

A Knowledge Graph-based System for Retrieval of Lifelog Data

Luca Rossetto, Matthias Baumgartner, Narges Ashena, Florian Ruosch,
Romana Pernisch, and Abraham Bernstein

Department of Informatics, University of Zurich, Zurich, Switzerland
{lastname}@ifi.uzh.ch

Abstract. Lifelogging is a phenomenon where practitioners record an increasing part of their subjective daily experience with the aim of later being able to use these recordings as a memory aid or basis for data-driven self improvement. The resulting lifelogs are, therefore, only useful if the lifeloggers have efficient ways to search through them. The logs are inherently multi-modal and semi structured, combining data from several sources, such as cameras and other wearable physical as well as virtual sensors, so representing the data in a graph structure can effectively capture all produced interrelations. Since annotating each entry with a sufficiently large semantic context is infeasible, either manually or automatically, alternatives must be found to capture the higher level semantics. In this paper, we demonstrate LifeGraph, a first approach of creating a Knowledge Graph-based lifelog representation and retrieval solution, able of capturing a lifelog in a graph structure and augmenting it with external information to aid with the association of higher-level semantic information.

Keywords: Lifelogging, Lifelog Retrieval, Multi-modal Retrieval, Multi-modal Graphs

1 Introduction

Various technological advances over the last few decades have led to a rapid increase in possibilities for mobile sensing, processing, and storing of various forms of data. One of the many consequences of these possibilities was the emergence of the quantified-self movement as well as the Lifelogging phenomenon. Lifelogging describes the act of capturing aspects of one's personal experience using cameras to continuously record one's point of view, sensors to quantify bio-feedback, or properties of one's immediate surroundings, purely software-based means to analyze one's interaction with the digital environment as well as more traditional methods, such as entries in a personal diary. Thus, the resulting lifelog is

Copyright © 2020 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

inherently multi-modal and can exhibit a high degree of interconnectivity. An appropriate, but currently under-explored approach, is, therefore, to represent a lifelog as a knowledge graph.

Lifelogs are commonly used as a memory aid or to monitor one’s own life in order to improve or maintain health. In order for these logs to be useful, it is paramount that lifeloggers have the possibility to efficiently search within the logs which becomes increasingly difficult due to the growing volume and diversity of available data. To catalyze research in the area of lifelog retrieval, a competitive benchmark called the Lifelog Search Challenge (LSC) [3] was established in 2018 as an annual workshop to evaluate interactive lifelog retrieval systems. In this paper, we demonstrate an interactive knowledge graph-based lifelog retrieval system called LifeGraph [5], which was first introduced in the context of the 2020 edition of the LSC. The following provides an overview of how the graph was constructed, how the system uses it to answer queries, and how users can expect to be able to interact with the system itself.

2 Constructing LifeGraph

The basis for LifeGraph is formed by the lifelog dataset made available in the context of LSC 2020, which consists of 114 days of lifelogging data made up of roughly 200k images taken by a wearable camera with accompanying metadata as well as various sensor data. Ideally, each entry in a lifelog, independent of its type and method of representation, would have rich semantic annotations describing not only its content but also placing it into a larger context. Since generating such annotations manually is infeasible and doing so automatically is beyond what is currently possible, alternatives need to be used to make the collected data searchable and, hence, usable. To overcome this limitation, we construct a graph structure which captures the relations between higher level semantic concepts and lower level labels, such as the type of environment or the presence of everyday objects within an image, both of which can be reliably detected with currently available technology [9,4].

A challenge of constructing LifeGraph is to link high-level search entities with low-level image features. We use COEL (Classification of Everyday Living) [1] as a starting point which provides a taxonomy of everyday actions, such as housework, sports, or food related activities. In order to combine COEL with detectable entities, we defined a mapping to Wikidata [8]. It was created semi-automatically, based on the similarity to Wikidata entity labels, whereas ambiguities were resolved manually.

To connect the mapped concepts to lower-level semantic image labels we combined them with additional data from Wikidata. For this, we created a graph initially consisting of the Wikidata entities mapped to COEL as well as all detectable entities. We then successively retrieved all entities from Wikidata that can be connected to at least two entities in the present graph, and added them to the graph. This procedure is repeated three times, and two additional times considering only class relations (`:instance-of`, `:subclassOf`). Finally, we dis-

carded disconnected components of the graph since they failed to link image labels to high-level concepts. We further removed entities intended for navigational or internal purpose, such as disambiguation or template pages.

We modeled the metadata which was already part of the dataset provided for LSC in a simple and flat way following the schema explained in [5]. Each metadata entry is associated with an image based on time. Using either an object or datatype property, we link all the metadata information with the image in the graph. Contrary to what was stated in [5], each image is linked to a day resource. This enables us to link day-specific information directly to the day resource instead of repeating this information for all ca. 2000 images of that day. Before populating the schema, we interpolated some aspects of the metadata such as the provided semantic location labels (“work” vs. “home”) but also detected objects [4] and places [9]. If the same concept is interrupted by a sufficiently small gap, we add associations to this concept to all entries in the gap.

To associate images with objects and places, we then directly establish a link between images and Wikidata entries using either the `:detected` or `:detected_place` property.

3 Exploring LifeGraph

The browser-based LifeGraph user interface is a modified version of *vitriivr-ng* [2] – the UI of the *vitriivr* [7] content-based multimedia retrieval stack – which has already been successfully used for effective retrieval of lifelog data [6]. Apart from the user interface, LifeGraph does not share any logic with the *vitriivr* stack. For the sake of simplicity, a user can specify queries by selecting nodes from the graph, representing the concepts in which they are interested, using a simple auto-completing text box. This query formulation scheme offers a reasonable compromise between specificity and efficiency while not requiring any technical expertise on the user side. Once specified, a query is sent to a back-end which performs its execution and the subsequent scoring of the results. The query is evaluated by traversing the graph, searching for paths from each of the specified start nodes to all *log entries* which could be part of the result. The maximum length of this path is continuously increased until an empirically determined threshold is reached, either in path length or number of potential results. Each log entry found in this way is scored by the sum of the inverse of the path lengths leading to it, normalized by the number of start points in the query. This means the shorter the distance to the highest number of start points, the higher the score of the retrieved item. These aggregated results are then returned to the UI, where they are presented as either a list of log entries sorted by score or a list of entries grouped by day, sorted by time within each. Since the user is supposed to be able to browse through the retrieved results in order to identify the desired elements, the inevitable false-positives produced by some of the underlying detectors are less of a problem than missing detections or false-negatives, since a large number of false-positives will cause users to spend an unnecessarily large amount of time browsing, while too many false-negatives could prevent them from finding

the desired result at all. All log entries are accompanied by some metadata, including technical information, like its originating time and date. Additionally, entries include lifelog specific semantic information, such as manual semantic annotations or bio feedback properties, e.g., heart-rate, number of steps taken within a specific interval, etc. The UI uses these metadata to provide additional filtering options to a user, excluding entries from the result if they do not match Boolean criteria relevant to the query. Since these filters can only reduce the size of the result set, they can be evaluated directly in the UI without the need for further back-end communication or interaction with the underlying graph.

4 Demonstration

During the demonstration, participants will have the opportunity to use the system in a setting analogous to the evaluation during LSC. They will be provided with a textual description of a point in the life of the lifelogger whose experiences serve as foundation for the dataset. They can then try to find the described life events using LifeGraphs retrieval and browsing functionality. A screenshot illustrating the system in action is shown in Figure 1.

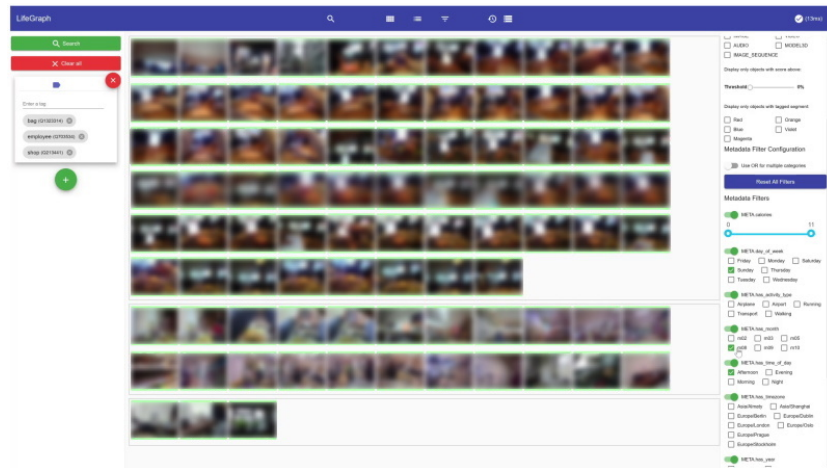


Fig. 1. Screenshot of the LifeGraph retrieval system in action. The shown user interface is a modified version of *vitivr-ng* [2]. The retrieved results have been made unrecognizable to comply with the privacy protection requirements of the dataset.

5 Conclusion and Outlook

In this paper, we demonstrated LifeGraph, a first step into the direction of using a Knowledge Graph in the context of the emerging phenomenon which is

Lifelogging. Due to the inherent multi-modality and semi-structured nature of lifelogs, we see great potential in the interaction of the Semantic Web and the Lifelogging communities and are looking forward to future developments in this area.

References

1. Bruton, P., Langford, J., Reed, M., Snelling, D.: Classification of everyday living version 1.0 (2019), <https://docs.oasis-open.org/coel/COEL/v1.0/os/COEL-v1.0-os.html>, last updated 23 January 2019
2. Gasser, R., Rossetto, L., Schuldt, H.: Towards an all-purpose content-based multimedia information retrieval system. arXiv preprint arXiv:1902.03878 (2019), <https://arxiv.org/abs/1902.03878>
3. Gurrin, C., Le, T.K., Ninh, V.T., Dang-Nguyen, D.T., Jónsson, B.P., Lokoš, J., Hürst, W., Tran, M.T., Schoeffmann, K.: Introduction to the Third Annual Lifelog Search Challenge (LSC'20). In: Proceedings of the 2020 International Conference on Multimedia Retrieval. pp. 584–585 (2020). <https://doi.org/10.1145/3372278.3388043>
4. Redmon, J., Farhadi, A.: YOLO9000: better, faster, stronger. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017. pp. 6517–6525. IEEE Computer Society (2017). <https://doi.org/10.1109/CVPR.2017.690>
5. Rossetto, L., Baumgartner, M., Ashena, N., Ruosch, F., Pernischová, R., Bernstein, A.: Lifegraph: A knowledge graph for lifelogs. In: Proceedings of the Third Annual Workshop on Lifelog Search Challenge. pp. 13–17 (2020). <https://doi.org/10.1145/3379172.3391717>
6. Rossetto, L., Gasser, R., Heller, S., Amiri Parian, M., Schuldt, H.: Retrieval of structured and unstructured data with vitriv. In: Proceedings of the ACM Workshop on Lifelog Search Challenge. pp. 27–31 (2019). <https://doi.org/10.1145/3326460.3329160>
7. Rossetto, L., Giangreco, I., Tanase, C., Schuldt, H.: vitriv: A flexible retrieval stack supporting multiple query modes for searching in multimedia collections. In: Proceedings of the 24th ACM international conference on Multimedia. pp. 1183–1186 (2016). <https://doi.org/doi.org/10.1145/2964284.2973797>
8. Vrandečić, D., Krötzsch, M.: Wikidata: a free collaborative knowledgebase. Communications of the ACM **57**(10), 78–85 (2014). <https://doi.org/10.1145/2629489>
9. 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). <https://doi.org/10.1109/TPAMI.2017.2723009>