Overview of the NTCIR-14 Lifelog-3 Task

Cathal Gurrin[1], Hideo Joho[2], Frank Hopfgartner[3], Liting Zhou[1], Van-Tu Ninh[1], Tu-Khiem Le[1], Rami Albatal[1], Duc-Tien Dang-Nguyen[4], and Graham Healy[1]

[1] Dublin City University, Ireland
[2] University of Tsukuba, Japan
[3] University of Sheffield, UK
[4] University of Bergen, Norway

Abstract. Lifelog-3 was the third instance of the lifelog task at NTCIR. At NTCIR-14, 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 activities of participating teams who took part in the challenges and we suggest next steps for the community.

Keywords: Lifelog · Information Retrieval · Test Collection

1 Introduction

NTCIR-14 hosted the third running of the Lifelog task. Over the three iterations of the task, from NTCIR-12 [10], NTCIR-13 [11] and this year, we note that nearly 20 participating research groups have taken part in 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 returning to the definition proposed by Dodge and Kitchin [6], 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 task was initially proposed because the organisers identified that technological progress had resulted in lifelogging becoming a potentially normative activity, thereby necessitating the development of new forms of personal data analytics and retrieval that are designed to operate on multimodal lifelog data. Additionally, the organisers note recent efforts to employ lifelogging, summarised in [4], as a means of supporting human memory [13] or facilitating large-scale epidemiological studies in healthcare [21], lifestyle monitoring [23], diet/obesity monitoring [25], or for exploring societal issues such as privacy-related concerns [14] and behaviour analysis [7].

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, 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.

We will now describe each task in detail.

2.1 LSAT SubTask

The LSAT subtask was a known-item search task applied over lifelog data. In this subtask, the participants had to retrieve a number of specific moments in a lifelogger's life in response to a query topic. We consider moments to be semantic events, or activities that happened at least once in the dataset. The task can best be compared to a known-item search task with one (or more) relevant items per topic. 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 SubTask

The aim of this subtask was to develop new approaches to the annotation of multimodal lifelog data in terms of activities of daily living. 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 where every image was annotated with one-or-more ground truth activity labels.

2.3 LIT SubTask

The LIT subtask was exploratory in nature and the aim of this subtask 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 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, the organisers prepared a new test collection that was specifically designed for the task and with a view to supporting future research into dietary [21] consumption of individuals. We developed this dataset following the process described in [5], with the following requirements in mind:

- To balance the size of the collection between being small enough to encourage participation and being large enough to provide challenging tasks.
- To include rich, multimodal lifelog data, gathered in free-living environments by a number of individuals, which can support many applications from ad-hoc retrieval to activity analytics and insight generation.
- To lower barriers-to-participation by including sufficient metadata, such as the visual annotations of visual content.
- To apply the principles of privacy-by-design [2] when creating the test collection, because personal sensor data (especially camera or audio data) carries privacy concerns [8], [14], [19].
- To include realistic topics representing real-world information needs of varying degrees of difficulty for various sub-tasks.

These requirements (refined from previous NTCIR-lifelog tasks) guided the test collection generation process.

3.1 Data Gathering Process

As with previous NTCIR-Lifelog tasks, the data was gathered by a number of lifeloggers (in this case, two) who wore the lifelogging devices and gathered biometric data for most (or all) of the waking hours in the day. One lifelogger gathered one month of data and one lifelogger gathered two weeks of data. The lifeloggers wore an OMG Autographer passive-capture wearable camera clipped to clothing or worn on a lanyard around the neck which captured images (from the wearer's viewpoint) and operated for 12-14 hours per day (1,250 - 4,500 images per day - depending on capture frequency or length of waking day). Additionally mobile apps gathered locations, physical movements and a record of music listening. Finally, additional wearable sensors provided health and wellness

Fig. 1. Examples of Wearable Camera Images from the Test Collection

data from continual heart-rate monitors, continuous (15-minute interval) blood glucose monitors, along with manual annotations of food and drink consumption.

Following the data gathering process, there were a number of steps (the same as in previous editions of the lifelog task) that were taken to ensure that test collection was both as realistic as possible, and took into account sensitivities associated with personal data:

- **Temporal Alignment.** All data was temporally aligned to UTC time.
- **Data Filtering.** Given the personal nature of lifelog data, it was necessary to allow the lifeloggers to remove any lifelog data that they may have been unwilling to share.
- **Privacy Protection.** Privacy-by-design [2] was one of the requirements for the test collection. Consequently, faces and screens were blurred and every image was also resized down to 1024×768 resolution which had the additional effect of rendering most textual content illegible.

3.2 Details of the Dataset

The data consists of a medium-sized collection of multimodal lifelog data over 42 days by the two lifeloggers. The contribution of this dataset over previously released datasets was the inclusion of additional biometric data, a manual diet log and the inclusion of conventional photos. In most cases the activities of the lifeloggers were separate and they did not meet. However on a small number of occasions the lifeloggers appeared in data of each other. The data consists of:

- **Multimedia Content**. Wearable camera images captured at a rate of about two images per minute and worn from breakfast to sleep. Accompanying this image data was a time-stamped record of music listening activities sourced

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

from `Last.FM`[5] and an archive of all conventional (active-capture) digital photos taken by the lifelogger.

- **Biometrics Data**. Using the FitBit fitness trackers[6], the lifeloggers gathered 24×7 heart rate, calorie burn and steps. In addition, continuous blood glucose monitoring captured readings every 15 minutes using the Freestyle Libre wearable sensor[7].
- **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 app[8], along with a time-stamped diet-log of all food and drink consumed.
- **Enhancements to the Data**. The wearable camera images were annotated with the outputs of a visual concept detector, which provided three types of outputs (Attributes, Categories and Concepts). Two visual concepts which include attributes and categories of the place in the image are extracted using PlacesCNN [24]. The remaining one is detected object category and its bounding box extracted by using Faster R-CNN [20] trained on MSCOCO dataset [16].

3.3 Topics

The LSAT task includes 24 topics with pooled relevance judgements. These LSAT topics were evaluated in terms of traditional Information Retrieval effectiveness measurements such as Precision, RelRet and MAP. An example of an LSAT topic is included as Figure 2. For the full list of the topics see Table 2.

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

[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/

TITLE: Ice Cream by the Sea
DESCRIPTION: Find the moment when U_1 *was eating ice cream beside the sea.*
NARRATIVE: To be relevant, the moment must show both the ice cream with cone in the hand of u1 as well as the sea clearly visible. Any moments by the sea, or eating an ice cream which do not occur together are not considered to be relevant.
EXAMPLES OF RELEVANT IMAGES FOUND BY PARTICIPANTS"

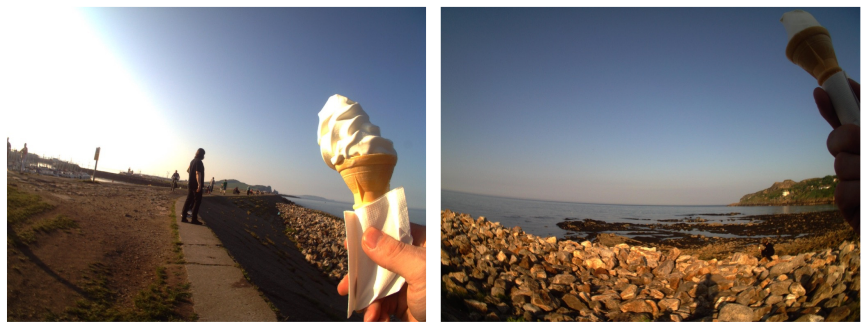

Fig. 2. LSAT topic example, including example results.

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

Topic Titles		
Ice cream by the Sea	Eating Fast Food	A New TV
Going Home by Train	Photograph of a Bridge	In a Toyshop
7* Hotel	Buying a Guitar	Empty Shop
Card Shopping	Croissant	Coffee and Scone for Breakfast
Cooking a BBQ	Flight Check-in	Mirror
Meeting with a Lifelogger	Seeking Food in a Fridge	Car Sales Showroom
Watching Football	Coffee with Friends	Dogs
Eating at the desk	Walking Home from Work	Crossing a Bridge

relevant image content from the collection. A full list of topics is available from the NTCIR-14 website[9] and replicated at the URL in the footnote[10].

For the LADT (Activity Detection) subtask, there were sixteen types of activities defined for annotation. 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 are:

- **traveling**: travelling (car, bus, boat, airplane, train, etc)

[9] http://research.nii.ac.jp/ntcir/ntcir-14/
[10] NTCIR-14 - Lifelog-3 Topics - http://ntcir-lifelog.computing.dcu.ie/resources/NTCIR-14-Lifelog-SubTask1-Topics-English.xml

- **face-to-face interacting**: face-to-face interaction with people at home or in the workplace (excluding social interactions)
- **using a computer**: using desktop computer / laptop / tablet / smartphone
- **cooking**: preparing meals (include making tea or coffee) at any location
- **eating**: eating meals in any location, but not including moments when drinking alone.
- **time with children**: taking care of children / playing with children
- **houseworking**: working in the home (e.g. cleaning, gardening)
- **relaxing**: relaxing at home (e.g. TV, having a drink)
- **reading**: reading any form of paper
- **socialising**: socialising outside the home or office
- **praying**: praying / worshipping / meditating
- **shopping**: shopping in a physical shop (not online)
- **gaming**: playing computer games
- **physical activities**: physical activities / sports (walking, playing sports, cycling, rowing, etc)
- **creative activities**: creative endeavours (writing, art, music)
- **other activities**: any other activity not represented by the fifteen labels above.

Each image can be tagged as belonging to one or more activities and the 'other activities' category was designed to take into account all activities that were not in the other fifteen.

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

3.4 Relevance Judgement and Scoring

Pooled binary relevance judgements were generated for all 24 LSAT topics. Scoring for the LSAT sub-task was calculated using the ubiquitous trec_eval toolkit [1]. A manually generated pooled ground-truth was generated for every topic, which formed the input for trec_eval programme. The pooling was done over the entire submissions from all official runs for the LSAT sub-task. Two custom applications were developed to support both the LSAT and LADT evaluation processes.

For the LADT topics/labels, a manual relevance judgement was 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 of the thirteen 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 managed to submit 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. We will now summarise the approaches taken by the teams.

NTU (Taiwan) took part in both the LSAT and LADT Tasks [9]. For the LSAT task, the NTU team developed an interactive lifelog retrieval system that automatically suggested to the user, a list of candidate query words and adopted a probabilistic relevance-based ranking function for retrieval. They enhanced the official concept annotations by applying the Google Cloud Vision API[11] 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 [22]. 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 [17]. 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.

[11] Google Cloud Vision API - https://cloud.google.com/vision/

Table 3. 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

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 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 lifelog data [18]. 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. A list of important difficulties were compiled from the user study and proposed as a set of requirements for future interactive lifelog retrieval systems.

It can be seen from Table 3 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. Without additional teams, there is little further analysis that we can do at this point.

4.2 LADT Task

The NTU group (Taiwan) took part in the LADT task [9] 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 too part, no comparison is possible between participants. Readers are referred to the NTU paper [9] for details of their different runs and the comparative performance of these.

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 [17], 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 non-visual data as well as one for visual 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 clearly showed the activities of the two users in a highly visual manner.

5 Learnings & Future Plans

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. We can summarise the learnings from this task as follows:

- After the previous NTCIR lifelog tasks, we still note that there is no standardised approach to retrieval of lifelog data, however, we do notice a number of emerging approaches that show promise. Firstly, the utilisation of **additional visual concept detectors** is considered a positive addition. Likewise we note the integration of **external WWW content** in many approaches. Finally, the lexical gap between user queries and concept annotations suggest that an **term expansion effort** is needed, and the current consideration is that this could be achieved using word embedding.
- Three of the four groups participating in the LSAT sub-task built **interactive retrieval systems** for lifelog data, highlighting the belief of the participants in the importance of the user in the retrieval process.
- The LSAT task is a valuable task and it continued to attract the majority of participants. This task is superseded by two related collaborative benchmarking activities, the Lifelog Search Challenge (LSC) [12], and the ImageCLEF Lifelog task[3].

6 Conclusion

In this paper, we described the data and the activities from the Lifelog-3 core-task at NTCIR-14. There were three sub-tasks prepared for this year. 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 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. 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 Science Foundation Ireland under grant number SFI/12/RC/2289 and Irish Research Council (IRC) under Grant Number GOIPG/2016/741. 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. Cavoukian, A.: Privacy by design: The 7 foundational principles. implementation and mapping of fair information practices. Information and Privacy Commissioner of Ontario, Canada (2010)
3. 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. org¡ http://ceur-ws. org¿, Avignon, France (September 10-14 2018) (2018)
4. Dang-Nguyen, D.T., Riegler, M., Zhou, L., Gurrin, C.: Challenges and Opportunities within Personal Life Archives. In: ACM International Conference on Multimedia Retrieval (2018)
5. 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)

6. 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
7. 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
8. Ferdous, M.S., Chowdhury, S., Jose, J.M.: Analysing privacy in visual lifelogging. Pervasive and Mobile Computing **"40**, 430 – 449 (2017)
9. 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)
10. 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)
11. 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)
12. 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)
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, 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
17. 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)
18. 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 nticr-14 lifelog-3 semantic access task. In: The Fourteenth NTCIR conference (NTCIR-14) (2019)
19. O'Hara, K.: Narcissus to a man: Lifelogging, technology and the normativity of truth (September 2010), event Dates: 16th September 2010
20. 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

21. 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
22. Suzuki, T., Ikeda, D.: Smart lifelog retrieval system with habit-based concepts and moment visualization. In: The Fourteenth NTCIR conference (NTCIR-14) (2019)
23. 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
24. 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)
25. 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)