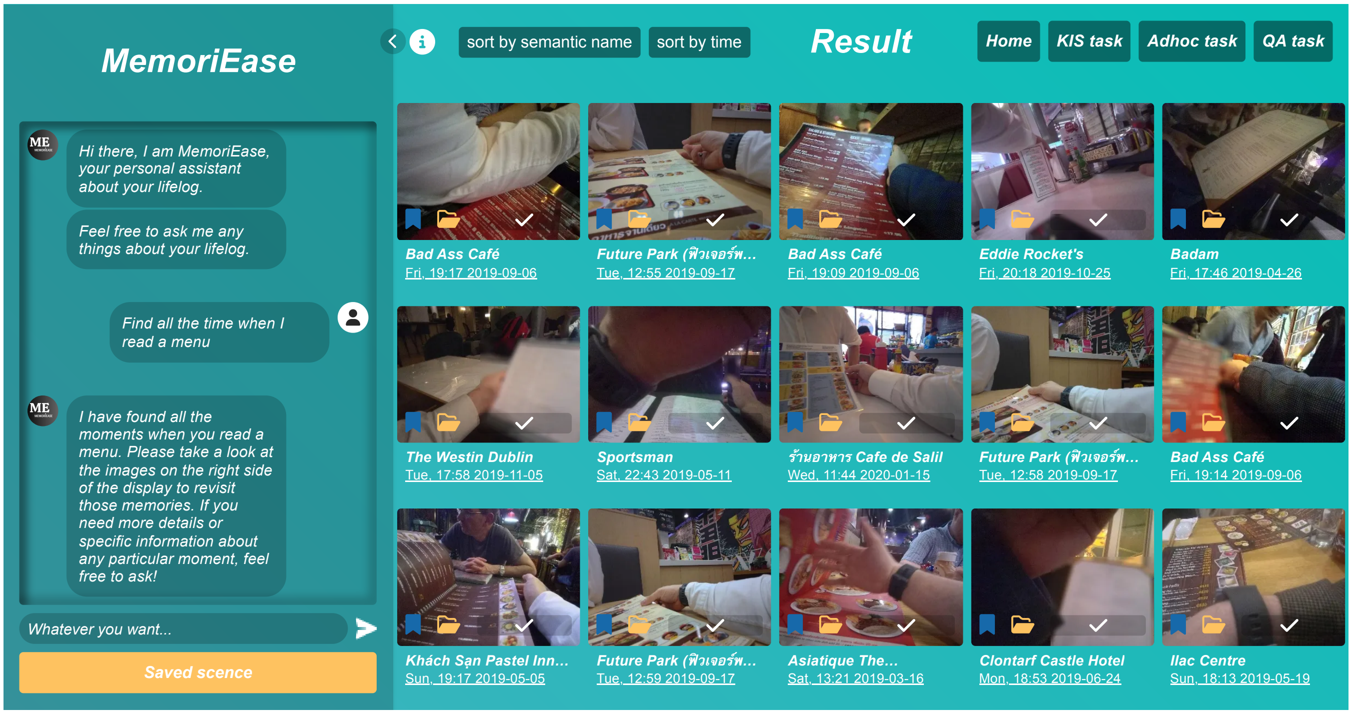

Figure 2: MemoriEase 3.0 UI

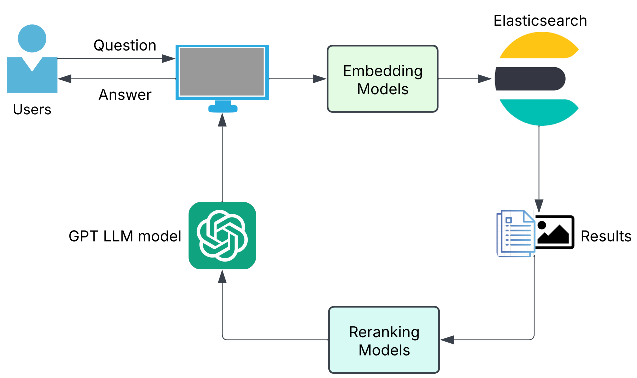

Figure 3: RAG pipeline for QA

Elasticsearch for cosine similarity between embeddings. However, the original query and related filters are not considered in the image-to-image retrieval. To address this problem, we extract the filters from the original query and apply them to later retrieval. In addition, for the pooled image embedding encoding, we average the query embedding with the pooled image embeddings with an appropriate weight to balance the original information and the user's choice information. These enhancements are expected to improve the performance of MemoriEase in the ad-hoc task.

4 Performance at LSC'25

The LSC'25 challenge attracted 12 systems and 21 participants (a system can be used by multiple users). The MemoriEase system participated remotely with only 1 participant. Remarkably, we achieved the third position in the leaderboard, following the MEMORIA and Snapseek teams. We also achieved the second-highest position in the QA sub-task with a score of 96 thanks to the RAG component. The score of KIS queries also increased by 10 points, but the performance on Adhoc sub-task remained the same compared to last year's competition. Figure 4 illustrates the leaderboard of LSC'25.

Specifically, we resolved 4 out of 6 QA queries, which is equal to top teams like MEMORIA, SnapSeek, and VitaChronicle, but thanks to the quick time of finding the answer, we ranked second place in the QA task, following the VitaChronicle team. There was a query that no other team found the correct answer except our system. This is "What is the name of the salesperson who sells me a Japanese car before June 2020?". We found the event of visiting a car showroom in May 2020, and submitted the name of the showroom called "Joe Duffy". However, the answer was wrong, and the feedback was the name of the salesperson instead of the name of the showroom. We used the browsing by time feature in our system to scroll to the image of talking to a salesperson and found his name tag in the table. We submitted the result at the last second and achieved the highest score for that query. This is an interesting insight that to find the correct answer, one needs an excellent combination of the good use of system functions and logical, reasoning thinking on the way to find the answer. Some answers for queries are not explicitly found on the first time of search but through refinement, the answer can be found.

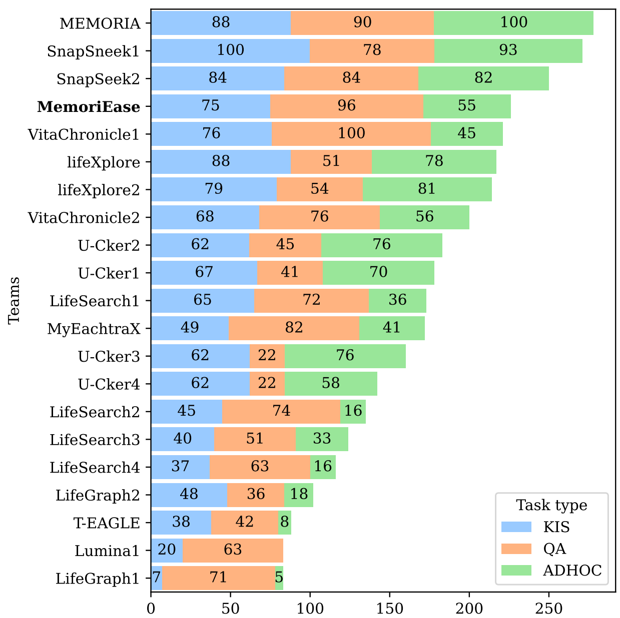

Figure 4: LSC'25 leaderboard.

On the KIS task, we also found 4 out of 6 correct answers, but it is far below the 6 out of 6 correct answers from top teams like SnapSeek or lifeXplore. We found out that if we can not find the correct answers in the first 3 hints, it is difficult to find the correct answers even when the time and location of queries are revealed in the last hint. The two wrong queries are "I was in the kitchen of a restaurant" and "I was shopping for a rat trap with my brother". The system failed to find the answer because it only found a place like a kitchen, but in a shop with a lot of products, the correct answer is nearly an empty kitchen with a lot of pipes. The temporal information did not help to find the answer. In the second query, the system found the mousetrap with the label "Mouse trap" instead of the rat trap. The correct answer showed the small rat trap, which may lead to the embedding model failing to encode it.

The Adhoc task in LSC'25 this year is challenging with 6 queries, but 5 of them have fewer than 10 correct images. Only the query "Drinking beer other than Guinness with other people" has more than 10 correct images. Due to this scenario, the relevance feedback for visual similarity search in our system did not work well to find as many visually similar images. We relied mostly on text-to-image search to find the answers and failed to find the answer for the query "Menu in a picture frame in a restaurant". The object of the menu in a picture frame is unique and strange to the embedding model, so it only found the image of a normal menu or a big menu hanging on the wall. If we could find one single correct image of the picture frame menu, we could find more similar images, but it failed in the first steps, so we cannot submit any correct images for this query. We achieved an average score of 55 in this task due to the slow submission time and wrong submissions.

Overall, we performed quite well in the LSC'25 with the third position. Thanks to the RAG for QA, we can find the answer quickly, and a little luck in the car salesperson's name helps us to achieve the second position in the QA task. Queries in LSC are becoming more and more difficult, which require not only good systems but also quick action and logical thinking. Some components in our system can be improved to deal with the weakness during competition and will be introduced in later versions.

5 Conclusions

This paper introduced the third version of the MemoriEase lifelog retrieval system for the LSC'25 challenge. We propose several enhancements in integrating CLIP and BLIP2 embedding models, improving the relevance feedback mechanism, and introducing a RAG approach for the QA task. Our system achieved 7 correct results out of 10 KIS topics and 6 correct answers from 9 QA questions, showing a clear improvement over previous versions. Despite these advancements, there are still several challenges in the LSC tasks. Difficult QA cases involving complex temporal or metadata reasoning still pose limitations for the current system. We will improve our RAG approach for future work by incorporating a better retrieval mechanism for complex queries and LLM reasoning for aggregation questions.

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]. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

References

[1] Naushad Alam, Yvette Graham, and Cathal Gurrin. 2024. Memento 4.0: A Prototype Conversational Search 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, 82–87. doi:10.1145/3643489.3661126

[2] Ahmed Alateeq, Mark Roantree, and Cathal Gurrin. 2024. Voxento-Pro: An Advanced Voice Lifelog Retrieval Interaction for Multimodal Lifelogs. 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, 105–110. doi:10.1145/3643489.3661130

[3] Giuseppe Amato, Paolo Bolettieri, Fabio Carrara, Fabrizio Falchi, Claudio Gennaro, Nicola Messina, Lucia Vadicamo, and Claudio Vairo. 2024. Will VISIONE Remain Competitive in Lifelog Image Search?. 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, 58–63. doi:10.1145/3643489.3661122

[4] Mehdi Cherti, Romain Beaumont, Ross Wightman, Mitchell Wortsman, Gabriel Ilharco, Cade Gordon, Christoph Schuhmann, Ludwig Schmidt, and Jenia Jitsev. 2023. Reproducible scaling laws for contrastive language-image learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2818–2829.

[5] 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.

[6] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805 [cs.CL] https://arxiv.org/abs/1810.04805

[7] 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.

[8] Cathal Gurrin, Björn Þór Jónsson, Duc Tien Dang Nguyen, Graham Healy, Jakub Lokoc, Liting Zhou, Luca Rossetto, Minh-Triet Tran, Wolfgang Hürst, Werner Bailer, et al. 2023. Introduction to the sixth annual lifelog search challenge, LSC'23. In *Proceedings of the 2023 ACM International Conference on Multimedia Retrieval*. 678–679.

[9] 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.

[10] 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.

[11] Cathal Gurrin, Liting Zhou, Graham Healy, Allie Tran, Luca Rossetto, Werner Bailer, Duc-Tien Dang-Nguyen, Steve Hodges, Björn Þór Jónsson, Minh-Triet Tran, and Klaus Schöffmann. 2025. Introduction to the 8th Annual Lifelog Search Challenge, LSC'25. In *Proceedings of the 2025 International Conference on Multimedia Retrieval* (Chicago, IL, USA) *(ICMR '25)*. Association for Computing Machinery, New York, NY, USA, 2143–2144. doi:10.1145/3731715.3734579

[12] Nico Hezel, Konstantin Schall, Bruno Schilling, Klaus Jung, and Kai Uwe Barthel. 2024. Libro - Lifelog Search Browser. 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, 70–75. doi:10.1145/3643489.3661124

[13] Minh-Quan Ho-Le, Huy-Hoang Do-Huu, Duy-Khang Ho, Nhut-Thanh Le-Hinh, Hoa-Vien Vo-Hoang, Van-Tu Ninh, and Minh-Triet Tran. 2024. SnapSeek: An Interactive Lifelog Acquisition 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, 24–29. doi:10.1145/3643489.3661116

[14] Omar Shahbaz Khan, Ujjwal Sharma, Hongyi Zhu, Stevan Rudinac, and Björn Þór Jónsson. 2024. Exquisitor at the Lifelog Search Challenge 2024: Blending Conversational Search with User Relevance Feedback. 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, 117–121. doi:10.1145/3643489.3661132

[15] 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.

[16] Hoang-Bao Le, Thao-Nhu Nguyen, Tu-Khiem Le, Minh-Triet Tran, Thanh-Binh Nguyen, Van-Tu Ninh, Liting Zhou, and Cathal Gurrin. 2024. LifeSeeker 6.0: Leveraging the linguistic aspect of the lifelog system in 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, 53–57. doi:10.1145/3643489.3661121

[17] 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

[18] 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

[19] Duy-Nam Ly, Dinh-Thuan Duong-Le, Gia Huy Vuong, Van-Son Ho, Van-Tu Ninh, Minh-Triet Tran, and Khanh-Duy Le. 2024. CollaXRSearch: A Collaborative Virtual Reality System for Lifelog Retrieval. 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, 76–81. doi:10.1145/3643489.3661125

[20] Thang-Long Nguyen-Ho, Onanong Kongmeesub, Minh-Triet Tran, Dongyun Nie, Graham Healy, and Cathal Gurrin. 2024. EAGLE: Eyegaze-Assisted Guidance and Learning Evaluation for Lifeloging Retrieval. 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, 18–23. doi:10.1145/3643489.3661115

[21] Eirik Pagani Vavik, Michela Wilhelmsen, Jenny Dal Østensen, Duc-Tien Dang-Nguyen, Van-Tu Ninh, Manoj Kesavulu, Cathal Gurrin, Gia-Huy Vuong, Van-Son Ho, and Minh-Triet Tran. 2024. VitaChronicle: Applying UX/UI Principles and Guidelines to Enhance Lifelog Retrieval System Design. 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, 30–35. doi:10.1145/3643489.3661117

[22] Martin Rader and Klaus Schoeffmann. 2024. lifeXplore at the Lifelog Search Challenge 2024. 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, 64–69. doi:10.1145/3643489.3661123

[23] 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

[24] Luca Rossetto, Athina Kyriakou, Svenja Lange, Florian Ruosch, Ruijie Wang, Kathrin Wardatzky, and Abraham Bernstein. 2024. LifeGraph 4 - Lifelog Retrieval using Multimodal Knowledge Graphs and Vision-Language Models. 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, 88–92. doi:10.1145/3643489.3661127

[25] Loris Sauter, Ralph Gasser, Laura Rettig, Heiko Schuldt, and Luca Rossetto. 2024. General Purpose Multimedia Retrieval with vitrivr at 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, 47–52. doi:10.1145/3643489.3661120

[26] Florian Spiess, Heiko Schuldt, and Luca Rossetto. 2024. Spatiotemporal Lifelog Analytics in Virtual Reality with vitrivr-VR. 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, 7–11. doi:10.1145/3643489.3661113

[27] Iselin Ågotnes Steffensen, Sara Tumey Celik Nystad, Margrethe Liahjell, Duc-Tien Dang-Nguyen, Manoj Kesavulu, Van-Tu Ninh, Cathal Gurrin, Gia-Huy Vuong, Van-Son Ho, and Minh-Triet Tran. 2024. T@Retrospect: A Journey Through Time: A User-Centered Prototype Enabling Seamless Information Retrieval Across Expertise Levels. 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, 36–40. doi:10.1145/3643489.3661118

[28] Allie Tran, Werner Bailer, Duc-Tien Dang-Nguyen, Graham Healy, Steve Hodges, Björn Þór Jónsson, Luca Rossetto, Klaus Schoeffmann, Minh-Triet Tran, Lucia Vadicamo, and Cathal Gurrin. 2025. The State-of-the-Art in Lifelog Retrieval: A Review of Progress at the ACM Lifelog Search Challenge Workshop 2022-24. arXiv:2506.06743 [cs.MM] https://arxiv.org/abs/2506.06743

[29] Ly Duyen Tran, Thanh-Binh Nguyen, Cathal Gurrin, and Liting Zhou. 2024. MyEachtraX: Lifelog Question Answering on Mobile. 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, 93–98. doi:10.1145/3643489.3661128

[30] Quang-Linh Tran, Binh Nguyen, Gareth Jones, and Cathal Gurrin. 2024. MemoriEase at the NTCIR-17 Lifelog-5 Task. (2024).

[31] Quang-Linh Tran, Binh Nguyen, Gareth Jones, and Cathal Gurrin. 2025. DCU MemoriEase at the NTCIR-18 Lifelog 6 Task. *(No Title)* (2025), none.

[32] 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

[33] 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*. 30–35.

[34] Maria Tysse Hordvik, Julie Teilstad Østby, Duc Tien Dang Nguyen, Manoj Kesavulu, Thao-Nhu Nguyen, Tu-Khiem Le, Cathal Gurrin, and Minh-Triet Tran. 2024. LifeLens 2.0: Improving Efficiency and Usability in Lifelog Retrieval Systems through UX/UI Design. 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, 41–46. doi:10.1145/3643489.3661119

[35] Lucia Vadicamo, Rahel Arnold, Werner Bailer, Fabio Carrara, Cathal Gurrin, Nico Hezel, Xinghan Li, Jakub Lokoc, Sebastian Lubos, Zhixin Ma, et al. 2024. Evaluating performance and trends in interactive video retrieval: Insights from the 12th vbs competition. *IEEE Access* (2024).

[36] Gia Huy Vuong, Van-Son Ho, Tien Thanh Nguyen Dang, Xuan-Dang Thai, Thang-Long Nguyen-Ho, Minh-Khoi Pham, Tu-Khiem Le, Van-Tu Ninh, Graham Healy, Cathal Gurrin, and Minh-Triet Tran. 2024. LifeInsight2.0: An Enhanced Approach for Automated Lifelog Retrieval in 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, 1–6. doi:10.1145/3643489.3661112

[37] Xiang Sean Zhou and Thomas S Huang. 2003. Relevance feedback in image retrieval: A comprehensive review. *Multimedia systems* 8 (2003), 536–544.