

Advances in Lifelog Data Organisation and Retrieval at the NTCIR-14 Lifelog-3 Task

Cathal Gurrin¹, Hideo Joho², Frank Hopfgartner³, Liting Zhou¹,
Van-Tu Ninh¹, Tu-Khiem Le¹, Rami Albatal¹, Duc-Tien Dang-Nguyen⁴, and
Graham Healy¹

¹ Dublin City University, Ireland

² University of Tsukuba, Japan

³ University of Sheffield, UK

⁴ University of Bergen, Norway

Abstract. Lifelogging refers to the process of digitally capturing a continuous and detailed trace of life activities in a passive manner. In order to assist the research community to make progress in the organisation and retrieval of data from lifelog archives, a lifelog task was organised at NTCIR since edition 12. Lifelog-3 was the third running of the lifelog task (at NTCIR-14) and the Lifelog-3 task explored three different lifelog data access related challenges, the search challenge, the annotation challenge and the insights challenge. In this paper we review the dataset created for this activity, activities of participating teams who took part in these challenges and we highlight learnings for the community from the NTCIR-Lifelog challenges.

Keywords: Lifelog · Information Retrieval · Test Collection

1 Introduction

Lifelogging refers to the process of digitally capturing a detailed trace of life activities in a passive manner [12]. NTCIR-14 hosted the third running of the Lifelog task which aimed to advance efforts at lifelog data analytics and retrieval. Over the three iterations of the task, from NTCIR-12 [9], NTCIR-13 [10] and this year, we report that nearly 20 participating research groups have submitted official runs for the various sub-tasks and we can identify progress in the approaches being made across all tasks, but especially so for the lifelog retrieval task.

Before we begin our review of the submissions for the lifelog task, we introduce the concept of lifelogging by presenting the definition of Dodge and Kitchin [4], who refer to lifelogging as ‘*a form of pervasive computing, consisting of a unified digital record of the totality of an individual’s experiences, captured multimodally through digital sensors and stored permanently as a personal multimedia archive*’. This lifelog task was initially proposed because the organisers identified that technological progress had resulted in lifelogging becoming a more commonplace activity, thereby necessitating the development of new forms

of data analytics, organisation and retrieval that are designed to operate over archives of multimodal lifelog data. Additionally, the organisers note recent efforts to employ lifelogging in many domains, such as a means of supporting human memory [13] or facilitating large-scale epidemiological studies in health-care [25], lifestyle monitoring [27], diet/obesity monitoring [30], or for exploring societal issues such as privacy-related concerns [14] and behaviour analysis [6]. The increasing uptake of lifelogging as a personal and practitioner technology have also lead to related activities in ImageCLEF [2], the Lifelog Search Challenge [11] and a related task at MediaEval 2019. These have all been designed to encourage comparative research into the annotation, retrieval and analysis of multimodal lifelog data.

At NTCIR-14 there were three lifelog sub-tasks, a semantic search sub-task (LEST), a lifelog annotation sub-task (LADT) and an insights sub-task (LIT), of which the LADT was the only new sub-task. In this paper we will provide an overview of the lifelog task, in terms of the dataset, the sub-tasks and the submissions submitted by participating organisations.

2 Task Overview

The Lifelog-3 task explored a number of approaches to information access and retrieval from personal lifelog data, each of which addressed a different challenge for lifelog data organization and retrieval. The three sub-tasks, each of which could have been participated in independently, are as follows:

- **Lifelog Semantic Access sub-Task (LSAT)** to explore search and retrieval from lifelogs.
- **Lifelog Activity Detection sub-Task (LADT)** to identify Activities of Daily Living (ADLs) from lifelogs, which have been employed as indicators of the health of an individual.
- **Lifelog Insight sub-Task (LIT)** to explore knowledge mining and visualisation of lifelogs in an open and topic agnostic manner.

We will now describe and motivate each task in detail.

2.1 LSAT Sub-Task

The LSAT sub-task was a known-item search task applied over lifelog data, the aim of which was to advance the state-of-the-art in retrieval of lifelog data in response to typical user information needs. This was created in the spirit of the ad-hoc task at TREC which had lead to the development of many enhanced text search techniques. In this sub-task, the participants had to retrieve a number of specific *moments* in a lifelogger’s life in response to a query topic that follows the traditional three-part TREC format of title, description and narrative. We consider moments to be semantic events, or activities that happened at least once in the dataset and are at least one-minute in duration. The task can best

be compared to a traditional known-item search task with one (or more) relevant items per topic, though operating over a multimodal dataset. Participants were allowed to undertake the LAST task in an interactive or automatic manner. For interactive submissions, a maximum of five minutes of search time was allowed per topic. The LSAT task included 24 search tasks, generated by the lifeloggers who gathered the data.

2.2 LADT Sub-Task

The aim of this sub-task was to develop new approaches to the annotation of multimodal lifelog data in terms of activities of daily living, with a motivation from various healthcare and public-health surveys carried out with lifelog data [25] [30]. An ontology of important lifelog activities of daily living, guided by Kahneman’s lifestyle activities [15] were provided as a multi-label classification task. The task required the development of automated approaches for multi-label classification of multimodal lifelog data. Both image content as well as provided metadata and external evidence sources were available to be used to generate the activity annotations. The submission was comprised of one or more activity labels for each image.

2.3 LIT Sub-Task

The LIT sub-task was exploratory in nature and the aim of this sub-task was to gain insights into the lifelogger’s daily life activities. Participants were requested to provide insights about the lifelog data that support the lifelogger in reflecting upon the data and provide for efficient/effective means of visualisation of the data. There was no explicit evaluation for this topic-agnostic task, so participants were free to analyse and describe the data in whatever manner they wished.

3 Description of the Lifelog-3 Test Collection

As with each of the previous two Lifelog NTCIR tasks, a purpose-built test collection was prepared and released, which was designed to support both ad-hoc retrieval and insights generation from lifelog data. This dataset was prepared following the process described in [3] and took measures to ensure temporal data alignment and privacy preservation of the lifeloggers and bystanders in their data. As with previous tasks, the data was gathered by a number of lifeloggers (in this case, two lifeloggers) who wore various sensing devices and gathered biometric data for most (or all) of the waking hours in the day, along with some manual annotations.

3.1 Details of the Dataset

The data consists of a medium-sized collection of rich multimodal lifelog data over 42 days by the two lifeloggers. The contribution of this dataset over previously released lifelog datasets was the inclusion of additional biometric data,



Fig. 1. Examples of Wearable Camera Images from the Dataset

a manual diet log and the inclusion of conventional photos. This makes it the richest personal lifelog dataset ever released. The data consists of:

- **Multimedia Content.** Wearable camera PoV images, captured at a rate of two images per minute by an OMG Autographer passive-capture wearable camera, and worn from breakfast to sleep. All recognisable faces and screens were blurred and every image was also resized down to 1024×768 resolution to ensure that text captured was illegible. For examples, see Figure 1. Additional multimedia content included a time-stamped record of music listening activities sourced from Last.FM⁵ and an archive of all conventional (active-capture) digital photos taken by the lifelogger.
- **Biometric Data.** Using FitBit fitness trackers⁶, the lifeloggers gathered 24×7 heart rate, calorie burn and steps as numeric lifelog data. In addition, continuous blood glucose monitoring captured readings every 15 minutes using the Freestyle Libre wearable sensor⁷, which provided a continuous record of the blood glucose levels of the individual, which would typically change based on foods consumed and physical activity rate.
- **Human Activity Data.** The daily activities of the lifeloggers were captured in terms of the semantic locations visited, physical activities (e.g. walking, running, standing) from the Moves app⁸, along with a time-stamped diet-log of all food and drink consumed which was manually recorded by the lifeloggers throughout the day.
- **Enhancements to the Data.** The wearable camera images were annotated with the outputs of a visual concept detector, which provided three types of

⁵ Last.FM Music Tracker and Recommender - <https://www.last.fm/>

⁶ Fitbit Fitness Tracker (FitBit Versa) - <https://www.fitbit.com>

⁷ Freestyle Libre wearable glucose monitor - <https://www.freestylelibre.ie/>

⁸ Moves App for Android and iOS - <http://www.moves-app.com/>

Table 1. Statistics of NTCIR-14 Lifelog Data

	Size
Number of Lifeloggers	2
Number of Days	43 days
Size of the Collection	14 GB
Number of Images	81,474 images
Number of Locations	61 semantic locations
Number of LSAT Topics	24 topics
Number of LADT Types	16 activities

outputs (attributes, categories and visual concepts). The image attributes and categories of the place in the image are extracted using PlacesCNN [29]. The visual concepts are detected object category and its bounding box extracted by using Faster R-CNN [22] trained on MSCOCO dataset [18]. These formed the official annotations of the visual content of the collection. In some cases, participants added additional non-official annotations using other sources, as outlined below, which in many cases was shown to provide additional beneficial metadata to assist the retrieval and organisation process.

3.2 Topics

The LSAT task includes 24 topics with pooled relevance judgements, as is often used in comparative benchmarking activities with large datasets. These LSAT topics were evaluated in terms of traditional Information Retrieval effectiveness measurements such as Precision, RelRet and MAP. These 24 topics were labelled as being one of two types, either precision-based or recall-based. Precision-based topics had a small number of relevant items in the dataset, whereas Recall-based topics would have had a larger number of relevant topics. Each topic was further labelled as being related to User 1, User 2 or both users. An example of a topic is shown in Figure 2, along with some example relevant image content from the collection. For the full list of the topics, please refer to the NTCIR14-Lifelog3 Overview paper [8] or the associated task website⁹.

For the LADT (Activity Detection) sub-task, there were sixteen types of activities defined for annotation, which aim to cover a wide-range of daily human activities. These were defined in order to make it easier for participants to develop event segmentation algorithms for the very subjective human event segmentation tasks. The sixteen types of activity were: *traveling, face-to-face interacting, using a computer, cooking, eating, time with children, houseworking, relaxing, reading, socialising, praying, shopping, gaming, physical activities, creative activities, and other activities* (i.e. those not represented by the previous fifteen labels). Each image can be tagged as belonging to one or more activities.

⁹ <http://ntcir-lifelog.computing.dcu.ie/>

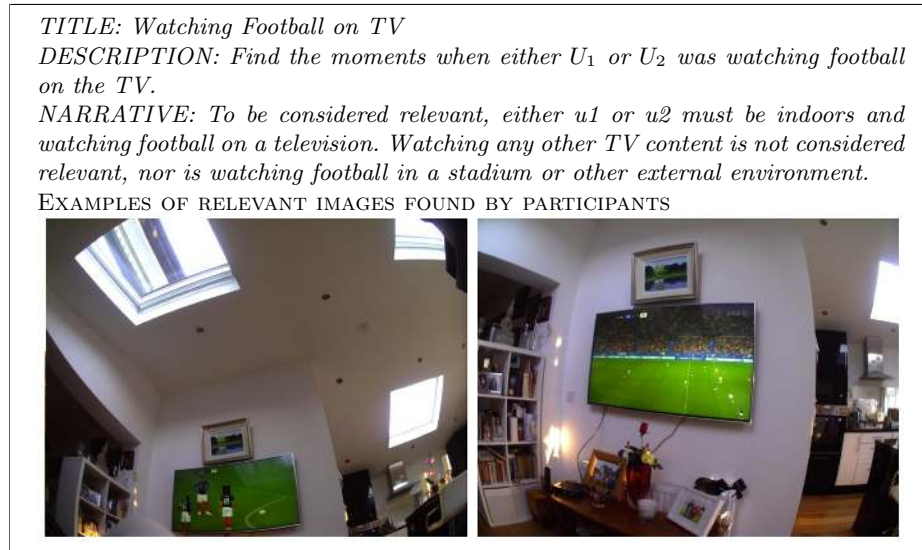


Fig. 2. LSAT topic example, including example results.

A description of each activity is provided in the NTCIR14-Lifelog3 Overview paper [8].

For the LIT task, there were no topics and participants were free to analyse the data in whatever manner they wished in order to extract meaningful insights. One group took part in the LIT task, which is outlined in the relevant section below.

3.3 Relevance Judgement and Scoring

As stated, pooled binary relevance judgements were generated for all 24 LSAT topics. Scoring for the LSAT sub-task was calculated using `trec_eval` [1]. Two custom applications were developed to support both the LSAT and LADT evaluation processes. For the LADT topics/labels, manual relevance judgements were performed over 5,000 of the images and these annotations were used in assessing participant performance. These images were chosen randomly from the collection and scores were calculated according to the following process. For each run, using the labelled subset of the test images, the score was calculated as the number of correctly predicted labels divided by the total number of labels in the ground truth collection (over all activities). It is worth noting that for some activities, the official runs did not include any labelled images i.e. gaming, praying, physical activity and time with children.

4 Participants and Submissions

In total fourteen participants signed up to the Lifelog-3 task at NTCIR-14, however only five participants submitted to any of the sub-tasks of the Lifelog task. We will now summarise the effort of the participating groups in the sub-tasks that they submitted to.

4.1 LSAT Sub-task

Four participating groups took part in the LSAT sub-task. **NTU (Taiwan)** took part in both the LSAT and LADT Tasks [7]. For the LSAT task, the NTU team developed an interactive lifelog retrieval system that automatically suggested a list of candidate query words to the user, and adopted a probabilistic relevance-based ranking function for retrieval. They enhanced the official concept annotations by applying the Google Cloud Vision API¹⁰ and pre-processed the visual content to remove images with poor quality and to offset the fish-eye nature of the wearable camera data. In the provided examples, this was shown to increase the quality of the non-official annotations. The interactive system facilitated a user to select from suggested query words and to restrict the results to a particular user and date/time interval. Three official runs were submitted, one automatic and two interactive. The first run (NTU-Run1) used an automatic query enhancement process using the top 10 nearest concepts to the query terms. The other two runs employed a user-in-the-loop (NTU-Run2 & NTU-Run3).

QUIK (Japan) from Kyushu University participated in the LSAT task with a retrieval system that integrates online visual WWW content in the search process and operated based on an underlying assumption that a lifelog image of an activity would be similar to images returned from a WWW search engine for similar activities [26]. The approach operated using only the visual content of the collection and used the WWW data to train a visual classifier with a convolutional neural network for each topic. For a given query, images from the WWW were gathered, filtered by a human and combined to create a new visual query (average of 170 images per query). In order to solve the lexical gap between query words and visual concept labels, a second run employed word embedding when calculating the similarities. Two runs were submitted. QUIK-Run1 used only visual concepts while QUIK-Run2 used the visual concepts as well as the query-topic similarity.

VNU-HCM (Vietnam) group took part in the LSAT task by developing an interactive retrieval system [19]. The research required a custom annotation process for lifelog data based on the identifiable habits of the lifeloggers. This operated by extracting additional metadata about each moment in the dataset, by adding in outputs of additional object detectors, manually adding in ten habit concepts, scene classification, and counting the number of people in the images. Associated with this new data source, the team developed a scalable and user-friendly interface that was designed to support novice users to generate queries

¹⁰ Google Cloud Vision API - <https://cloud.google.com/vision/>

Table 2. LSAT results for NTCIR-14 Lifelog-3 subtask.

Group ID	Run ID	Approach	MAP	P@10	RelRet
NTU	NTU-Run1	Automatic	0.0632	0.2375	293
NTU	NTU-Run2	Interactive	0.1108	0.3750	464
NTU	NTU-Run3	Interactive	0.1657	0.6833	407
DCU	DCURun1	Interactive	0.0724	0.1917	556
DCU	DCU-Run2	Interactive	0.1274	0.2292	1094
HCMUS	HCMUS-Run1	Interactive	0.3993	0.7917	1444
QUIK	QUIK-Run1	Automatic	0.0454	0.1958	232
QUIK	QUIK-Run2	Automatic	0.0454	0.1875	232

and browse results. One run was submitted (HCMUS-Run1), which was the best performing run at Lifelog-3.

DCU (Ireland) group took part in the LSAT task by developing an interactive retrieval engine for the lifelog data with the official annotations [20]. The retrieval engine was designed to be used by novice users and relied on an extensive range of facet filters for the lifelog data and limited search time to five minutes for each topic. The results of a query were displayed in 5 pages of 20 images, and for any given image, the user could browse the (temporal) context of that image in order to locate relevant content. The user study and subsequent questionnaire illustrated that the interface and search supports provided were generally liked by users.

It can be seen from Table 2 that the results could be analysed by considering both automatic and interactive runs. For automatic runs, NTU achieve the best scores in all three measures: MAP, P@10 and RelRet of 6.32%, 23.75% and 293 respectively while QUIK also generates competitive results. For interactive runs, the team from HCMUS obtains the highest scores of all three measures, which are also the highest results in two approaches with MAP, P@10 and RelRet of 39.93%, 79.17% and 1444 respectively. Whether this performance is due to higher quality annotations or the intuitive interface is not yet clear. While NTU focused on increasing P@10 of their interactive system (68.33%), DCU concentrated on increasing the recall measure by returning as many number of relevant images as possible (RelRet: 1094 images). Both teams managed to achieve the second highest scores of the corresponding measure system. We can hypothesise that these findings suggest that enhancing the official visual annotations with higher-quality non-official annotations seems to lead to an enhanced performance in both automatic and interactive retrieval. This finding is in line with similar findings from previous editions of the lifelog task at NTCIR-12 [9] and NTCIR-13 [10], as well as findings from the related LSC search challenge [11].

4.2 LADT Task

The NTU group (*Taiwan*) took part in the LADT task [7] and developed a new approach for the multi-label classification of lifelog images. In order to train the

classifier, the authors manually labelled four days, which were chosen because they covered most of the activities that the lifeloggers were involved in. It is noted that there is no training data generated for some of the activities for user 1 and user 2. Since only one group took part, no comparison is possible between participants. Readers are referred to the NTU paper [7] for details of their different runs and the comparative performance of these runs.

4.3 LIT Sub-task

For the LIT task, there were no submissions to be evaluated in the traditional manner; rather the LIT task was an exploratory task to explore a wide-range of options for generating insights from the lifelog data. One group took part in the LIT task. **THUIR (China)** developed a number of detectors for the lifelog data to automatically identify the status/context of a user [19], which could be used in many real-world applications, especially so for forms of assistive technology. There were three detectors developed for inside/outside status, alone/not alone status and working/not working status. These detectors were designed to operate over both non-visual and visual lifelog data. A comparison between the two approaches showed that the visual features (integrating supervised machine learning) were significantly better than non-visual ones based on metadata. Finally the authors presented a number of statistics of users' activities for all three detectors, which highlighted the activities of the two users in a highly visual manner.

5 Advancing State-of-the-art Over Three Editions

Lifelog-3 was the third in a series of collaborative benchmarking exercises for lifelog data at NTCIR. It attracted five active participants, four for the automatic LSAT sub-task, one for the LADT sub-task and one for the LIT sub-task. Over the course of the three editions, we can identify a number of successful techniques that can be applied over lifelog data to enhance the performance of both automatic and interactive systems. We can summarise the the advances in state-of-the-art as follows:

- The utilisation of **non-official visual concept detectors** is considered a positive addition. The official visual concepts released with the test collections were using off-the-shelf approaches, but customised techniques or employing latest approaches from the field were shown to significantly advance performance. For example, at NTCIR-12, the LIG-MRM group (France) performed significantly ahead of all other submissions, by focusing on enhancing the performance of the visual concept detectors to be used for retrieval [23]. Likewise the VCI2R at NTCIR-13 proposed a general framework to bridge the semantic gap between lifelog data and the event-based LSAT topics [17] by enhancing the visual annotations and employing temporal smoothing of annotations, which proved to be the most successful approach at NTCIR-13. At NTCIR-14, the VNU-HCM (Vietnam) group developed an interactive

retrieval system [19] that used enhanced visual metadata (including human annotations), which outperformed all other approaches.

- We note the integration of **additional metadata sources** in some approaches, which was also considered to be beneficial. For example, at NTCIR-12 the QUIK team (Japan) integrated online visual WWW content in the search process to enhance the performance of image-based retrieval by using the WWW data to train a visual classifier with a convolutional neural network for each topic [26]. Also at NTCIR-12, the VTIR team identified that location was a very important component in the information retrieval process [28] and thus enhanced location semantic descriptions were used to facilitate retrieval.
- Another source of enhancement was taking measures to **reduce the lexical gap** between user queries and concept annotations using some form of term expansion, and the current consideration is that this could be achieved using word embedding approaches. At NTCIR-12, the QUT group took an approach to retrieval that utilised spreading annotations (via visual similarity) to generate long, descriptive paragraphs of text to annotate the lifelog content, as opposed to the conventional tag-based approach [24]. The IDEAS Institute for Information Industry (Taiwan) took a textual approach to retrieval [16] at NTCIR-14, utilising word2vec to better match visual concepts to user queries (an approach referred to as bridging the lexical gap at NTCIR-14) via query expansion.
- We note the increasing use of **interactive systems** throughout the editions to address the challenges posed by the LSAT task. At NTCIR-12, there was one interactive retrieval system presented by the team from the University of Barcelona who developed an interactive retrieval engine that integrated a semantic-content tagging tool to enhance the quality of the annotations [21], which naturally outperformed all automatic approaches. At NTCIR-13, the DCU team employed a human-in-the-loop to translate the provided queries into system queries for their retrieval engine, in one of their runs [5]. However, at NTCIR-14, we note that three of the participants developed interactive systems and a fourth participant also integrated the human-in-the-loop query enhancement. This interest in interactive retrieval systems has led to the development of a parallel benchmarking activity, the Lifelog Search Challenge for interactive retrieval systems [11].

At this point, after the three editions of the lifelog task, the main approaches that the organisers’ consider to be valuable for lifelog access is the use of enhanced visual concept detectors to improve indexing and the application of approaches to bridging the lexical gap, either via some form of index term expansion or query-expansion.

6 Conclusion

In this paper, we described the data and the activities from the Lifelog-3 core-task at NTCIR-14, which was the third and final edition of the NTCIR-Lifelog

task, which included three sub-tasks, each of which addressed a different challenge for lifelog organisation and retrieval. For the LSAT sub-task, four groups took part and produced eight official runs including five interactive and three automatic runs. The approach taken by HCMUS, of enhancing the provided annotations with additional object detectors, habits, scenes and people analytics, along with an intuitive user interface, ensured that their runs were significantly better than the runs of any other participant. The LADT and LIT tasks attracted one participant each, so we are not in a position to draw any conclusions at this point.

After this, the third instance of the NTCIR-Lifelog task, we are beginning to see some learnings from the comparative benchmarking exercises. It can be seen that additional visual concept detectors, integrating external sources and addressing the lexical gap between users and the systems are priority topics for the research community to address. Likewise we note the interest in the community of developing interactive (user-in-the-loop) approaches to lifelog data retrieval. What we have not yet seen is widespread use of the temporally aligned metadata that accompanies the lifelog datasets. We hope that participants and readers will continue the effort to develop new approaches for the organisation and retrieval of lifelog data, and take part in future NTCIR, LSC and ImageCLEF efforts within the domain.

Acknowledgements

This publication has emanated from research supported in part by research grants from Irish Research Council (IRC) under Grant Number GOIPG/2016/741 and Science Foundation Ireland under grant numbers SFI/12/RC/2289 and SFI/13/RC/2106. We acknowledge the support and input of the DCU ethics committee and the risk & compliance officer.

References

1. Buckley, C.: Treceval ir evaluation package (2004)
2. Dang-Nguyen, D.T., Piras, L., Riegler, M., Zhou, L., Lux, M., Gurrin, C.: Overview of ImageCLEFlifelog 2018: daily living understanding and lifelog moment retrieval. In: CLEF2018 Working Notes. CEUR Workshop Proceedings, CEURWS, Avignon, France (September 10-14 2018) (2018)
3. Dang-Nguyen, D.T., Zhou, L., Gupta, R., Riegler, M., Gurrin, C.: Building a Disclosed Lifelog Dataset: Challenges, Principles and Processes. In: Content-Based Multimedia Indexing (CBMI) (2017)
4. Dodge, M., Kitchin, R.: 'Outlines of a world coming into existence': Pervasive computing and the ethics of forgetting. *Environment and Planning B: Planning and Design* **34**(3), 431–445 (2007). <https://doi.org/10.1068/b32041t>
5. Duane, A., Zhou, L., Dang-Nguyen, D.T., Gurrin, C.: Dcu at the ntcir-13 lifelog-2 task. In: Proceedings of NTCIR-13, Tokyo, Japan (2017)
6. Everson, B., Mackintosh, K.A., McNarry, M.A., Todd, C., Stratton, G.: Can wearable cameras be used to validate school-aged childrens lifestyle

- behaviours? *Children* **6**(2) (2019). <https://doi.org/10.3390/children6020020>, <http://www.mdpi.com/2227-9067/6/2/20>
7. Fu, M.H., Chia-Chun, C., Huang, G.H., Chen, H.H.: Introducing external textual knowledge for lifelog retrieval and annotation. In: The Fourteenth NTCIR conference (NTCIR-14) (2019)
 8. Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L., Ninh, T., Le, T.K., Albatal, R., Dang-Nguyen, D.T., Healy, G.: Overview of the ntcir-14 lifelog-3 task. In: The Fourteenth NTCIR conference (NTCIR-14) (06 2019)
 9. Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L., Albatal, R.: Overview of ntcir-12 lifelog task. In: Kando, N., Kishida, K., Kato, M.P., Yamamoto, S. (eds.) Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies. pp. 354–360 (2016)
 10. Gurrin, C., Joho, H., Hopfgartner, F., Zhou, L., Gupta, R., Albatal, R., Dang-Nguyen, D.T.: Overview of NTCIR-13 Lifelog-2 Task. In: The Thirteenth NTCIR conference (NTCIR-13). pp. 6–11 (2017)
 11. Gurrin, C., Schoeffmann, K., Joho, H., Zhou, L., Duane, A., Leibetseder, A., Riegler, M., Piras, L.: Comparing Approaches to Interactive Lifelog Search at the Lifelog Search Challenge (LSC2018). *ITE Transactions on Media Technology and Applications* **7**(2), 46–59 (2019)
 12. Gurrin, C., Smeaton, A.F., Doherty, A.R.: LifeLogging: Personal Big Data. *Foundations and Trends® in Information Retrieval* **8**(1), 1–125 (2014). <https://doi.org/10.1561/1500000033>, <http://www.nowpublishers.com/articles/foundations-and-trends-in-information-retrieval/INR-033>
 13. Harvey, M., Langheinrich, M., Ward, G.: Remembering through lifelogging: A survey of human memory augmentation. *Pervasive and Mobile Computing* **27**, 14–26 (2016). <https://doi.org/10.1016/j.pmcj.2015.12.002>, <http://dx.doi.org/10.1016/j.pmcj.2015.12.002>
 14. Hoyle, R., Templeman, R., Armes, S., Anthony, D., Crandall, D., Kapadia, A.: Privacy behaviors of lifeloggers using wearable cameras. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. pp. 571–582. *UbiComp '14*, ACM, New York, NY, USA (2014)
 15. Kahneman, D., Krueger, A.B., Schkade, D.A., Schwarz, N., Stone, A.A.: A survey method for characterizing daily life experience: The day reconstruction method. *Science* **306**, 1776–1780 (2004)
 16. Lin, H.L., Chiang, T.C., Chen, L.P., Yang, P.C.: Image searching by events with deep learning for ntcir-12 lifelog. In: Proceedings of NTCIR-12, Tokyo, Japan (2016)
 17. Lin, J., del Molino, A.G., Xu, Q., Fang, F., Subbaraju, V., Lim, J.H.: Vci2r at the ntcir-13 lifelog semantic access task. In: Proceedings of NTCIR-13, Tokyo, Japan (2017)
 18. Lin, T., Maire, M., Belongie, S.J., Bourdev, L.D., Girshick, R.B., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft COCO: common objects in context. *CoRR* **abs/1405.0312** (2014), <http://arxiv.org/abs/1405.0312>
 19. Nguyen, I.V.K., Shrestha, P., Zhang, M., Liu, Y., Ma, S.: Thuir at the ntcir-14 lifelog-3 task: How does lifelog help the users status recognition. In: The Fourteenth NTCIR conference (NTCIR-14) (2019)
 20. Ninh, V.T., Le, T.K., Zhou, L., Healy, G., Venkataraman, K., Tran, M.T., Dang-Nguyen, D.T., Smith, S., Gurrin, C.: A baseline interactive retrieval engine for the ntcir-14 lifelog-3 semantic access task. In: The Fourteenth NTCIR conference (NTCIR-14) (2019)

21. de Oliveira Barra, G., Ayala, A.C., Bolaos, M., Dimiccoli, M., Aghaei, M., Carn, M., Giro-I-Nieto, X., Radeva, P.: Lemore: A lifelog engine for moments retrieval at the ntcir-lifelog lsat task. In: Proceedings of NTCIR-12, Tokyo, Japan (2016)
22. Ren, S., He, K., Girshick, R.B., Sun, J.: Faster R-CNN: towards real-time object detection with region proposal networks. CoRR **abs/1506.01497** (2015), <http://arxiv.org/abs/1506.01497>
23. Safadi, B., Mulhem, P., Qunot, G., Chevallet, J.P.: Mrim-lig at ntcir lifelog semantic access task. In: Proceedings of NTCIR-12, Tokyo, Japan (2016)
24. Scells, H., Zuccon, G., Kitto, K.: Qut at the ntcir lifelog semantic access task. In: Proceedings of NTCIR-12, Tokyo, Japan (2016)
25. Signal, L.N., Smith, M.B., Barr, M., Stanley, J., Chambers, T.J., Zhou, J., Duane, A., Jenkin, G.L., Pearson, A.L., Gurrin, C., Smeaton, A.F., Hoek, J., Ni Mhurchu, C.: KidsCam: An Objective Methodology to Study the World in Which Children Live. *American Journal of Preventive Medicine* **53**(3), e89–e95 (2017), <http://dx.doi.org/10.1016/j.amepre.2017.02.016>
26. Suzuki, T., Ikeda, D.: Smart lifelog retrieval system with habit-based concepts and moment visualization. In: The Fourteenth NTCIR conference (NTCIR-14) (2019)
27. Wilson, G., Jones, D., Schofield, P., Martin, D.J.: The use of a wearable camera to explore daily functioning of older adults living with persistent pain: Methodological reflections and recommendations. *Journal of Rehabilitation and Assistive Technologies Engineering* **5**, 2055668318765411 (2018). <https://doi.org/10.1177/2055668318765411>
28. Xia, L., Ma, Y., Fan, W.: Vtir at the ntcir-12 2016 lifelog semantic access task. In: Proceedings of NTCIR-12, Tokyo, Japan (2016)
29. Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., Torralba, A.: Places: A 10 million image database for scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017)
30. Zhou, Q., Wang, D., Mhurchu, C.N., Gurrin, C., Zhou, J., Cheng, Y., Wang, H.: The use of wearable cameras in assessing children’s dietary intake and behaviours in china. *Appetite* **139**, 1 – 7 (2019)