

# LifeGraph 5: A SPARQL-ing User Interface for Advanced Multimodal Lifelog Querying

Florian Ruosch<sup>1</sup>, Luca Rossetto<sup>2</sup>

<sup>1</sup>FPR Consulting, Zurich, Switzerland

<sup>2</sup>Dublin City University, Dublin, Ireland

## Abstract

Lifelogging, the practice of recording parts of the subjective daily life, generates rich, multimodal data, but poses significant challenges for efficient retrieval. Building upon the *LifeGraph* series of knowledge graph-based lifelog retrieval systems, this paper presents the fifth iteration, which introduces a novel user interface to facilitate intuitive and powerful querying of lifelogs from multimodal knowledge graphs. We showcase how this frontend, powered by our custom MediaGraph store *MeGraS*, seamlessly exposes and leverages SPARQL capabilities. Through interactive demonstration scenarios, we illustrate how users can easily construct complex and expressive queries that also include advanced features such as similarity-based search, near-duplicate detection, and dynamic content extraction, all the while using native SPARQL syntax. This work highlights *LifeGraph 5*'s user-centric design and *MeGraS*'s role in bridging gaps between complex knowledge graph operations and accessible multimodal lifelog exploration.

## Keywords

Multimodal Knowledge Graph, Graph-based Retrieval, Interactive Retrieval, Query Builder, User Interface

## 1. Introduction

Lifelogging [1] is the practice of continuously capturing an individual's subjective daily experiences through various means of recording. Prominent data sources are wearable cameras to collect first-person-view images, to track the GPS location, sensors to record information such as heart rate, but also lists of consumed media and their associated metadata. This generates vast and inherently multimodal datasets, holding immense potential for memory augmentation and data-driven insights. However, the sheer volume, diversity, and semi-structured nature of lifelog data pose significant challenges for efficient and effective retrieval. Multimodal knowledge graphs have emerged as a powerful tool for structuring such complex and interconnected information, enabling sophisticated querying and semantic understanding: the *LifeGraph* series [2, 3, 4, 5].

We developed *LifeGraph 5* [6] for the 8th Lifelog Search Challenge (LSC'25) [7], an annual competition for document retrieval and question answering on a large multimodal dataset of lifelogs. It represents the latest iteration of our system, which pushes for enhanced query capabilities by extending SPARQL with concepts like implicit and derived relations. Having come a long way since the first demonstration of *LifeGraph* [8], a core component of *LifeGraph 5*'s advancements is the newly designed user interface and the underlying custom MediaGraph Store (*MeGraS* [9]), which enables access to the contents of the documents in the multimodal knowledge graph.

In this demo, we focus on the interactive aspects and the user experience. We aim to showcase how our new frontend transforms complex querying using extended SPARQL into an intuitive and accessible process for users. Through this, we will highlight how *LifeGraph 5* enables the formulation of advanced queries that leverage multimodal features and non-materialized relations, thereby unlocking deeper insights into the lifelog data.

ISWC 2025 Companion Volume, November 2–6, 2025, Nara, Japan

✉ ruosch@ifi.uzh.ch (F. Ruosch); luca.rossetto@dcu.ie (L. Rossetto)

id 0000-0002-0257-3318 (F. Ruosch); 0000-0002-5389-9465 (L. Rossetto)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

## 2. LifeGraph 5

This section first describes our custom multimodal knowledge graph store, *MeGraS*, then explains our multimodal knowledge graph *LifeGraph* and, finally, its user interface.

*LifeGraph 5* represents a significant evolution from the foundational work of *LifeGraph 1* [8]. Our development has been guided by the requirements of the Lifelog Search Challenges [7] (LSC), an annual competition wherein participants aim to efficiently retrieve documents and answer questions from a multimodal, large-scale lifelog dataset. It is comprised of over 700,000 images taken from a first-person point of view with a wearable camera, additional data from sensors, and rich metadata, presenting a significant challenge for traditional retrieval systems. The specific tasks in the LSC range from known-item search, where participants must find a specific item, to question answering about events or objects, and complex ad-hoc searches, where as many images as possible should be retrieved that fit the given description.

While *LifeGraph 1* established the foundation and the principle of using multimodal knowledge graphs for lifelog retrieval, *LifeGraph 5* introduces key architectural and user-facing advancements to handle the multimodal nature of the LSC more effectively. The most significant innovations can be summarized as follows:

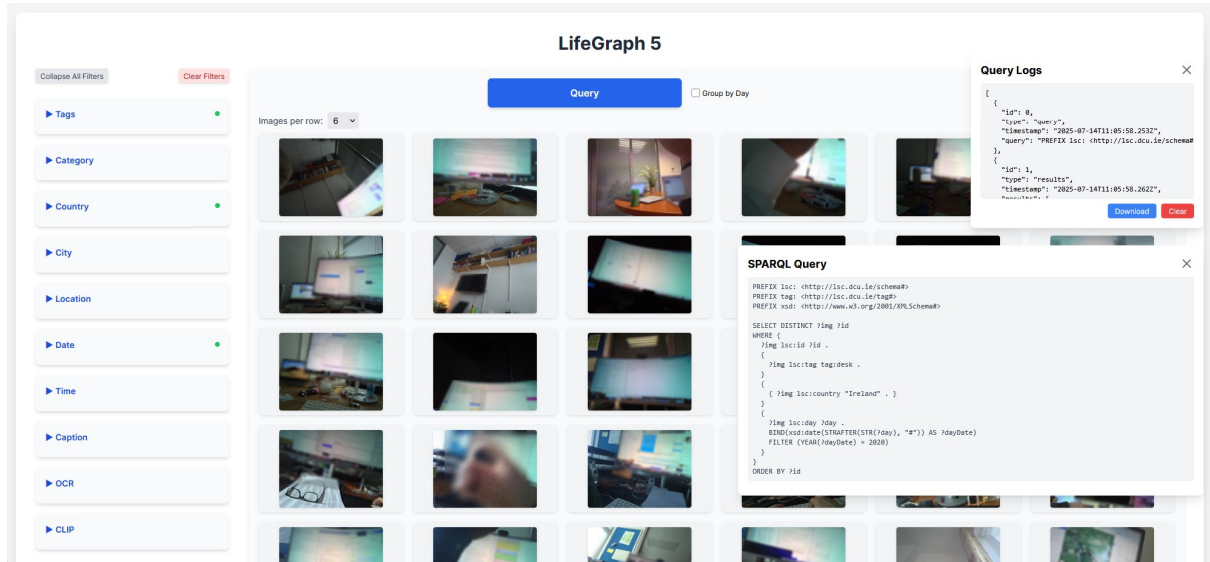
- **Custom Multimodal Knowledge Graph Store:** Unlike its predecessors, which treated multimedia as external, *LifeGraph 5* is built upon a novel, custom-designed backend called the MediaGraph Store (*MeGraS*). It elevates multimedia documents to “first-class citizens” within the knowledge graph, allowing the query engine to directly access and process their content. This is a fundamental architectural shift that is critical for meeting the demands of the LSC, where content-based retrieval, such as finding specific objects or actions within an image, is paramount.
- **Extended SPARQL Capabilities:** *MeGraS* directly extends the SPARQL query language with native support for advanced operations. This is vital for the LSC, as it enables complex functions like k-nearest neighbor searches, which can find semantically similar images based on their vector embeddings, and the ability to detect near-duplicates. These capabilities provide a powerful toolset for participants to tackle the challenge’s retrieval tasks with greater precision and efficiency.
- **Dynamic and Unmaterialized Relations:** The system can handle derived and implicit relations that are not necessarily materialized in the graph but computed at query time and persisted, if necessary. This provides a more flexible data model for rapid exploration.
- **Intuitive User Interface:** *LifeGraph 5* features a completely newly designed frontend that transforms complex, SPARQL-based querying into an intuitive process for users. This user-centric design is particularly beneficial in the time-constrained environment of the LSC, as it allows participants to quickly build and refine queries using a series of selectors and filters. The interface renders the SPARQL query in real-time, providing transparency into the underlying logic without requiring the user to manually write code.

### 2.1. MediaGraph Store

*MeGraS* [9], our MediaGraph Store, serves as both the RDF graph store as well as the query engine and aims to elevate multimedia documents to first-class citizens in knowledge graphs. Contrary to traditional graph stores, it makes the linked documents available through assigned URIs and gives its query engine access to their contents, allowing for advanced manipulation such as segmentation or feature extraction. It also supports the Unified Multimedia Segmentation scheme [10] which allows the direct addressing of arbitrary parts of these documents.

Interactions with the graph happen through the RESTful API, but *MeGraS* also offers a SPARQL endpoint leveraging Apache Jena. Furthermore, it natively supports vector operations such as k-nearest neighbor or similarity search, which is also available through SPARQL. *MeGraS* also handles derived and implicit information, which is not necessarily materialized and can be computed at query time.

Derived relations may be available in the graph ex ante, but are not required to be, and always have a graph node as the subject and a literal as the object. They result from predefined functions, such as



**Figure 1:** A screenshot of the LifeGraph 5 user interface in action.

computing the embedding of an image or extracting features such as visible text. If a requested derived relation is not available at query time, *MeGraS* runs the functions and persists its output, avoiding the need for repeated calculations.

Implicit relations are defined to be always between two graph nodes, but never materialized, and are inferred from other information that is available to the query engine. They are not persisted as they may depend on other nodes in the graph and can change based on additions or removals. Examples include k-nearest neighbor, spatial (e.g., behind, above), and temporal (e.g., during, after) relations.

All these mechanisms allow for advanced and more expressive querying, purely through SPARQL. Combined with the capability of handling tens of millions of triples through the integrated PostgreSQL database, this makes *MeGraS* an optimal backend for our multimodal lifelog retrieval.

## 2.2. Ontology of LifeGraph

The structure of the multimodal knowledge graph that contains the data of *LifeGraph 5* is described in a dedicated ontology.<sup>1</sup> It serves as the formal schema for the image-related information within the LSC dataset, and its core is the Image class representing individual images. For the object properties, the ontology links images to instances of days for temporal facts and to tags to associate them with descriptive keywords that do not carry any semantics. The data properties capture literal attributes. They consist of the automatically extracted OCR (Optical Character Recognition), the running numbering, the identifier, and additional annotations like the VLM-generated caption or the manually curated category. Furthermore, they also include spatial references such as location name, city, and country. Finally, we also employ a custom float vector datatype to precisely represent numerical vector data (e.g., CLIP [11] image embeddings). Adherence to this ontological structure ensures that *LifeGraph 5*'s data is consistent, well-defined, and semantically rich.

## 2.3. User Interface

This subsection explores *LifeGraph 5*'s frontend in detail. First, we focus on how the SPARQL query is constructed. Then, we describe the results display and its advanced functionalities.

<sup>1</sup><https://github.com/MediaGraphOrg/LifeGraph5/blob/main/LifeGraph5.owl>

### 2.3.1. Query Builder

Figure 1 shows an overview of the user interface and its components. The SPARQL query, created based on the filters and selectors on the left, is rendered in real time in the dedicated area on the right. The criteria represented in the query and their order can be configured for the user interface. By default, the entire list is enabled, which can be divided into two classes: selectors (for tag, category, country, city, and location) and filters (date, time, caption, OCR, and CLIP).

The selectors allow for searching for and selecting a value, which is then used as the object for the associated predicate in the query. Their design can be seen on the left in Figure 2: the search bar, the scrollable list, and the selected objects. Tags are manually annotated labels for the images, based on their content (e.g., bedroom, building). Categories, meanwhile, are associated with the type of location of the image (e.g., airport terminal, bakery). Country, city, and location are all inferred from the location metadata of the image and then mapped to entities in Wikidata based on the smallest distance.

The filters can be distinguished into two types: simple data-based search and matching possibly unmaterialized relations. The dates are all extracted from the metadata of the images and can be searched either with a range, by year, month, or weekday. The time filter also uses a range of start and end. The caption was generated using Vision-Language Models for *LifeGraph 4* [5] and can be full-text searched.

Most of the filtering can be done by constructing sequences of simple triple patterns, based on the selected criteria, of the following form: “?s lsc:predicate lsc:object .” This is shown in the SPARQL query area in Figure 1, and the active criteria are indicated on the left with a green dot. Furthermore, the query can also be specified to retrieve the dates for all results, allowing for hierarchically grouping them by year, month, and day, for easier access.

A more sophisticated query can be seen in Listing 1. Thereby, the CLIP [11] embedding of a textual description is computed and compared to the embedding of images using a custom SPARQL function. The resulting set is then limited to the eight nearest images in the embedding space.

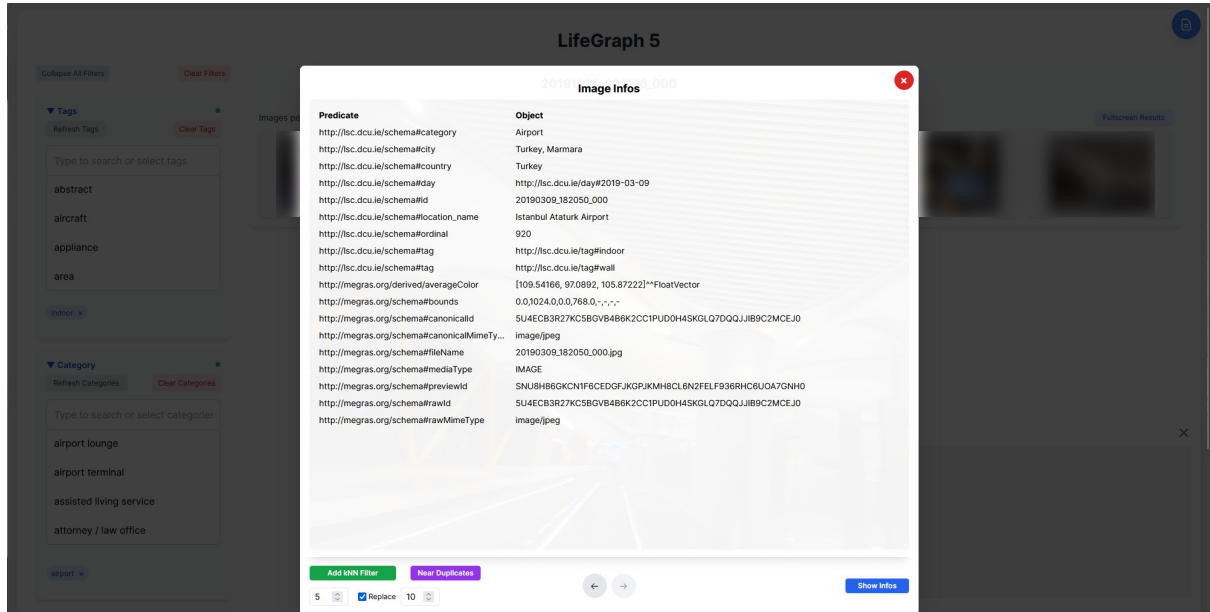
This showcases two important aspects leveraging *MeGraS*: the derived relation `clipEmbedding`, which may be unmaterialized and can be computed and persisted at query time, and the two custom functions `CLIP_TEXT` (computing the CLIP embedding of text) and `COSINE_SIM` (calculating the cosine distance of two vectors). Hence, we can achieve finding images matching a textual description based on their distance in the CLIP embedding space using pure SPARQL. Similarly, the OCR relation does not need to be precomputed but can be extracted at query time, if not yet available.

Listing 1: Finding the eight images for which the CLIP [11] embeddings are the most similar to the embedding of the textual description.

```
SELECT ?img
WHERE {
  BIND (megras:CLIP_TEXT("A man walking his dog on a rainy day.") as ?textVec)
  ?img derived:clipEmbedding ?clipVec .
  BIND (megras:COSINE_SIM(?textVec, ?clipVec) as ?cosSim)
}
ORDER BY DESC(?cosSim)
LIMIT 8
```

### 2.3.2. Results Display and Interaction

In Figure 1, the grid of results is visible, which allows for responsive and efficient exploration of the query results. Clicking on the preview of an image brings up the result overlay, as shown in Figure 2. The result overlay offers functionality for the exploration of the results set by navigating with either buttons or the arrow keys. Also, all triple relevant for the displayed image can be retrieved with the “Show Infos” button in the bottom right. Clicking on a line in the table of predicates and objects refines the query by adding the corresponding criterion, allowing for user feedback on the result set.



**Figure 2:** A screenshot of the result overlay of LifeGraph 5 showing all relevant triples from the multimodal knowledge graph for a retrieved image.

Furthermore, the buttons in the bottom left can be used for similarity search. It offers a CLIP-embedding-based k-nearest neighbor filter, whereby the k can be set dynamically (see Listing 2 for the pure SPARQL query). Again, the CLIP embedding does not need to have been materialized ex ante but can be computed and persisted at runtime with *MeGraS*. Likewise, a chosen number of near duplicates can be retrieved, whereby the relation is implicit and not materialized in the graph.

Listing 2: Detecting the implicit relation of k-nearest neighbors. In this example, k is equal to 5.

```
SELECT ?img
WHERE {
  lsc:imgID implicit:clip5nn ?img .
}
```

### 3. Demonstration

*LifeGraph 5* is open source, free to download,<sup>2</sup> and intended to be used together with *MeGraS* [9].<sup>3</sup> The latter can be compiled from source or downloaded as a Docker container.<sup>4</sup> Furthermore, a demo video is available that showcases *LifeGraph 5*'s capabilities.<sup>5</sup>

During the demonstration, participants will have the opportunity to interact with *LifeGraph 5*. The setting will be analogous to that during the evaluation of the Lifelog Search Challenge [7]: participants will have to find images based on textual descriptions and answer questions about events in the lifelog data. The user interface can be used to explore, browse, and retrieve the relevant images.

Participants will engage in an interactive session demonstrating the system's core features. They will use the various selectors and filters on the left-hand side to construct a query. The real-time generation of the corresponding SPARQL query will be visible throughout the process, providing a clear link between the user's actions and the underlying query logic. After execution, the results are displayed in

<sup>2</sup><https://github.com/MediaGraphOrg/LifeGraph5>

<sup>3</sup><https://megras.org>

<sup>4</sup><https://megras.org/docker>

<sup>5</sup><https://megras.org/2025iswcdemo>



a responsive grid. Participants can then select a specific item to bring up a detailed overlay showing the image and its node neighborhood. This also provides the ability to refine the query with new criteria based on the displayed information. The interactive process, combined with advanced features like similarity search and near-duplicate detection, will highlight how *LifeGraph 5* simplifies complex multimodal lifelog exploration.

**Illustrative Case Study: Who did I go to dinner with?** Imagine a user wants to find images from a specific event, recalling only a few details: “I went out to dinner in France in October 2019, but I cannot remember who I was with. I think we had seafood and wine.” Now, how can the user find the picture taken at that event to identify the person he was with?

- Step 1 The user interacts with the user interface’s selectors and filters. They set the date filter to Year: 2019 and Month: October. They also select Country: France, and use the selectors for Tags and Category, specifying wine and seafood restaurant, respectively.
- Step 2 As these criteria are selected, the *LifeGraph 5* frontend constructs the corresponding SPARQL query in real-time, rendering it in the dedicated query area. This live generation provides transparency into the underlying query logic.
- Step 3 After executing the query, the system returns a grid of relevant images. The user can select a specific image to look at to find instances of the event they were searching for. Selecting an item brings it up in its full resolution and will allow the user to identify the person on it, answering their question.

## 4. Conclusion

In this paper, we presented *LifeGraph 5* [6], a novel user interface designed to simplify and enhance the retrieval of lifelog data from multimodal knowledge graphs. By seamlessly integrating with our powerful MediaGraph Store (*MeGraS* [9]), it empowers users to perform sophisticated queries leveraging extended SPARQL capabilities. We demonstrated how this intuitive interface allows for complex knowledge graph operations such as similarity-based search, near-duplicate detection, and dynamic content extraction from documents.

*LifeGraph 5* combined with *MeGraS* represents a step forward in making knowledge-graph-based lifelog exploration more accessible and user-friendly. While this and previous iterations of *LifeGraph* highlighted the potential of our approach, it also exhibited areas for further optimization, particularly concerning query performance in extremely large datasets. Future work will focus on improving these backend limitations and further enhancing the user interface, potentially exploring natural-language-to-SPARQL interfaces such as NLQxform [12]. *LifeGraph 5* and *MeGraS* aim to serve as foundational tools for advancing research and practical applications in the complex and evolving landscape of multimodal knowledge management and retrieval.

## Acknowledgments

This work was partially funded by the Swiss National Science Foundation through Project “MediaGraph” (Grant Number 202125).

## Declaration on Generative AI

During the preparation of this work, the authors used Gemini and Grammarly for drafting content, improving writing style, as well as grammar and spell check. After using these tools, the authors reviewed and edited the content as needed, and they take full responsibility for the publication’s content.

## References

- [1] C. Gurrin, A. F. Smeaton, A. R. Doherty, Lifelogging: Personal big data, *Found. Trends Inf. Retr.* 8 (2014) 1–125. doi:10.1561/15000000033.
- [2] L. Rossetto, M. Baumgartner, N. Ashena, F. Ruosch, R. Pernischová, A. Bernstein, Lifegraph: A knowledge graph for lifelogs, in: *Proceedings of the Third Annual Workshop on Lifelog Search Challenge, LSC '20*, Association for Computing Machinery, New York, NY, USA, 2020, p. 13–17. doi:10.1145/3379172.3391717.
- [3] L. Rossetto, M. Baumgartner, R. Gasser, L. Heitz, R. Wang, A. Bernstein, Exploring graph-querying approaches in lifegraph, in: *Proceedings of the 4th Annual on Lifelog Search Challenge, LSC '21*, Association for Computing Machinery, New York, NY, USA, 2021, p. 7–10. URL: <https://doi.org/10.1145/3463948.3469068>.
- [4] L. Rossetto, O. Inel, S. Lange, F. Ruosch, R. Wang, A. Bernstein, Multi-mode clustering for graph-based lifelog retrieval, in: *Proceedings of the 6th Annual ACM Lifelog Search Challenge, LSC 2023*, Thessaloniki, Greece, June 12–15, 2023, ACM, 2023, pp. 36–40. doi:10.1145/3592573.3593102.
- [5] L. Rossetto, A. Kyriakou, S. Lange, F. Ruosch, R. Wang, K. Wardatzky, A. Bernstein, 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, LSC '24*, Association for Computing Machinery, New York, NY, USA, 2024, p. 88–92. doi:10.1145/3643489.3661127.
- [6] F. Ruosch, L. Rossetto, A SPARQL in the dark: Shining a light on multimodal lifelogs with LifeGraph 5, in: *Proceedings of the 8th Annual ACM Workshop on the Lifelog Search Challenge (LSC '25)*, LSC '25, ACM, New York, NY, USA, 2025. doi:10.1145/3729459.3748694.
- [7] C. Gurrin, L. Zhou, G. Healy, A. Tran, L. Rossetto, W. Bailer, D.-T. Dang-Nguyen, S. Hodges, B. Þór Jónsson, M.-T. Tran, K. Schöffmann, Introduction to the 8th annual lifelog search challenge, lsc'25, in: *Proceedings of the 2025 International Conference on Multimedia Retrieval, ICMR '25*, Association for Computing Machinery, New York, NY, USA, 2025, p. 2143–2144. doi:10.1145/3731715.3734579.
- [8] L. Rossetto, M. Baumgartner, N. Ashena, F. Ruosch, R. Pernischová, A. Bernstein, A knowledge graph-based system for retrieval of lifelog data, in: K. L. Taylor, R. S. Gonçalves, F. Lécué, J. Yan (Eds.), *Proceedings of the ISWC 2020 Demos and Industry Tracks: From Novel Ideas to Industrial Practice co-located with 19th International Semantic Web Conference (ISWC 2020)*, Globally online, November 1–6, 2020 (UTC), volume 2721 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2020, pp. 223–228. URL: <https://ceur-ws.org/Vol-2721/paper557.pdf>.
- [9] L. Rossetto, F. Ruosch, MeGraS: An Open-Source Store for Multimodal Knowledge Graphs, in: *Proceedings of the 33rd ACM International Conference on Multimedia (MM '25)*, MM '25, Association for Computing Machinery, New York, NY, USA, 2025. doi:10.1145/3746027.3756872.
- [10] J. Willi, A. Bernstein, L. Rossetto, Unified multimedia segmentation - A comprehensive model for uri-based media segment representation, *TGDK 2* (2024) 1:1–1:34. doi:10.4230/TGDK.2.3.1.
- [11] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, I. Sutskever, Learning transferable visual models from natural language supervision, in: M. Meila, T. Zhang (Eds.), *Proceedings of the 38th International Conference on Machine Learning, ICML 2021, 18–24 July 2021, Virtual Event*, volume 139 of *Proceedings of Machine Learning Research*, PMLR, 2021, pp. 8748–8763. URL: <http://proceedings.mlr.press/v139/radford21a.html>.
- [12] R. Wang, Z. Zhang, L. Rossetto, F. Ruosch, A. Bernstein, Nlqxform: A language model-based question to SPARQL transformer, in: D. Banerjee, R. Usbeck, N. Mihindukulasooriya, G. Singh, R. Mutharaju, P. Kapanipathi (Eds.), *Joint Proceedings of Scholarly QALD 2023 and SemREC 2023 co-located with 22nd International Semantic Web Conference ISWC 2023*, Athens, Greece, November 6–10, 2023, volume 3592 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2023. URL: <https://ceur-ws.org/Vol-3592/paper2.pdf>.