Papers101: Supporting the Discovery Process in the Literature Review Workflow for Novice Researchers

Kiroong Choe   Seokweon Jung   Seokhyeon Park *   Hwajung Hong †   Jinwook Seo ‡

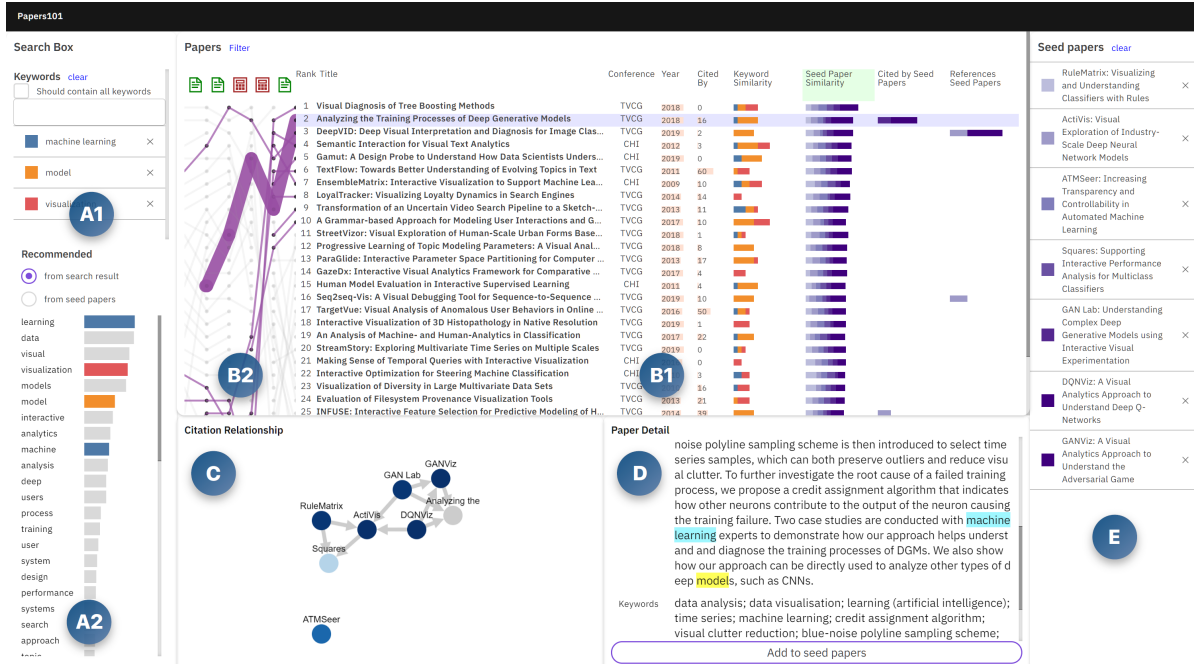Seoul National University, Seoul, Republic of Korea

Figure 1: An overview of Papers 101, an interactive system that accelerates the discovery of literature for novice researchers. (A1) Keyword Search Box and Keyword List which receive search keywords and list them; (A2) Recommended keywords section which recommends other relevant keywords to current search query; (B1) Ranking view which visualizes the ranking of the papers; (B2) History view which visualizes up to five previous rankings histories in parallel coordinates; (C) Citation Relationship view which enables users to validate the cohesiveness of the seed paper set; (D) Paper Detail view which provides additional information of the selected paper; and (E) Seed paper list which shows a list of seed papers selected by users.

ABSTRACT

A literature review is a critical task in performing research. However, even browsing an academic database and choosing must-read items can be daunting for novice researchers. In this paper, we introduce Papers101, an interactive system that supports novice researchers' discovery of papers relevant to their research topics. Prior to system design, we performed a formative study to investigate what difficulties novice researchers often face and how experienced researchers address them. We found that novice researchers have difficulty in identifying appropriate search terms, choosing which papers to read first, and ensuring whether they have examined enough candidates. In this work, we identified key requirements for the system dedicated to novices: prioritizing search results, unifying the contexts of multiple search results, and refining and validating the search queries. Accordingly, Papers101 provides an opinionated perspective on selecting important metadata among papers. It also visualizes how the priority among papers is developed along with the users' knowledge discovery process. Finally, we demonstrate the potential usefulness of our system with the case study on the metadata collection of papers in visualization and HCI community.

Index Terms: Applied computing—Document management and

text processing—Document management—Document metadata; Information systems—Information systems applications—Digital libraries and archives

1 INTRODUCTION

A literature review is a critical, but daunting task for novice researchers. Once the research topic is decided, the literature review process should follow; researchers need to review papers from various perspectives and synthesize the knowledge. However, novice researchers often face difficulties even at the initial phase, browsing relevant work from academic databases and choosing must-read items among them. This can be for many reasons such as vague ideas about appropriate search terms and less knowledge on exploiting academic metadata [15].

What further intensifies the problem is that the discovery of papers is not a single task but a process. For example, novice researchers who are less knowledgeable about terminologies in a certain domain often attempt to obtain a list of candidate papers by searching with common, broad keywords. Through the iterative queries, they eventually learn more suitable academic terms to use for the following search. Thus, until they reach a final understanding, various bottlenecks of novice researchers occur repeatedly and simultaneously. If individual challenges are handled by separate tools, it will increase a burden to integrate and switch between multiple contexts. The

*e-mail: {krchoe, swjung, shpark}@hcil.snu.ac.kr

†e-mail: hwajunghong@snu.ac.kr

‡e-mail: jseo@snu.ac.kr

absence of integrated tools is a problem that is pointed out among researchers [16], and it should be particularly addressed in systems intended for beginners.

We performed a formative study and design iterations with researchers from varying level of experience. As a result, we identified not only the unique tasks and needs that exist in the process, but also the empirical heuristics that can resolve it. This lead to the design of Papers101[1], an interactive system that supports novice researchers to discover papers relevant to their research topic. Papers101 provides opinionated perspectives on the selection of important metadata along with the stage and purpose within the discovery process. It also visualizes how the priority among papers are composed of those metadata and how it has changed over the users' path of knowledge discovery. In addition, Papers101 provides ways to validate and refine the current search query so that the users can be guided beyond their local knowledge and interests.

The contribution of our study is as follows.

1. We identified several types of needs and challenges that novice researchers face in the discovery stage of literature review process.

2. We built an interactive system that is dedicated to novice researchers so that they can accelerate the discovery of related works by simplified task flow and refined information.

3. We designed several components that visualize the process of knowledge discovery in order to facilitate novice researchers to obtain ideas and skills of the literature review.

2 RELATED WORKS

2.1 Metadata Collections and Exploration Tools

There have been attempts to build a collection of academic metadata and corresponding exploration tools. Vispubdata [8] provides a dataset of IEEE VIS papers from 1990 to 2018, including metadata such as titles, authors, references. Using the dataset, several exploration tools have been suggested [8, 9, 18, 20]. CiteVis2 and CiteMatrix, introduced in the same research, visualize an overview of reference and citation relationship within the dataset. KeyVis [9] aggregates author keywords of papers to recommend other related keywords when users input a keyword. Keshif [20] is a general exploration tool for rapid exploration of tabular data. Other exploration tools visualize ranking change over time [18], the distance between papers in terms of reference relationship [11]. Personal Informatics Paper Browser [5] is another metadata collection of papers in the personal informatics field, which includes human-labeled tags. These datasets and tools put primary focus on a static overview of the entire dataset, which can be viewed as an "explore" query where the location and target are both unknown. Our tool, however, is more dedicated to "browse" query, where the location(i.e., research topic) is known and the target is unknown.

2.2 Visualization of Academic Collections

In 2004, the Infovis competition was held to invite useful visualizations for the metadata of InfoVis papers. There has been various visualization research since then, which can be classified into two major categories: general or task-specific. The former research [1, 7, 10, 13, 22] express the innate characteristics of the academic dataset, such as chronological feature and network relationship. For example, most research utilize node-link visualizations to represent the reference relationship. While this type of visualizations can give users insight into the entire dataset, the utility can be limited depending on users and tasks. Task-specific visualizations [2–4, 12, 14, 21, 23] instead focus on the specific tasks or attributes regarding the dataset, such as the influence of individual researchers [14] and the selection of academic reviewers [12]. However, there are few visualizations directly target the literature review process.

2.3 Systems for Literature Review Workflow

Systems that support the literature review as a workflow are as follows. PaperQuest [11] inputs the papers from a user and recommends other related papers based on reference relationships. Users can keep updating the context by specifying papers of interest or papers to be read on the list. Sturm [15] surveyed on the (meta-)requirements of systematic literature search systems, and designed LitSonar which supports building expressive nested query over multiple data sources. LitSense [16] is designed to directly support the literature review workflow, such as discovering papers, organizing the library by categories, and sensemaking the library through several overview visualizations. While we target a similar research problem addressed by LitSense, we tried a different approach, putting more weight on paper recommendations and keyword extensions for novice researchers.

3 FORMATIVE STUDY

Prior to system design, we conducted a formative study to examine what difficulties novice researchers have and how experienced researchers address them.

3.1 Challenges of Novice Researchers

First of all, we interviewed four researchers with less than one year of academic research experience to investigate the difficulties which novice researchers experience in the discovery stage of the paper. Participants were 25 years old on average and yet to publish papers. Their field of study mainly lay within HCI. The specific research interests include algorithmic music generation, information visualization, and explainable AI. The interview questions are as follows.

- How and where they find materials related to the research topic (paper, article, websites, etc.)
- Which aspects of the paper they prioritize while searching for related researches (title, author, abstract, figures, etc.)
- What difficulties they have in identifying relevant research papers.

All of the interviewees were found to use academic search engine such as Google Scholar to identify relevant papers. While using such academic search engines, some interviewees had difficulty in searching because they were less knowledgeable about the terminology used in the field. They also felt that it was inefficient to go through a large number of candidate papers without priorities. A participant said the same holds for searching for references and citations one by one. Another challenge was that they lack confidence in whether the search results they obtained included all the candidate papers.

3.2 Heuristics of Experienced Researchers

Next, we conducted the following interviews with nine experienced researchers with more than three years of research experience, for those who have completed a master's degree. The interviewees had five to ten years of research experience, with an average of 6.7 years. They published an average of 4.8 papers in international conferences and journals.

- How they prioritize among a number of papers and authors.
- How they identify appropriate search keywords.
- How they determine whether they have found enough papers that are worth reading.

Experienced researchers mentioned the discovery of a small number of seed papers that are closest to the topic is the most important task. Seed papers was prioritized based on whether it is a survey paper, whether it is published recently, whether the citation count is high, and whether it is published in conferences or journals the

---

[1]The demo is available at https://kiroong.github.io/Papers101/.

researchers were interested in. While some researchers tried to look for search terms by referring to general articles prior to delving into the academic database, in most cases they found new topics and keywords through iterative search process. In addition, most researchers valued exploration through references and citations, and it was also a basis for determining whether the list of candidate papers has converged or not.

4 Design Considerations

Based on the findings from formative study, we derived the requirements of our systems as follows.

R1. Prioritizing Search Result

Most of the researchers we interviewed began a literature review from a number of candidate papers to collect various perspectives on the topic. However, novice researchers, unlike experienced researchers, read one paper by one without exploiting metadata and found it overwhelming. This process can be facilitated if they were helped to determine which paper to invest limited resources in reading. Therefore, the system targeting novices should have its criteria of prioritizing papers to not only reduce the time of the review process, but also encourage users to learn skills of assessing papers. Based on the feedback from experienced researchers, we chose to use the following criteria.

- **Keyword similarity:** how many times search keywords occur in the title, abstract, and author keywords of the paper.
- **Seed paper similarity:** how much the paper shares similar words with those in the current seed papers.
- **Publication year:** how recently the paper was published
- **Citation count:** how many papers in the database references the paper
- **Cited by seed papers:** the number of seed papers which is cited by the paper
- **References seed papers:** the number of seed papers which references the paper

R2. Unifying the Contexts of Multiple Search Results

Novice researchers reported that they felt lost while going back and forth between multiple search results. This problem is further aggravated when they started to explore the references of each paper. Thus, our system should support users to maintain a single unified overview against the frequent change of the search query and the perspective.

For this purpose, we adopted a simplified view where every priority criteria or perspective ends up being calculated as a single score assigned to each paper. For example, the task of finding papers that cite or are cited by a particular paper can be considered as giving one point to such papers and zero to those that are not. It enables us to take every result of user action as a list of papers with ranking. At the next section, we introduce how we exploited tabular view and parallel coordinates to visualize the state of a ranked list of papers as well as the change in ranking over time.

R3. Recommending Search Keywords

Novice researchers have vague ideas about academic terminologies. They often have to try searching with several general terms until they find a specific keyword that are prevalent in the academic field. Therefore, a 'keywords recommendation' feature could be an effective approach to guiding novice researchers to identify a set of relevant keywords candidates beyond their pre-defined search terms.

R4. Validating Seed Papers

Another challenge for novice researchers lied in the validation of seed papers. They were less confident about the query results: whether a set of candidate papers they identified were enough (i.e.,

representative), or whether there were irrelevant papers in the list (i.e., cohesive). The former requirement is directly addressed by our main goal. For the latter requirement, we adopted additional node-link visualization for the purpose of examining cohesiveness of seed papers.

R5. Other Requirements Derived from Pilot Study

As a practice of iterative design process, we demonstrated our primitive prototype and received feedback from two expert researchers (one postdoctoral student and one professor). We asked them to recall one of their past publications and find the most relevant papers using our prototype. Although they already knew what papers should be found, but they began by searching with broad search terms and justified each step of interaction. By observing them, we found that the following types of query were frequently used.

- **Query 1.** Listing all papers cited by, or referencing a specific paper.
- **Query 2.** Listing all papers sorted by a single field, not combination of them
- **Query 3.** Listing all papers published in a certain range of year
- **Query 4.** Listing all papers of a specific author.

In addition, we observed that they searched for a paper to accomplish one of the following different objectives.

- **Discovery:** Finding papers that cover various trends in the research topic.
- **Expansion:** Finding other relevant papers when a few seed papers were already found.
- **Seminal Paper:** Finding papers that has had a significant impact on the development of the research topic.
- **Serendipity:** Finding papers that are seemingly unrelated but may have potential relevance.

5 Papers101

In this section, we describe the final design of our system, named Papers101. Our system is composed of seven different views, all of which are designed according to the design requirements defined in the previous section.

A1. Keyword Search Box and List

The search box (Fig. 1:A1) inputs search keywords from users. Current search keywords are shown in the keyword list view below the search box. Each keyword is assigned a unique color distinguished by the hue channel, which is used in other views.

A2. Recommended Keywords

Recommended keywords view (Fig. 1:A2) aims to recommend other relevant keywords to the current search query (R3). It shows the most frequent words from either the top 30 papers in the current ranking or from seed papers. Keywords that are already in the search keywords are displayed by their color. Users can switch between the two modes by radio buttons.

B1. Ranking View

The ranking view(Fig. 1:B1) shows the current ranking of the papers. As described in section 4.1, there are six different sub-scores to assess the priority of each paper as defined in the previous section: keyword similarity, seed paper similarity, publication year, citation count, cited by seed papers, and references seed papers (R1). We adopted tabular layout, as it was frequently used for visualizing rankings of multi-attribute data [6, 19]. Six columns at the right part of the main view correspond to the sub-scores and represent their value with stacked bar charts. For example, the third column among them indicates the keyword similarity, the number of occurrence of

178

search keywords in that paper. A single stacked bar in that column represents the occurrences of all search keywords in the paper.

As a small multiple composed of stacked bar charts, this view visualizes the contribution of each keyword and seed paper to the current composition of ranking. The broader the area filled with its corresponding color, the greater the contribution of the keyword or seed paper. By default, the ranking is determined by the sum of all six sub-scores, where each sub-score is normalized to be between 0 and 1. Users can click header cells of the six columns to sort by values in the column (R5).

B2. History View

In the history view(Fig. 1B2), we visualize up to five previous ranking histories. This allows users to comprehend the change of ranking against multiple search queries and perspectives(R2). Several design choices were available for this purpose. For example, we could use either parallel coordinates or Lasagna plots [17] to represent a change in ranking. We used parallel coordinates because the representation can be considered familiar by novice researchers as it has the same encoding as the line charts that are widely used. To avoid visual clutters, only the parallel coordinate of current top 10 papers and a hovered element are made salient and other lines are faded out. At the top of each history, a glyph icon is shown to inform the type of query which resulted in the change. The magnifying glass icon indicates a search keyword, the document icon indicates a seed paper, and the calculator icon indicates the change in weight or filter. The color of the icon indicates the type of the operation: The green color indicates addition, and the red color indicates a deletion.

C. Citation Relationship View

Citation relationship view(Fig. 1:C) is designed to enable users to validate the cohesiveness of the seed paper set (R4). First, this view provides an overview of the citation relationship among seed papers with node-link visualizations. Users can find a seed paper that works as a herb of other seed papers, or a seed paper that is isolated from the others. Also, the color of a node indicates the cohesiveness of the paper. The brighter the color is, it is more unlikely for the paper to be ranked high when it is excluded from seed papers. Users can also check for nodes with low cohesiveness and verify if it is suitable for the seed paper set.

D. Paper detail View

Paper detail view (Fig. 1:D) allows users to check out additional information about the selected paper such as abstract, and author keywords. To help users determine whether the paper is suitable for their needs or not, search keywords in the abstract are highlighted by their corresponding color. It also contains a button to add the paper to the seed paper at the bottom.

E. Seed paper list

Similar to the keyword list view, this view shows a list of seed papers selected by users(Fig. 1:E). Each seed paper is also assigned a unique color, distinguished by brightness.

6 CASE STUDY

We collected the metadata of all papers published in CHI and TVCG before 2020. The following scenarios demonstrate the potential usefulness of our system.

6.1 Topic: Machine Learning Model Visualization

We assume a user interested in visualizing machine learning models. This scenario is also illustrated in Fig. 1. The user starts by entering three search keywords, *machine learning*, *model*, and *visualization*, and filter only the papers containing all three keywords. The system then ranks the papers based on published year, citation count, and the occurrence of search keywords. The user looks at the paper at
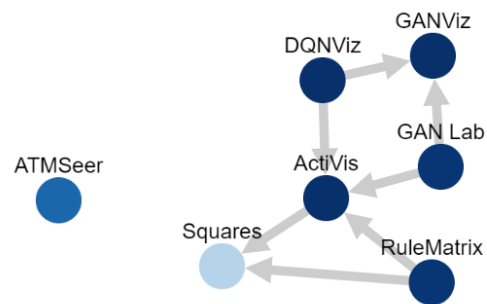

Figure 2: Citation relationship view in the case study 1

Figure 3: Ranking view in the case study 2

the top rank and adds it to the seed paper because they think it is related to machine learning visualization. This repeats four times, adding *RuleMatrix*, *ActiVis*, *Squares*, and *ATM Seer* into the seed paper set.

As the user found some seed papers, now they sort the entries by "references seed paper" column to browse which papers reference the seed papers. This results in studies such as *GAN Lab*, *DQNViz*, and *GAN-Vis* placed in the upper ranks. The user adds them to the seed paper set. To see if all seven seed papers are meaningful, the user can sort the papers in descending order of the seed paper similarity. Then, from the citation relationship view (Fig. 2), it can be found that *ATM Seer* does not have a reference or citation relationship with other seed papers. Also, *Squares* shows low cohesiveness, which means that it would not be ranked high in the current perspective if it was excluded from seed papers alone. The user removes the two papers from the seed paper set. Indeed, *ATM Seer* and *Squares* put more focus on performance analysis while other papers are more about inspecting the behavior of a machine learning model.

Finally, the user can select a "serendipitous" perspective where the papers are sorted by the sum of all sub-scores to have a balanced list of candidate papers. This results in the top ranking with papers closely related to the current seed paper, such as *LSTMVis*, along with papers with associated topics such as analyzing the training process of models.

6.2 Topic: Peer Support for People with Mental Health Problems

In this scenario, we show how Papers101 can facilitate users to find other related search keywords. We assume a user whose research topic is about peer support against mental disease. They can start from the general keywords, *mental health*, and *peer*. However, the user finds that there are only two papers containing both keywords in the dataset.

The user now tries to find other search terms. For this purpose, the user first adds the two papers to the seed paper set and change the perspective to show all papers that are cited by seed papers. Then, the word *communities*, *support*, *online*, and *feedback* can be found in recommended keywords.

Now the user adds those four keywords to the search keyword set. Then, among eight papers that are cited by the seed papers, one paper stands out to be containing many of those keywords (Fig. 3), suggesting itself as a option for the third seed paper.

7 DISCUSSION & CONCLUSION

In this paper, we designed Papers101, an interactive system to address difficulties in the literature discovery for novice researchers. We followed an iterative design process, and feedback from two experts showed the potential applicability of Papers101 into several usage scenarios such as discovery, expansion, and serendipity. Our case study also demonstrated possible use-cases in the research field of visualization and HCI. Still, there are opportunities to augment the contribution of our work.

Utilizing contextual information Firstly, the utility of search queries can be further increased by adding more contextual information to it. For example, our keyword recommendation is currently based on unigrams, which often bears ambiguity in its semantic meaning depending on what words come before and after it. In future works, we could contextualize search keywords by using n-grams or, ideally, contemporary natural language models. There will also be better distance metrics than the word frequency to compare between keywords and documents, or between documents. Similarly, the reference and citation relationship, which forms the heavy basis of exploiting seed papers, can be categorized into several citing motivation, such as survey, fundamental, or application.

Towards a systematized design study Secondly, our study has limitation on the design rationale and evaluation of the system. In terms of formative study, the feedback from four novice researchers may not be sufficient to draw robust conclusion. Formal study with a sufficient number of participants will not only strengthen the design rationale, but also lead to better understanding of users, such as their common strategies on searching papers and managing library. It is also recommended to gather more feedback from professional researchers to establish system principles that can be widely generalized. In terms of evaluation, a comparative evaluation would further demonstrate the effectiveness of our system. Possible options are comparing ours with popular academic search engines as Google Scholar, library management tools as Mendeley, and academic metadata collections with human-labeled categories as the Personal Informatics Paper Browser.

REFERENCES

[1] H. Alfraidi, W.-S. Lee, and D. Sankoff. Literature visualization and similarity measurement based on citation relations. In *2015 19th International Conference on Information Visualisation*, pp. 217–222. IEEE, 2015.

[2] G. Choi, S. Lim, T. Yoon, and K. Lee. Citation network visualization of reference papers based on influence groups. In *2018 IEEE 8th Symposium on Large Data Analysis and Visualization (LDAV)*, pp. 96–97. IEEE, 2018.

[3] J.-K. Chou and C.-K. Yang. Papervis: Literature review made easy. In *Computer Graphics Forum*, vol. 30, pp. 721–730. Wiley Online Library, 2011.

[4] A. G. Dias, E. E. Milios, and M. C. F. de Oliveira. Trivir: A visualization system to support document retrieval with high recall. In *Proceedings of the ACM Symposium on Document Engineering 2019*, pp. 1–10, 2019.

[5] D. A. Epstein, C. Caldeira, M. C. Figueiredo, X. Lu, L. M. Silva, L. Williams, J. H. Lee, Q. Li, S. Ahuja, Q. Chen, et al. Mapping and taking stock of the personal informatics literature. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(4):1–38, 2020.

[6] S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. Lineup: Visual analysis of multi-attribute rankings. *IEEE transactions on visualization and computer graphics*, 19(12):2277–2286, 2013.

[7] F. Heimerl, Q. Han, S. Koch, and T. Ertl. Citerivers: Visual analytics of citation patterns. *IEEE transactions on visualization and computer graphics*, 22(1):190–199, 2015.

[8] P. Isenberg, F. Heimerl, S. Koch, T. Isenberg, P. Xu, C. D. Stolper, M. Sedlmair, J. Chen, T. Möller, and J. Stasko. vispubdata. org: A metadata collection about ieee visualization (vis) publications. *IEEE transactions on visualization and computer graphics*, 23(9):2199–2206, 2016.

[9] P. Isenberg, T. Isenberg, M. Sedlmair, J. Chen, and T. Möller. Visualization as seen through its research paper keywords. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):771–780, 2016.

[10] R. Nakazawa, T. Itoh, and T. Saito. A visualization of research papers based on the topics and citation network. In *2015 19th International Conference on Information Visualisation*, pp. 283–289. IEEE, 2015.

[11] A. Ponsard, F. Escalona, and T. Munzner. Paperquest: A visualization tool to support literature review. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 2264–2271, 2016.

[12] M. Salinas, D. Giorgi, F. Ponchio, and P. Cignoni. Reviewernet: A visualization platform for the selection of academic reviewers. *Computers & Graphics*, 2020.

[13] Z. Shen, M. Ogawa, S. T. Teoh, and K.-L. Ma. Biblioviz: a system for visualizing bibliography information. In *Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation-Volume 60*, pp. 93–102. Citeseer, 2006.

[14] M. Shin, A. Soen, B. T. Readshaw, S. M. Blackburn, M. Whitelaw, and L. Xie. Influence flowers of academic entities. In *2019 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 1–10. IEEE, 2019.

[15] B. Sturm and A. Sunyaev. Design principles for systematic search systems: a holistic synthesis of a rigorous multi-cycle design science research journey. *Business & Information Systems Engineering*, 61(1):91–111, 2019.

[16] N. Sultanum, C. Murad, and D. Wigdor. Understanding and supporting academic literature review workflows with litsense. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pp. 1–5, 2020.

[17] B. J. Swihart, B. Caffo, B. D. James, M. Strand, B. S. Schwartz, and N. M. Punjabi. Lasagna plots: a saucy alternative to spaghetti plots. *Epidemiology (Cambridge, Mass.)*, 21(5):621, 2010.

[18] R. Vuillemot and C. Perin. Investigating the direct manipulation of ranking tables for time navigation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pp. 2703–2706. ACM, New York, NY, USA, 2015. doi: 10.1145/2702123. 2702237

[19] E. Wall, S. Das, R. Chawla, B. Kalidindi, E. T. Brown, and A. Endert. Podium: Ranking data using mixed-initiative visual analytics. *IEEE transactions on visualization and computer graphics*, 24(1):288–297, 2017.

[20] M. A. Yalçın, N. Elmqvist, and B. B. Bederson. Keshif: Rapid and expressive tabular data exploration for novices. *IEEE transactions on visualization and computer graphics*, 24(8):2339–2352, 2017.

[21] T. Yoon, H. Han, H. Ha, J. Hong, and K. Lee. A conference paper exploring system based on citing motivation and topic. In *2020 IEEE Pacific Visualization Symposium (PacificVis)*, pp. 231–235. IEEE, 2020.

[22] W. Zeng, A. Dong, X. Chen, and Z.-l. Cheng. Vistory: interactive storyboard for exploring visual information in scientific publications. *Journal of visualization*, pp. 1–16, 2020.

[23] X. Zhang, Y. Qu, C. L. Giles, and P. Song. Citesense: supporting sensemaking of research literature. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 677–680, 2008.