Personal Knowledge Graphs: Use Cases in e-learning Platforms

Eleni Ilkou

supervised by Prof. Dr. Wolfgang Nejdl

ilkou@l3s.de

L3S Research Center, Leibniz University Hannover

Hanover, Germany

ABSTRACT

Personal Knowledge Graphs (PKGs) are introduced by the semantic web community as small-sized user-centric knowledge graphs (KGs). PKGs fill the gap of personalised representation of user data and interests on the top of big, well-established encyclopedic KGs, such as DBpedia [21]. Inspired by the widely recent usage of PKGs in the medical domain to represent patient data, this PhD proposal aims to adopt a similar technique in the educational domain in e-learning platforms by deploying PKGs to represent users and learners. We propose a novel PKG development that relies on ontology and interlinks to Linked Open Data. Hence, adding the dimension of personalisation and explainability in users' featured data while respecting privacy. This research design is developed in two use cases: a collaborative search learning platform and an e-learning platform. Our preliminary results show that e-learning platforms can get benefited from our approach by providing personalised recommendations and more user and group-specific data.

CCS CONCEPTS

• **Information systems → Personalization**; **Collaborative search**; **Personalization**; • **Applied computing → Collaborative learning**; **E-learning**; **Distance learning**; • **Computing methodologies → Knowledge representation and reasoning**.

KEYWORDS

e-learning, collaborative search, collaborative learning, personalised knowledge graphs

1 INTRODUCTION

Nowadays, societies focus on the digital transformation of education and skill development in online learning platforms [8]. Millions

of learners use daily online learning platforms for their formal education, especially during the COVID-19 pandemic. The need for online teaching and lifelong learning tools has gained momentum. The same happens with online collaborative learning and search. Collaborative search happens when two or more people team up in a search task online and perform synchronous or asynchronous searches. Collaborative work online and collaborative learning are more important than ever; however, platforms supporting web collaboration lack semantic features, such as interconnections between the data and semantic recommendations. These platforms are mainly customised to facilitate learning applications and usually ignore the description and documentation of modelled concepts. As a result, current approaches cannot exploit common understanding encoded either in domain ontologies or knowledge graphs. At the same time, there is an increased need for systems with high personalised capabilities, personalised collaborative search [5], and more productive and impactful platforms that can support collaborative learning. Semantic web (SW) technologies can assist in this effort.

Knowledge Graphs (KGs), although well-established and widely used, such as the DBpedia and YAGO, do not directly contain personal information and are not usually designed to accommodate users' personal data. The same issue occurs also in domain KGs, such as in educational domain, where the need for personalisation is significant [13]. The recently suggested personal knowledge graphs (PKGs) [22] come to fill this gap by offering pocket-sized knowledge graphs related to users' interests. The same issue also occures in educational KGs where . PKGs in online learning environments can benefit users by providing personalised features and connecting their actions on the web with KGs and Linked Data.

The proposed line of work aims to assist the efforts of the SW community in setting the standards of new applications with deploying further research in SW technologies, such as the PKGs, in no classical computer science domains of application; in our case, the educational domain and the collaborative learning and e-learning sub-domains. Also, based on our knowledge, this is the first attempt to utilise PKGs for purely educational and learning purposes and have them developed in e-learning platforms. This work can assist researchers in both fields of education and computer science to create better personalised systems, and encourage developers in e-learning platforms to utilise SW technologies. This proposal aims to explore the syntax and semantics of PKGs in learning environments and increase the personalisation and explainability of the learning system's actions to users.

2 RESEARCH PROBLEM

The PhD thesis addresses the problem related to the personalisation of the web searches and to the enhancement of e-learning environments with personalised features linked to the SW. For

example in e-learning platforms the importance of personalised recommendations is increasing; the PKGs could assist in this effort an e-learning platform which uses a knowledge base for its content. Also, collaborative searching environments and collaborative learning usually offer features adopted from non-collaborative settings. Not many features are developed directly for collaborative learning environments, which creates a gap between the needs of collaborative environments and the actual technology offered. A solution could be the connection of the collaborative search with the Linked Data. This can occur by utilising PKGs to connect individuals' and group's search with the semantic web and offer a better collaborative and personalised experience.

The thesis is formulated around the analysis of the characteristics related to the deployment of PKGs for educational applications. However, we can especially classify our use cases in collaborative learning environments and e-learning platforms. Therefore, we formulate the research questions of the thesis as follows:

- RQ1: How to syntactically and semantically represent a PKG in the e-learning domain?
- RQ2: How can e-learning platforms offer better semantically enhanced personalised features, such as semantic recommendations, with the usage of PKGs?
- RQ3: How can collaborative search learning environments offer more personalised features with the usage of PKGs?
- RQ4: Can collaborative learning platforms offer better collaboration with the usage of PKGs?

3 STATE OF THE ART

3.1 Personal Knowledge Graphs

Personal or Personalised Knowledge Graphs (PKGs) are small graphs on top of KGs which contain user's related data. They are the natural complementary effort to address the challenges raised by the personal information management (PIM) [18], about users retrieving and organising personal information. PKGs are currently mostly used in the medical domain, on top of medical KGs to represent patients data [26].

In the broad domain, the works of Personalised Knowledge Graph Summarization [9, 22] were introduced. The PKG summarisation constructs personal summaries of users from a KG that contain the relevant facts of the users' interests. These summaries support personalised content queries and utility. These works are linked to KGs, such as the DBpedia and Yago, and provide the theoretical formulations for constructing PKG summaries from users' past queries. Moreover, similar work to the PKG summaries direction suggests graph-based approaches for users activity discovery from heterogeneous personal information collections, such as emails and files [23]. It proposes a method for unsupervised setting that perceives privacy. Influenced by these works, we intend to create PKGs containing information about users' and groups' queries and actions while paying attention to privacy.

3.2 Personalisation in e-learning

Personalisation in e-Learning is about creating an adaptive environment for the learner with adaptive content that derives mainly from the curriculum, educational resources, learning path, learning preferences, and cognitive state. Then the personalisation usually happens with a recommender system that suggests relevant content to the user, in terms of content type and characteristics as well as previous actions and performance of the user. The goal is to create user-centric applications that aim towards a positive learning experience. There are different methods suggested for achieving personalisation in e-learning; KGs is one of them [13]. The two main approaches are the symbolic, with the usage of ontologies, and a sub-symbolic approach, in which they handle recommender as a black box. Also, the combination of the symbolic and sub-symbolic exists [12]; however, due to space limitation we will not extend in this part.

The symbolic approaches are deployed with the usage of ontologies and semantic frameworks. WASPEC [3] utilises a semantic framework that models learner profiles and categorises personalisation parameters to learning preferences and accessibility. Our latest publication, EduCOR ontology [11], contributes in that line as it is analysed in Section 6. Another interesting approach is the formulation of the learning session graph to achieve personalisation in informal learning settings in information wikis [16]. Their framework achieves high scores in recommendation relevance. Encouraged by their design and findings, we target integrating session graphs in modelling user activity in the PKG.

Another approach is a "black-box" educational recommender system can benefit from open learner models providing a better sense of learning [1]. An open learner model assists students to understand their learning process and their peers. It can be classified as a cognitive tool approximating students' abilities with a score similar to learning analytics dashboards. Abdi et al. [1] showed that open learner models could increase the explainability and transparency of a recommendation; however, there were concerns regarding the fairness and feedback of the recommendation. Their study suggests that users' input and feedback into the recommendation process can improve satisfaction. Motivated by their findings, we aim to implement a transparent model in PKGs which justifies system actions and considers users' input and feedback in the recommendation process while offering additional personalisation via semantic enrichment.

3.3 Collaborative Search and Learning

Collaborative search can occur as a learning activity; Searching as Learning (SaL) explores this direction and links collaborative search with collaborative learning. SaL in a collaborative setting is formulated around project-based or team-based scenarios [24]. Research on understanding students online activity in SaL [17] suggests that the majority of students primarily choose content-based web pages, like Wikipedia, and that visualisations are essential, both for students and teachers who later can better monitor students' behaviour. This finding supports our idea of implementing PKGs and semantic technologies in the collaborative search and learning systems, and connecting the search with KGs and entities.

Furthermore, different systems have been introduced in the literature for collaborative search. One of the newest advancements is the QueryTogether [2], a system for entity-centric collaborative search in spontaneous settings that showed that entities could portray a vital role as interactive search objects. Also, their study confirmed that a common ground provided in collocated interaction supports
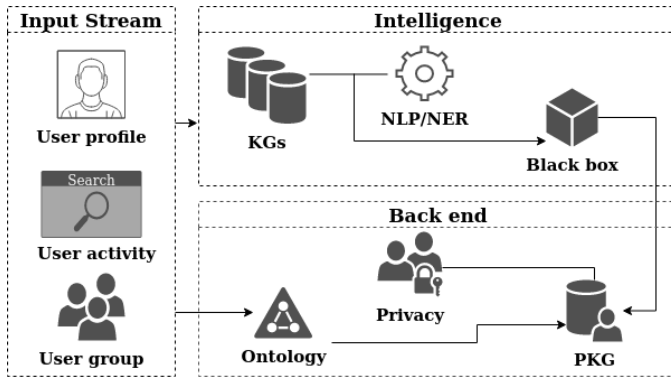
Figure 1: The overall architecture that includes the input stream, intelligence computation and back end creation of a personal knowledge graph (PKG).

group awareness, which has a positive impact on collaboration. Inspired by their work, we plan to emphasise the entities in the group search and implement features that aid group participation and awareness. Moreover, the Learnweb platform had implemented a semantic-based approach the LogCanvas [27], a visualisation tool linking to KGs for collaborative search. However, although the semantically enriched visual is a technically useful direction, the implementation lacks a collaborative and user-friendly interface and design. We will enhance this work by understanding the past pitfalls and designing collaborative semantic features linked to KGs.

4 PROPOSED APPROACH

We propose an architecture for the creation of a PKG for a user in the back end of an e-learning system in Figure 1 (RQ1). The input stream consists of the user's generated data, which is then passed to the intelligence part. In combination with an ontology, the input data are processed with the named entity recognition (NER) and natural language processing (NLP) software to identify the concepts and entities from the KGs. Because user's actions change through time, there is a weighted algorithm that recalculates the main points of interest for the user in each period of time. The black box calculates those weights and offers a filter in the recognised entities from the KGs, which will be passed to the PKG. Then the PKG is created with respect to user's and user group privacy. By implementing the PKGs as a structural element of a platform, we interconnect the activities that happen in the platform with Linked Data and KGs. The research work in the creation of a PKG intends to address the questions related to syntax and semantics as structural elements and regarding the represented data in each application context (RQ1).

4.1 Use case 1: Collaborative search

Applications using web search can get benefited by the usage of PKGs, which can offer semantically enhanced features and person-alisation capabilities. However, currently PKGs cannot play a role in the filter bubble of the search results provided to the users. We develop the collaborative semantically enhance search use case in the e-learning platform Learnweb [20]. We link the input stream

data with KGs to attain semantic relations between the data and identify the most important entities of the collaboration.

The proposed approach, in combination with a user-friendly interface and graphics, can help us reveal more personalised as well as group-project needs and develop collaborative features that are boosted with rich metadata and interconnections with the SW. Therefore, we would be able to examine how a collaborative e-learning platform can benefit from the addition of PKGs in their database to offer advanced SW features, personalisation (RQ3), and better collaboration in general (RQ4). The potential applications we can implement with the usage of PKGs include but are not limited to better users' credibility and understanding of the group's activity, learning analytics, recommendations, and feedback.

4.2 Use case 2: e-Learning

In this use case, we utilise the knowledge base of the eDoer platform [19], an open learning recommendation system prototype that connects the labour market skills with open educational resources (OERs). The eDoer is currently focused on Data Science related skills; however, the research targets into system's integration of a knowledge base in general domain OERs. The implementation of our approach could allow users to receive personalised recommendations based on their learning preferences and needs, accessibility needs and access, and semantic-based solutions, such as relevant to their topics learning content (RQ2).

4.3 Opportunities and Challenges

Our proposed approach subsist on some novel items and opportunities. So far, the current problem has not been solved with the usage of PKGs. This research might increase the interest in integrating PKGs in other domains, as well as the influence of the usage of knowledge bases in the e-learning community to utilise publicly available KGs and PKGs in their applications. An opportunity stands in the exploration of better collaborative features via the PKGs. Besides the plethora of collaborative search systems, no collaborative search interface has received major attention and has become long-established. This might happen due to the features offered in collaborative systems being adopted from single user purpose interfaces and are not directly developed for collaborative environments [10].

However, there are plenty of challenges this research needs to tackle. At first, we need to face the puzzle of knowledge acquisition, entity recognition and linking, storage and maintenance of the PKGs and connection to external PKGs. We select to start our alpha version by selecting DBpedia KG [21] and DBpedia Spotlight [6] to store our data locally in the Learnweb platform and maintain the data based on the processes provided by the Learnweb platform. In the future, we will focus on linking with more KGs and different NER software to improve our computation time and increase our coverage of search topics into more domains and languages. However, if the platform becomes widely used by thousands of users every day, we will be forced to expand our local implementation into a cloud service for computation and storage.

Also, by deploying PKGs, we need to raise privacy as a first-class citizen and ensure users' consent to select data from the personal

profile, store user actions, and share part of them with the collaborators to assist the collaboration and offer richer features. Also, the right to be forgotten based on General Data Protection Regulation [25] should be provided to the single user as well as the group results. We target to face this obstacle by keeping the users' data only for a limited period of time, nonetheless long enough to serve the team's objectives. Another vital factor to consider is that PKGs are highly time dependant [4]. Their information might change, particularly depending on the computation time, which affects the applications offered based on PKGs back end. Overall, we aim to address this part by setting time constrains on PKG computation time falls to ensure some level of confidence on the computed data. We discuss further challenges in future work.

5 METHODOLOGY

The methodology for approaching the research questions combines qualitative and quantitative analysis. In order to understand how the SW technologies are used in the e-learning domain, we plan to conduct an interdisciplinary examination of the state-of-the-art solutions relevant to our studied problems. This investigation includes the literature in the user profiling ontologies and schemes, literature in PKGs, and publications of solutions that apply personalisation in KGs and e-learning. Following, we aim to formalise our research questions and hypotheses and define potential solutions while identifying the innovative contributions. The formal characteristics of the proposed approach will be demonstrated, and the implemented solution will be made available to the community concerning users' privacy.

Mainly for evaluating our implementation, we will involve human participants in our research after receiving their consent based on national and international privacy laws. A similar work to ours [23], evaluated their method with requiting some participants (10 people). People with some technical background could perform experiments in the platform, which will allow us to collect and process their data and have their own experiences of the applications. We will develop documented in the form of interviews and questionnaires, and perform qualitative analysis; it will measure users' experience and satisfaction of the implemented features. Experts and professionals might also be included in our studies to receive precise feedback regarding the technical implementation, the learning aspects, and the user interface.

Nonetheless, we are facing a limitation because there are no goldstandards and baseline metrics for evaluating collaborative search and SaL. Among the different metrics that have been proposed are the effects on users' short and long term memory and gained knowledge, the time it takes to find information online and others. Several benchmarks have been suggested for a sub-domain of potential application tasks like the recommendations, such as the EdNet [7]; but, they do not fit the nature of SaL activities, project-based and team-based learning. However, we could compare our method with the state-of-the-art recommendation software by adopting them into a collaborative search learning environment.

6 RESULTS

The current line of research is still in an early stage. In our first steps in user modelling in e-learning platforms, we have published
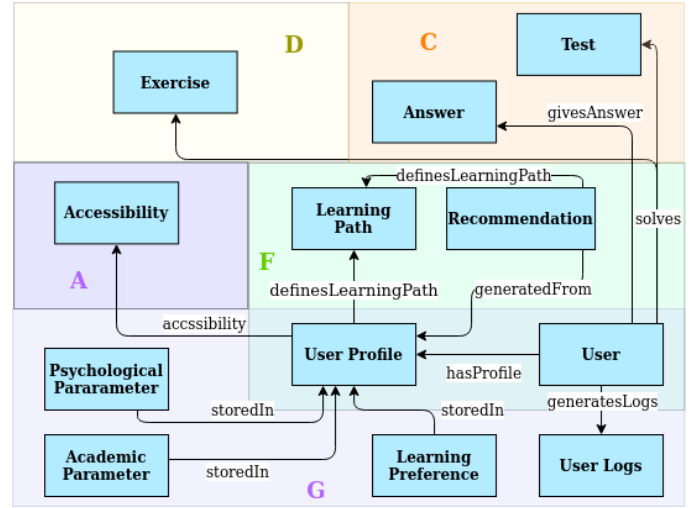

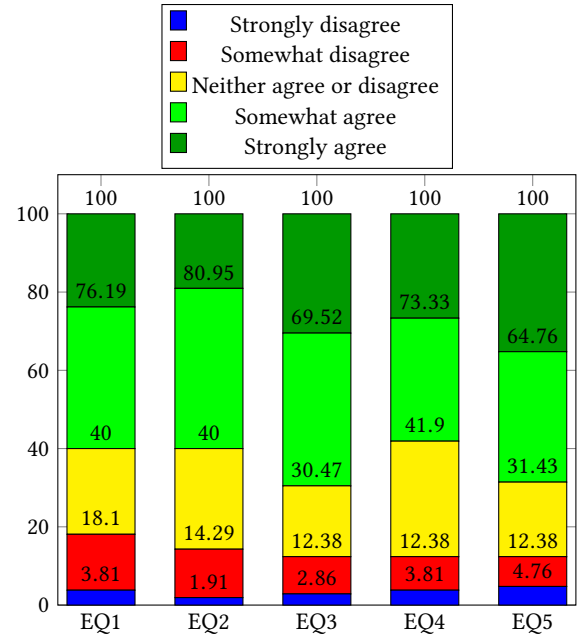Figure 2: User Profile pattern from EduCOR [11] ontology.

Figure 3: User general feedback from CollabGraph [15]. EQ: "I like the" 1: the group results visualized in a graph, 2: the summary of the team-members results, 3: the graph visualizations , 4: graph visualization next to the list view of the search results, 5: combination of the list and graph view.

EduCOR ontology [11] for personalised recommendations of educational resources. EduCOR consists of different parts and has a focus on user profiling, as can be seen in Figure 2. This part of the overall ontology has been extended in our latest submission to facilitate the necessary components for the creation of PKG ontology for web search [14] with respect to accessibility parameters, such as content access rights and privacy.

In our latest paper, which is currently under submission [15], we dive into the collaborative search and propose a collaborative search graph summary visualisation alongside the classical list-view of search results [15]. Our back end suggests the development of PKGs for each user which capture users' activity, connect to a KG in order to identify concepts and entities, and propagate the top entities to be visualised in the group summary graph. We evaluated our system on six different learning scenarios among 105 valid participants in a well-established user experience questionnaire and some evaluation questions we developed based on the parameters we wanted to rate. Our system shows high likability among users, as can be seen in Figure 3. However, it is still to be investigated how users perceive our system in an on-site experiment, and how the implementation of PKGs is superior to a simpler approach with the extensions we foresee for our system.

7 CONCLUSIONS AND FUTURE WORK

We presented the research plan on personal knowledge graphs in collaborative search environments and e-learning platforms. We outlined our design based on the related work and discussed the methodology and proposed implementation in the two different use cases in e-learning domain. We argued that the proposed approach could benefit collaborative learning search systems and e-learning platforms which are connected to knowledge bases by connecting them to semantic technologies, and we suggested a few applications to be deployed. However, there are is a broad domain of future applications such systems can develop.

We believe this work opens a new line of research in web search and PKGs. At first, this research can explore the knowledge acquisition processes as well as the maintaining, creation and update factors of PKGs.Further, privacy is a constant concern of personalised features. This issue could be addressed with a collaboration with legal researchers. Moreover, editor's and author's data could align with the data offered in the KG and provide the next generation KGs which offer advanced content credibility, access rights and privacy. Additionally, this work could be further developed to offer semantic personalised recommendations. These could be related to further web search items in collaborative search setting or suggested topics and educational resources in e-learning platforms.

From a general viewpoint, this research could be deployed towards the broader e-learning field and the human factors in computing systems or human-computer interaction. One case could be the additional annotations and features to support learning, such as direct feedback from the teacher in highlighted text and comments, and more user-centric visualisations. Another suggestion from the educational perspective could be the investigation of PKGs outcomes in knowledge building spaces.

ACKNOWLEDGMENTS

REFERENCES

[1] Solmaz Abdi, Hassan Khosravi, Shazia Sadiq, and Dragan Gasevic. 2020. Complementing educational recommender systems with open learner models. In *Proceedings of LAK'20*.

[2] Salvatore Andolina, Khalil Klouche, Tuukka Ruotsalo, Patrik Floréen, and Giulio Jacucci. 2018. Querytogether: Enabling entity-centric exploration in multi-device collaborative search. *Inf. Process. Manag.* (2018).

[3] Ufuoma Chima Apoki. 2021. The design of WASPEC: A fully personalised Moodle system using semantic web technologies. *Computers* (2021).

[4] Krisztian Balog and Tom Kenter. 2019. Personal Knowledge Graphs: A Research Agenda. In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval, ICTIR 2019, Santa Clara, CA, USA, October 2-5, 2019*.

[5] Senthilkumar N. C. and Ch. Pradeep Reddy. 2019. Collaborative Search Engine for Enhancing Personalized User Search Based on Domain Knowledge. *J. Medical Syst.* (2019).

[6] Mohamed Chabchoub, Michel Gagnon, and Amal Zouaq. 2018. FICLONE: Improving DBpedia Spotlight Using Named Entity Recognition and Collective Disambiguation. *Open J. Semantic Web* (2018).

[7] Youngduck Choi, Youngnam Lee, Dongmin Shin, Junghyun Cho, Seoyon Park, Seewoo Lee, Jineon Baek, Chan Bae, Byungsoo Kim, and Jaewe Heo. 2020. EdNet: A Large-Scale Hierarchical Dataset in Education. In *Artificial Intelligence in Education - 21st International Conference, AIED 2020, Ifrane, Morocco, July 6-10, 2020, Proceedings, Part II (Lecture Notes in Computer Science)*. Springer.

[8] H Davies, V Lehdonvirta, A Margaryan, J Albert, and LR Larke. 2020. Developing and matching skills in the online platform economy: Findings on new forms of digital work and learning from Cedefop's CrowdLearn study. (2020).

[9] Lukas Faber, Tara Safavi, Davide Mottin, Emmanuel Müller, and Danai Koutra. 2018. Adaptive Personalized Knowledge Graph Summarization. In *Proceedings of KDD Workshop on Mining and Learning with Graphs (MLG)*.

[10] Marti A. Hearst. 2014. What's Missing from Collaborative Search? *Computer* (2014).

[11] Eleni Ilkou, Hasan Abu-Rasheed, MohammadReza Tavakoli, Sherzod Hakimov, Gábor Kismihók, Sören Auer, and Wolfgang Nejdl. 2021. EduCOR: An Educational and Career-Oriented Recommendation Ontology. In *The Semantic Web - ISWC 2021 - 20th International Semantic Web Conference, ISWC 2021, Virtual Event, October 24-28, 2021, Proceedings (Lecture Notes in Computer Science)*. Springer.

[12] Eleni Ilkou and Maria Koutraki. 2020. Symbolic Vs Sub-symbolic AI Methods: Friends or Enemies?. In *CIKM (Workshops)*.

[13] Eleni Ilkou and Beat Signer. 2020. A Technology-enhanced Smart Learning Environment based on the Combination of Knowledge Graphs and Learning Paths.. In *CSEDU (2)*. 461–468.

[14] Eleni Ilkou, Davide Taibi, Marco Fisichella, and Tetiana Tolmachova. 2022. Personal Knowledge Graph Ontology for Web Search. (2022).

[15] Eleni Ilkou, Tetiana Tolmachova, Marco Fisichella, and Davide Taibi. 2022. CollabGraph: A graph-based collaborative search summary visualisation. (2022).

[16] Heba M Ismail, Boumediene Belkhouche, and Saad Harous. 2019. Framework for personalized content recommendations to support informal learning in massively diverse information Wikis. *IEEE Access* (2019).

[17] Roope Jaakonmäki, Jan vom Brocke, Stefan Dietze, Hendrik Drachsler, Albrecht Fortenbacher, René Helbig, Michael Kickmeier-Rust, Ivana Marenzi, Angel Suarez, and Haeseon Yun. 2020. Understanding Students' Online Behavior While They Search on the Internet: Searching as Learning. In *Learning Analytics Cookbook*.

[18] William Jones. 2007. Personal Information Management. *Annu. Rev. Inf. Sci. Technol.* (2007).

[19] TIB Labs. 2021. *eDoer Education Portal*. https://labs.tib.eu/edoer/

[20] Learnweb. 2021. *Learnweb - Learnweb*. https://learnweb.l3s.uni-hannover.de/

[21] Jens Lehmann, Robert Isele, Max Jakob, Anja Jentzsch, Dimitris Kontokostas, Pablo N. Mendes, Sebastian Hellmann, Mohamed Morsey, Patrick van Kleef, Sören Auer, and Christian Bizer. 2015. DBpedia - A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web* (2015).

[22] Tara Safavi, Caleb Belth, Lukas Faber, Davide Mottin, Emmanuel Müller, and Danai Koutra. 2019. Personalized Knowledge Graph Summarization: From the Cloud to Your Pocket. In *2019 IEEE International Conference on Data Mining, ICDM 2019, Beijing, China, November 8-11, 2019*.

[23] Tara Safavi, Adam Fourney, Robert Sim, Marcin Juraszek, Shane Williams, Ned Friend, Danai Koutra, and Paul N. Bennett. 2020. Toward Activity Discovery in the Personal Web. In *WSDM '20: The Thirteenth ACM International Conference on Web Search and Data Mining, Houston, TX, USA, February 3-7, 2020*. ACM.

[24] Tetiana Tolmachova, Eleni Ilkou, and Luyan Xu. 2020. Working Towards the Ideal Search History Interface. In *Proceedings of the CIKM 2020 Workshops co-located with 29th ACM International Conference on Information and Knowledge Management (CIKM 2020), Galway, Ireland, October 19-23, 2020*. CEUR-WS.org.

[25] Paul Voigt and Axel Von dem Bussche. 2017. The EU general data protection regulation (GDPR). *A Practical Guide, 1st Ed., Cham: Springer International Publishing* (2017).

[26] Huaqiong Wang, Xiaoyu Miao, and Pan Yang. 2018. Design and implementation of personal health record systems based on knowledge graph. In *2018 9th International Conference (ITME)*.

[27] Luyan Xu, Zeon Trevor Fernando, Xuan Zhou, and Wolfgang Nejdl. 2018. LogCanvas: Visualizing Search History Using Knowledge Graphs. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, SIGIR 2018, Ann Arbor, MI, USA, July 08-12, 2018*. ACM.