A Multi-layer Event Visualization for Exploring User Search Patterns in Literature Discovery with PUREsuggest

Solveig Rabsahl
Fabian Beck
s_rabsahl@stud.uni-bamberg.de
fabian.beck@uni-bamberg.de
University of Bamberg
Bamberg, Germany

ABSTRACT

Understanding user behavior is at the heart of user interface design, but can only be quantified to some extent. Qualitatively analyzing individual usage sessions is especially important in open-ended tasks like literature search. In this paper, we present a visual representation of logging data that provides the basis for an in-depth analysis and annotation of search and exploration sessions. The visualization was developed to evaluate the citation-based literature discovery tool PUREsuggest and is thus aimed at visualizing logging data of a literature search system. Events are represented on a timeline in different layers as bars and icon-based glyphs, and contextualized by the additional visualization of item states and active user-set search modifiers such as filters or keywords throughout the session. We demonstrate the applicability of the visualization by evaluating excerpts of two user sessions as an example.

CCS CONCEPTS

• **Human-centered computing** → **Visualization design and evaluation methods**; **Visualization systems and tools**; **Information visualization**; • **Information systems** → **Recommender systems**.

KEYWORDS

User Behavior, Event Visualization, Literature Search

ACM Reference Format:
Solveig Rabsahl and Fabian Beck. 2024. A Multi-layer Event Visualization for Exploring User Search Patterns in Literature Discovery with PUREsuggest. In *Proceedings of Mensch und Computer 2024 (MuC '24), September 01–04, 2024, Karlsruhe, Germany.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3670653.3677502

1 INTRODUCTION

The increasing availability of citation metadata, as offered by open APIs like *Crossref* [7] and *OpenCitations* [17], has led to the emergence of a new class of tools supporting literature search. For recommending new publications, citation-based literature discovery tools leverage citation links of seed publications provided by the user and allow incremental buildup of literature collections. Examples include, but are not limited to *CitationGecko* [19], *Litmaps* [14], *Connected Papers* [9], *ResearchRabbit* [18], *Inciteful* [20], and *PUREsuggest* [3, 4]. They account for both the exploratory nature of literature search and the importance of citation relationships, which help discover related publications.

To develop tools that align with users' workflows and research strategies, it is important to understand the users' search processes. Visual literature exploration, however, is inherently open-ended and aims at a broad understanding of an academic topic. Traditional quantitative measures like answer accuracy and time can only partly capture user performance of interactive visualization systems [16]. Qualitative methods that analyze individual sessions, in contrast, promise to contribute a richer picture of the search strategies applied by users. However, as recordings of interaction data can easily grow too large to be manually analyzed in their raw format, a more effective data representation is needed through which the sessions can be systematically analyzed and annotated.

To support researchers in a qualitative analysis of users' search strategies through a visual representation of the interaction logging data from a literature search process, we developed the visualization approach presented in this paper. The visualization was created for evaluation of *PUREsuggest* and is tailored to the visualization of logging data produced by *PUREsuggest*. As shown in Figure 2, the visualization is organized into layers, summarizing the interaction events at different levels of granularity. Identified usage patterns can be interactively annotated in the visualization to incrementally analyze the events of the whole session. The visualization has been developed in the context of a user evaluation of *PUREsuggest* [3], but has not been described as part of previous publications.

2 RELATED WORK

Visualizing action sequences on timelines is a common technique for analyzing behavioral data—*timeline-based visualizations* are described as a standard method in literature surveys on event visualization [12, 21]. As our approach shows the detailed behavior of individual usage sessions (data scale: *sequence* [12]), we are focusing the discussion of approaches that use the timeline to draw a rich picture of the events of a session, but omit approaches rather aggregating the data or mining patterns from it. We observe similarities to our approach, for instance, regarding the use of layered timelines to show different levels of granularity [2], icon images to represent events [13], dedicated timeline rows for different data items [2, 15], and annotation features [8].

MuC '24, September 01–04, 2024, Karlsruhe, Germany

ACM ISBN 979-8-4007-0998-2/24/09
https://doi.org/10.1145/3670653.3677502

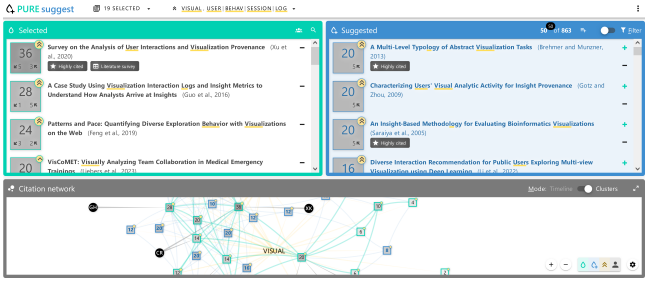

Figure 1: The interface of *PUREsuggest* comprises the list of selected publications (left), the list of suggested publications (right), and a citation network view (bottom); keywords can be set to steer the search (header).

While we have not found any generic frameworks for visualizing sessions details through interaction log data, there are various application-specific approaches. Similar to our solution, some approaches focus on representing analysis sessions where people work with visual analysis systems [10], for instance, trying to understand reasoning processes [8], connecting derived insights with interaction logs [11], or integrating eye tracking and interaction data [6]. But also other types of user sessions can be visualized similarly. For instance, Agarwal et al. [1] describe a design space of solutions to visualize mixed reality user sessions, also providing examples with timeline-based encodings. Minelli et al. [15] visualize software development sessions in a multi-window IDE system. Also related are timeline-based approaches that visualize AI agent behavior in games [2] or human teams in professional trainings [13]. Like us, these application-specific approaches all leverage specialized event encodings on timelines. However, we are not aware of solutions for literature search or visual recommendation systems.

3 BACKGROUND AND DATA COLLECTION

The interface of *PUREsuggest*, shown in Figure 1, consists of panels that depict selected (green) and suggested (blue) publications, as well as the citation network (gray) spanned by currently selected and suggested publications. A literature search starts by adding a few seed publications (either known in advance or found through a keyword-based search dialog). The system then makes recommendations based on the number of references from or to the selected publications. Users can enter keywords that highlight and boost the ranking of publications with matching titles (yellow). Ranking scores are visually explained in glyphs preceding the publication entry. Upon user selection, publications will be queued and only added to the selection upon an update request by the user. This way, the users can incrementally build a literature collection. They may use the citation network visualization, which is available in a timeline and cluster mode, for analyzing the collected data from different perspectives, and to link publications with the main authors and keywords, which are also represented as nodes in the network.

We instrumented the publicly available code of *PUREsuggest*[1] to record user actions concerning publication interactions and adjustments to the interface (Table 1), alongside status updates for all publications (such as that a publication was suggested). Log messages hold a timestamp, the respective action and additional information such as the component where an action was performed.

[1]https://github.com/fabian-beck/pure-suggest

Table 1: Selected examples of logged events and their corresponding icons used in the visualization.

Type	Event
Publication interactions	Publication activated
	Publication queued for selection
	Publication excluded
	Abstract opened
	DOI clicked
Interface actions	*Recommendation steering*
	Keywords updated
	Filter applied
	Interface adjustments
	Network mode changed
	Network expanded
	Author page opened

4 VISUALIZATION OF EVENT LOGS

For the qualitative analysis of interaction logs of literature search sessions, we propose a tailored event visualization and annotation approach, particularly supporting the identification of related actions and employed search strategies. The visualization consists of a multi-layer timeline view (temporal axis from left to right), as illustrated in Figure 2. It mainly shows user actions (Table 1) and contextual information about the status of the interface (e.g., active filters) and session (i.e., number of selected works). While user actions are the core of the analysis, contextual information supports the correct interpretation of observed actions. User actions are visualized as circular glyphs on the horizontal timelines, reflecting their inherently atomic nature. Status information is abstracted as numerical variables (such as number of active filters) and visualized using line, area, and bar charts. Since the approach is tailored to evaluate *PUREsuggest* sessions, colors correspond to the interface colors shown in Figure 1, as well as event icons match—where possible—the respectively used icons from the interface. Although the static image of the timeline already provides a rich summary of the session, interactions further deepen the exploration options and correspondences between events in separate layers can be traced. For instance, hovering over publication-related events in Layer B highlights the respective publication timeline and all related events in Layer B. Further details about single events are accessible on hover. We implemented the approach as a web application using *d3.js*; icons are based on *Google Material Design*.

Status chart. At the top of the timeline, Layer A provides an overview of the number of currently selected publications as a stacked area chart, discerning selected publications (green) and those queued to be selected (light blue). Hence, this chart allows tracking additions and removals of selected publications during the session. Additional line charts display the number of currently

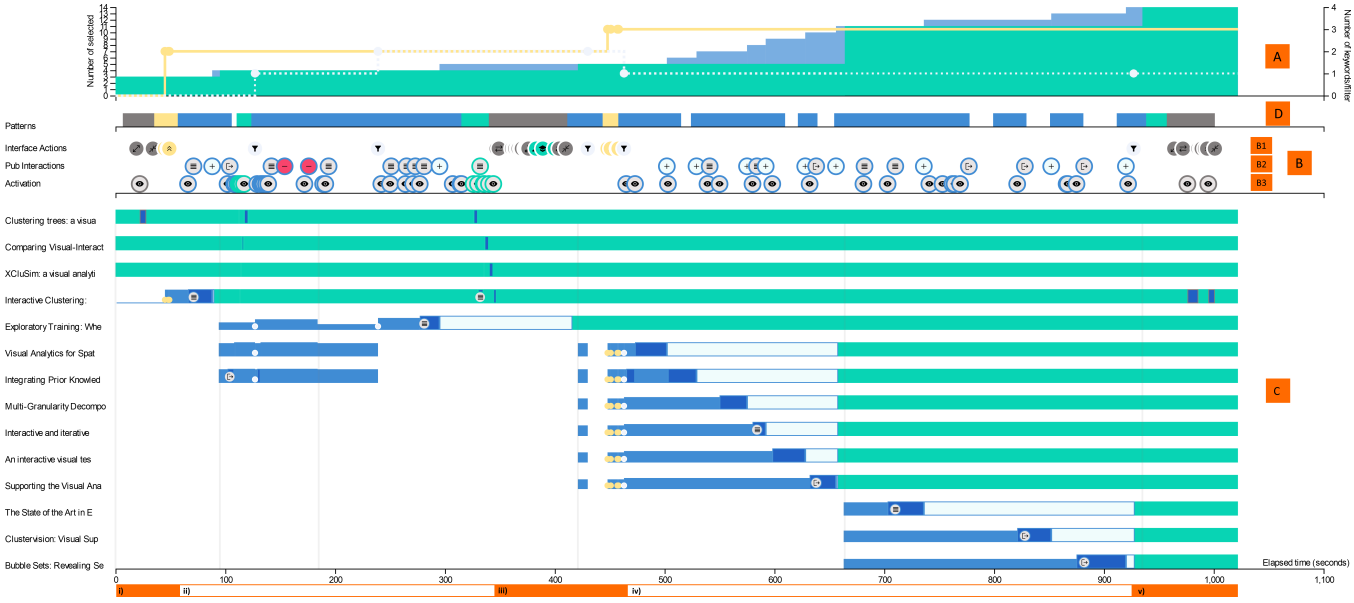

Figure 2: Timeline-based visualization of a literature search session organized into layers: (A) summarized numbers of selected publications as well as filters and keywords; (B) icons representing interaction events categorized in three rows by their type of interaction; (C) publication timelines providing information about the ranking and selection of individual publications; (D) annotated interaction patterns as coded by the analyst.

active filters (gray) and keywords (yellow), and reflect user adjustments to steer the ranking according to their individual search scope. Although area and line charts concern different analysis targets, they are superimposed onto a single diagram to save space.

Interaction timelines. Layer B, below, summarizes user interactions represented as simple icon-based circular glyphs. Circle colors indicate affected components or interface functionality. Stroke colors symbolize the component of *PUREsuggest* within which the action was performed. While this is a simple design, the icons make it easy to incorporate various actions into the visualization in an easy-to-interpret way. We group interactions into three categories: Line B1 shows *interface actions*, which affect the suggestion ranking (e.g., filtering, keyword updates) or the current view (e.g., network interactions). We also discern between general *publication interactions* (B2), which are substantial actions tied to specific publications, and *activation interactions* (B3), which refer to the inspection of publication details (e.g., clicking on the publication entry). Separating interactions into distinct timelines helps to get a notion of user goals at a glance. For instance, activity in line B3 without corresponding activity in line B2 suggests a focus on exploration, whereas concurrent activity in both lines hints at a focus on expanding the collection.

Publication timelines. As the most detailed part consuming most screen space, Layer C shows the states of the selected publications over the duration of the session. Publications are arranged vertically according to the time of first selection (or when they were first queued to be selected, respectively). States are visualized as horizontal bars with lengths scaled according to the duration of the state. Typically, a publication progresses through the states *suggested* (medium blue), *activated* (dark blue), *queued for selection* (light blue) and *selected* (green), as can well be seen in line 4 of the publication timelines. For suggested publications, bar heights indicate the position of the publication in the suggestion ranking, with taller bars corresponding to higher suggestion scores. Additionally, filter and keyword matches are displayed as gray and yellow dots on the timelines of corresponding publications. Icons representing publication-specific user actions, as seen in Layer B, are repeated on the corresponding publication timelines.

Annotation timeline. Finally, Layer D provides a timeline for interactively annotated usage patterns. It is placed at the top, between Layer A and Layer B, as it rather abstracts and summarizes the user behavior. Users annotate events and assign a code by brushing over the relevant time frame on the interaction timelines. Using a simple dropdown menu and text fields (not pictured), analysts can set predefined or custom codes. Custom codes bear a custom name, default color, and a short description, if desired. Once a code is selected, the code will appear as a colored bar segment of fixed height on the annotation timeline. In the current prototype of the tool, codes are not intended to overlap.

5 APPLICATION EXAMPLE

To demonstrate the applicability of the approach, we analyze two user sessions, which were drawn from the user study of *PUREsuggest* [3]. We selected them as interesting representatives of the 27 recorded sessions and as they complement each other regarding the temporal granularity of interesting patterns, analyzing the sessions in greater depth than in the previous publication [3]. In the first session (Figure 2), we perform an overall analysis on two scales of

temporal granularity. In the second session, we focus on an interesting session segment of a different user (Figure 3) to demonstrate how actions displayed in Layer B can be contextualized using Layer A and Layer C. Both users—a graduate student and a PhD student—were asked to build/extent a literature collection for a given topic based on a small set of seed papers.

In a first broad analysis of the first session, we categorize different search phases, focusing either on deliberate examination of publications to expand the selection or on exploration of the literature. To identify these phases, we use the ratio of activity between the different categories of the interaction timeline (B), namely, contrasting activations (B3) with interface actions (B1) and publication interactions (B2). During three time intervals at the beginning, mid, and end of the session (i, iii, v) the majority of activity relates to interface actions (B1) with almost no publication interactions (B2), indicating that the user explored the current selection and network representation of the literature. By contrast, during the two longer phases in between (ii, iv), the user focused on expanding the collection. The balanced ratio of activity regarding activation (B3) and publication interaction (B2) shows that most publication activations were accompanied by further interaction, implying that the user systematically examined individual publications. These contrasting search phases are also evident in the status chart (A), especially with the number of selected publications steadily increasing during the second expansion phase (v).

To get a deeper understanding of the identified exploration and expansion phases, using the visual interface, we manually assigned codes to user activities of finer granularity. The codes we employed are (*c1*) *suggestion review*, (*c2*) *selection exploration*, (*c3*) *keyword revision*, and (*c4*) *citation network analysis*. As they correspond to specific components in the *PUREsuggest* interface, we choose the colors for annotation codes accordingly (*c1*: blue, *c2*: green, *c3*: yellow, *c4*: gray). The final coding is depicted in the annotation timeline (D) of Figure 2. We observe that the user began by expanding and collapsing the network (B1), activating a publication in between (B3). We consider this as a brief exploration of the network (*c4*). Next, the user updated the keywords (B1,*c3*). They then began a longer period of reviewing suggestions (*c1*), which was interrupted by a departure from the tool via DOI-click (which opens the DOI-referenced publication page in a new tab) around the 120-second mark and two brief periods of exploring the current selection (*c2*; around second 130 and second 320). They then spent some time exploring the network functionality (B1, *c4*), before continuing to review suggestions (*c1*; seconds 410–940). During this extended suggestion review, the user updated the keywords (B1, *c3*) and occasionally opened publications via DOI-click. Around the 950-second mark, they briefly scrolled through the selected literature (*c2*; not visible in the visualization but derived from screen recording of the session). Finally, the user explored the keywords and literature in the network view (*c4*; the visually dense section in the interface timeline corresponds to multiple keyword drags in the network).

Analyzing an excerpt of the second session on an even finer temporal granularity, we now focus on individual actions to examine how users employ the tool's functionalities to steer the ranking according to their search scope. The cutout view in Figure 3 illustrates how the combination of filtering and keyword boosting affects the publication ranking. The status chart (A) shows that, initially, the

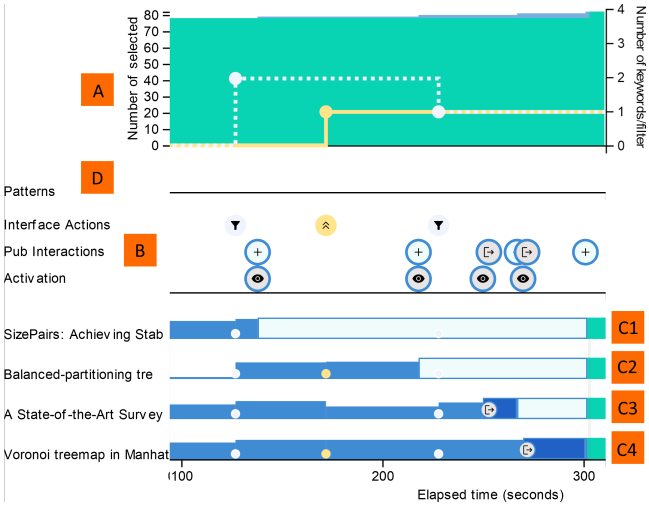

Figure 3: Detail of a second user session. The effects of changes in the number of active filters and keywords (A) resulting from user actions (B) are evident in the bar height of suggested publications (C1–C4). Publications not relevant to the example are omitted.

user set two filters, affecting the publication in line C2 in particular. The increasing height of the bar representing this publication indicates a leap from a low to a high recommendation score (ca. second 125). The user then selected a publication and added a keyword (A, B), which improved the ranking of the publications in lines C2 and C4. Note how the application of keywords or filters can also lower the ranking of mismatching publications, as in line C3. After adding the publication in line C2 to the selection, the number of active filters was reduced to one (A), which improved the ranking of the publication in line C3. The user selected both filtered publications. The observation that the publications that were affected by the steering actions are immediately added to the selection implies that the usage of the filter- and keyword-option was successful.

6 DISCUSSION AND CONCLUSION

In this paper, we presented a visualization approach supporting the analysis of logging data gathered during literature search sessions using the citation-based literature search tool *PUREsuggest*. We described design considerations and demonstrated in an application example that the visualization supports data interpretation at different temporal resolutions and the systematic annotation of user behavior. These examples already illustrate how an analyst can efficiently analyze user sessions as part of a qualitative study design. The annotation timeline (D) captures the results of the analysis and can be cut out and used to compare different sessions. Aside to assessing the suggested visualization itself in a user study with several analysts, the following limitations and extensions can be addressed as part of future work.

The approach relies on the scope and granularity of the logging data. Despite already providing rich information, it is not possible to infer from the data, for instance, whether periods of inactivity are spent idle, outside the tool or with activities that were not logged,

such as scrolling through publication lists or looking at the network. While these issues could be mitigated by expanding the logging scope to also include continuous data such as scroll and hover data, or even gaze tracking, the additional data might be difficult to interpret and visualize. Alternatively, additional screencasts can be recorded and used as reference (as done in the application example). In the future, the recorded videos can be integrated with the timeline representation, similar as described in related approaches [2, 13]. Further relevant extensions could include the support for comparing and abstracting sessions by multiple users, as partly explored in other event visualizations already [5, 13].

The presented visualization is specific to *PUREsuggest*, hence, does not directly generalize to other applications. However, when studying similar exploration and search processes, the proposed concepts can be reused, for instance, the different layers of the timeline. At the core, different items that are the target of the search can also be presented as rows, with an adequate encoding of their first time of appearance and following user interactions that relate to them. Still many changes would be necessary, for instance, states of discovered items might be different and would require a revised encoding, as well as event icons need to be adapted to different types of events. Generally, we advocate the approach of visually analyzing logging data as a complement to more traditional quantitative study designs and an enhancement of qualitative methods, which otherwise often rely on manual video analysis only.

ACKNOWLEDGMENTS

The work has been partially supported by University of Bamberg through internal project funding.

REFERENCES

[1] Shivam Agarwal, Jonas Auda, Stefan Schneegaß, and Fabian Beck. 2020. A Design and Application Space for Visualizing User Sessions of Virtual and Mixed Reality Environments. https://doi.org/10.2312/vmv.20201194

[2] Shivam Agarwal, Günter Wallner, and Fabian Beck. 2020. Bombalytics: Visualization of Competition and Collaboration Strategies of Players in a Bomb Laying Game. *Computer Graphics Forum* 39, 3 (2020), 89–100. https://doi.org/10.1111/cgf.13965

[3] Fabian Beck. 2024. PUREsuggest: Citation-based Literature Search and Visual Exploration with Keyword-controlled Rankings. to appear. *IEEE Transactions on Visualization and Computer Graphics* (2024).

[4] Fabian Beck and Cedric Krause. 2022. Visually Explaining Publication Ranks in Citation-based Literature Search with PURE Suggest. In *EuroVis 2022 – Posters*. The Eurographics Association, 3 pages. https://doi.org/10.2312/evp.20221110

[5] Tanja Blascheck, Fabian Beck, Sebastian Baltes, Thomas Ertl, and Daniel Weiskopf. 2016. Visual analysis and coding of data-rich user behavior. In *2016 IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE, 141–150. https://doi.org/10.1109/vast.2016.7883520

[6] Tanja Blascheck, Markus John, Kuno Kurzhals, Steffen Koch, and Thomas Ertl. 2016. VA2: a Visual Analytics Approach for Evaluating Visual Analytics Applications. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 61–70. https://doi.org/10.1109/tvcg.2015.2467871

[7] Crossref. (accessed: 2024/07). Crossref. https://www.crossref.org/.

[8] Wenwen Dou, Dong Hyun Jeong, Felesia Stukes, William Ribarsky, Heather Richter Lipford, and Remco Chang. 2009. Recovering Reasoning Processes from User Interactions. *IEEE Computer Graphics and Applications* 29, 3 (2009), 52–61. https://doi.org/10.1109/mcg.2009.49

[9] Alex Tarnavsky Eitan, Eddie Smolyansky, Itay Knaan Harpaz, and Sahar Perets. (accessed: 2024/07). Connected Papers. https://www.connectedpapers.com/.

[10] Omar Eltayeby and Wenwen Dou. 2016. A Survey on Interaction Log Analysis for Evaluating Exploratory Visualizations. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*. Association for Computing Machinery, 62–69. https://doi.org/10.1145/2993901.2993912

[11] Hua Guo, Steven R. Gomez, Caroline Ziemkiewicz, and David H. Laidlaw. 2016. A Case Study Using Visualization Interaction Logs and Insight Metrics to Understand How Analysts Arrive at Insights. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 51–60. https://doi.org/10.1109/tvcg.2015.2467613

[12] Yi Guo, Shunan Guo, Zhuochen Jin, Smiti Kaul, David Gotz, and Nan Cao. 2022. Survey on Visual Analysis of Event Sequence Data. *IEEE Transactions on Visualization and Computer Graphics* 28, 12 (2022), 5091–5112. https://doi.org/10.1109/tvcg.2021.3100413

[13] Carina Liebers, Shivam Agarwal, Maximilian Krug, Karola Pitsch, and Fabian Beck. 2023. VisCoMET: Visually Analyzing Team Collaboration in Medical Emergency Trainings. *Computer Graphics Forum* 42, 3 (2023), 149–160. https://doi.org/10.1111/cgf.14819

[14] Litmap Ltd. (accessed: 2024/07). Litmaps. https://www.litmaps.com.

[15] Roberto Minelli, Andrea Mocci, Michele Lanza, and Lorenzo Baracchi. 2014. Visualizing Developer Interactions. In *2014 Second IEEE Working Conference on Software Visualization*. IEEE, 147–156. https://doi.org/10.1109/vissoft.2014.31

[16] Carolina Nobre, Dylan Wootton, Zach Cutler, Lane Harrison, Hanspeter Pfister, and Alexander Lex. 2021. reVISit: Looking Under the Hood of Interactive Visualization Studies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, 13 pages. https://doi.org/10.1145/3411764.3445382

[17] OpenCitations. (accessed: 2024/07). OpenCitations. https://opencitations.net/.

[18] Research Rabbit. (accessed: 2024/07). ResearchRabbit. https://www.researchrabbit.ai/.

[19] Barney Walker. (accessed: 2024/07). Citation Gecko. https://citationgecko.azurewebsites.net/.

[20] M. Weishuhn. (accessed: 2024/07). Inciteful. https://inciteful.xyz/.

[21] Anton Yeshchenko and Jan Mendling. 2024. A survey of approaches for event sequence analysis and visualization. *Information Systems* 120 (2024), 102283. https://doi.org/10.1016/j.is.2023.102283