Enhancing Wikidata Exploration: Usability Insights and Design Reflections from WikiframeVG Users

Darnelle Melvin Cory Lampert
University of Nevada, Las Vegas

Abstract

In this research, we propose a mixed-methods usability study to identify WikiframeVG issues, enhance Wikidata community satisfaction, and inform future designs. Targeting Wikidata editors and users, our aim is to improve knowledge graph-based visualization tools' efficiency, effectiveness, learnability, and Wikidata community satisfaction.

Introduction

Within the broader semantic web and galleries, libraries, archives, museums, and special collection communities (GLAMS), there still lie barriers for mass adoption and consumption of linked data technologies [1-4]. These barriers impeding general use include:

- Lack of expertise in rdf modeling and writing SPARQL
- Lack of deep knowledge of the semantic web stack
- Lack of domain knowledge
- Lack of tools for linked data discovery and visualization.

To lower the barrier for adoption and consumption and based on community feedback and needs, Wikiframe Visual Graph (WikiframeVG) [5] was developed and released for community use. WikiframeVG is an open-source search and discovery tool aimed at helping Wikidata editors and general users visualize semantic relationships between Wikidata entities of interest using a no-coding, template-based approach. The initial prototype focuses on cultural heritage organizations; but other research domains can use it with modification to the mapping template. For a demo of WikiframeVG see presentation from the 2023 LD4 Conference on Linked Data [6]. In this project we propose a usability study to identify WikiframeVG usability issues, improve community satisfaction and retention, optimize conversion and task completion, and inform future designs.

Enhancing WikiframeVG's usability is vital for Wikidata projects and the wider community. As a crucial tool for Wikidata editors and users, improving its usability directly boosts the efficiency of semantic visualization tools. This aligns with the Wikimedia Foundation's commitment to democratizing knowledge, making linked data technologies more accessible. The research benefits Wikidata editors by streamlining workflows, fostering increased satisfaction, and retention.

In our usability study we intend to focus our research along six themes, including effectiveness, efficiency, learnability, satisfaction, barriers to adoption, and use in different domains. Some research questions worth investigation include:

1. How effectively does WikiframeVG enable users to visualize semantic relationships between Wikidata entities of interest?

2. How efficiently can users complete common tasks (e.g., creating visualizations, exploring relationships, view maps) using WikiframeVG?
3. How quickly can users, especially those with varying levels of expertise, learn to use WikiframeVG to create and interpret visualizations?
4. To what extent are users satisfied with the overall usability and user interface design of WikiframeVG?
5. What are the primary barriers preventing users from adopting WikiframeVG, and how can these barriers be mitigated?
6. How adaptable is WikiframeVG for users from different research domains, particularly beyond cultural heritage organizations?

Date: June 1, 2024 thru June 30, 2025.

Related work

In related work, Grether and Witschel [7] conducted two qualitative user studies focusing on interaction patterns with a graph-based discovery tool on a medical-themed knowledge graph. Their research investigates the impact of guidance mechanisms on improving exploration and discovery of unfamiliar information. Kuric, Fernandez, and Drozd [8] proposed a mixed methods approach to explore the overall user experience in navigating a knowledge graph with a query builder. Their study utilized a think-aloud approach and a subsequent usability evaluation, with participants interacting with the query builder tool while verbalizing thoughts, providing insights into usability and user experience aspects of navigation. Additionally, García et al. [9] presented a benchmark for graphical user interfaces in semantic data, offering a set of user tasks and metrics to assess interface performance, contributing to the evaluation landscape in the field.

Methods

Our research employs a mixed-methods approach, combining quantitative and qualitative data sources. Surveys, designed to gauge WikiframeVG usability, will be administered to Wikidata editors and users selected through purposive sampling. Recruitment will be via LD4 and Wikidata community channels. Concurrently, a usability study will be conducted, involving participants interacting with WikiframeVG while providing insights. Quantitative data will undergo descriptive statistical analysis, while qualitative data, including usability study outcomes, will be subjected to thematic analysis. Synthesizing these findings will allow us to derive comprehensive insights into WikiframeVG's usability, addressing research questions and contributing to future design improvements.

Expected output

- ➔ Insights for decision-making and future WikiframeVG designs
- ➔ Publication Venues:
  - ◆ Information Technology and Libraries
  - ◆ Journal of Web Semantics
  - ◆ Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering, and Knowledge Management
  - ◆ Proceedings of the Wikidata Workshop co-located with the International Semantic Web Conference
- ➔ Conference Participation:
  - ◆ American Library Association Annual Conference: Linked Data Interest Group

- Digital Library Federation Forum
- LD4 Conference on Linked Data
- Wikidata Workshop co-located with the International Semantic Web Conference
- Wikimania

In summary, Wikidata editors, general users, cultural heritage organizations, and knowledge graph researchers from diverse domains are the key target audiences for both the publication venue and conference participants.

Risks

This research project involves minimal risk to participants. There will be an exempt human subjects research proposal sent to UNLV's Institutional Research Board for review.

Other risks include holding community feedback events and teaching/partnering with other organizations. Both PI and co-PI have experience in conducting research projects (including usability studies), leading virtual and in-person events and teaching (including formal graduate program courses).

Community impact plan

- Plan to work with LD4 Wikidata Community for ongoing feedback via planned meetings
- Plan to work with community volunteers who participate as usability tests (5-7 volunteers) who have worked on a Wikidata project
- Plan to work with (1-3) Wikidata project managers to discuss adoption of Wikiframe

This work in the community will be tracked on the UNLV Wikidata Project page and shared openly.

Evaluation

Measures of success for the research include:

- Completion of usability study (participant data and formal report)
- Present results via conference or publication
- List of development priorities (publicly shared) for Wikidata gained from project activities to pass on to developer

Budget

A preliminary budget can be found on the budget draft [21].

Prior contributions

Prior contributions worth noting include: open-source software [20], instructional wikis for Wikimedia community reuse [16-19], monographs [10], conference proceedings [5], chapters [1], articles [2], presentations [6,11-13], and invited talks [14-15].

References

1. Melvin, D., & Lampert, C. (2023). Ethical explorations using Wikidata and Wikidata tools to expose underrepresented special collection materials. In A. Provo, K. Burlingame, & B.M. Watson (Eds.), Ethics in linked data (pp. 411-435). Library Juice Press. https://osf.io/eyjdr
2. OCLC Research Archives and Special Collections Linked Data Review Group. (2020). Archives and Special Collections Linked Data: Navigating between Notes and Nodes. Dublin, OH: OCLC Research. https://doi.org/10.25333/4gtz-zd88
3. Smith-Yoshimura, K. (2018) Analysis of 2018 International Linked Data Survey for Implementers. Code4lib Journal, 42. https://journal.code4lib.org/articles/13867

4. Smith-Yoshimura, K. (2016) Analysis of International Linked Data Survey for Implementers. D-Lib Magazine, 22(7/8). https://www.doi.org/10.1045/july2016-smith-yoshimura
5. Melvin, D. Hulet, A., & Lampert, C. K. (2023). WikiframeVG: A SPARQL Template-based Framework for Wikidata Graph Exploration and Visualization. In L-A Kaffee, S. Razniewski, G. Amaral, & K. S. Alghamdi (Eds.), *Proceedings of the 4th Wikidata Workshop 2023 co-located with the 22nd International Semantic Web Conference (ISWC2023)*. CEUR Workshop Proceedings. https://osf.io/z2k45
6. Melvin, D., & Hulet, A. (2023, July 12). WikiframeVG: A SPARQL template-based wikiframe for Wikidata graph exploration. [Conference presentation]. LD4 Conference on Linked Data. https://www.youtube.com/live/VYoD6-aw80M?feature=share&t=215
7. Kuric, E., Fernández, J.D., Drozd, O. (2019). Knowledge Graph Exploration: A Usability Evaluation of Query Builders for Laypeople. In: M. Acosta, P. Cudré-Mauroux, M. Maleshkova, T. Pellegrini, H. Sack, & Y. Sure-Vetter (Eds.), Semantic Systems. The Power of AI and Knowledge Graphs. SEMANTiCS 2019. Lecture Notes in Computer Science, vol 11702. Springer, Cham. https://doi.org/10.1007/978-3-030-33220-4_24
8. Grether, L., & Witschel, H. F. (2022). Studying interaction patterns for knowledge graph exploration. Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K2022), Valletta, Malta, 1, 257-264. https://doi.org/10.5220/0011548600003335
9. García, R., Gil, R., Bakke, E., & Karger, D. R. (2020). A benchmark for end-user structured data exploration and search user interfaces. *Journal of Web Semantics, 65*, 100610. https://doi.org/10.1016/j.websem.2020.100610
10. Carlson, S., Lampert, C., Melvin, D., & Washington, A. (2020). Linked data for the perplexed librarian. Chicago: ALA Editions. https://id.loc.gov/resources/works/21317899.html
11. Melvin, D. (2023, July 11). Proactive strategies to improve underrepresentation in public knowledge graphs: A Wikidata sprint in UNLV Special Collections highlighting LGBTQ+ Las Vegas. LD4 Conference on Linked Data. LD4 Community. Retrieved from https://doi.org/10.17605/OSF.IO/2GE9H
12. Lampert, C., & Ott, K. (2023, November 6). Pride and Joy: Using Wikidata to Showcase LGBTQ+ Archives at UNLV. Nevada Library Association Conference.
13. Melvin, D. O. (2021, August). What came first the SPARQL or the egg?: Managing SPARQL informed Wikidata productions for special collections and archives. LD4 Conference on Linked Data. LD4 Community. Retrieved from: https://doi.org/10.17605/OSF.IO/ZQ25V
14. Melvin, D., Lampert, C. K., & Hulet, A. (2023, October). WikiframeVG: A search and discovery application for Wikidata projects. LD4 - Wikidata Affinity Group.
15. Melvin, D., Hulet, A. (2023, September). WikiframeVG with Darnelle Melvin and Andre Hulet. LD4 - Rare Materials Affinity Group.

16. PCC Wikidata Pilot/UNLV: https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/UNLV
17. QuickStatements: https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/UNLV/QuickStatements
18. Introduction to Cradle: https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/UNLV/Introduction_to_Cradle
19. SPARQL Examples and Concepts: https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/UNLV/SPARQL_Examples_and_Concepts
20. Wikidata-discovery project: https://github.com/UNLV-Libraries/wikidata-discovery-project
21. Budget Draft: https://docs.google.com/document/d/1yM_WqMAWqcM-BJzRTa-VV7JOi0-uV3MvZ-2K6F2lCSM/edit?usp=sharing