

PaperOrganizer: A Visual Bibliography to Support Structuring Scientific Literature

Bachelor's Thesis
submitted to the
Media Computing Group
Prof. Dr. Jan Borchers
Computer Science Department
RWTH Aachen University

by
Markus Horlemann

Thesis advisor:
Prof. Dr. Jan Borchers

Second examiner:
Prof. Dr.-Ing. Ulrik Schroeder

Registration date: 22.01.2021
Submission date: 19.04.2021

Eidesstattliche Versicherung

Statutory Declaration in Lieu of an Oath

Horlemann, Markus

Name, Vorname/Last Name, First Name

381304

Matrikelnummer (freiwillige Angabe)

Matriculation No. (optional)

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Arbeit/Bachelorarbeit/
Masterarbeit* mit dem Titel

I hereby declare in lieu of an oath that I have completed the present paper/Bachelor thesis/Master thesis* entitled

PaperOrganizer: A Visual Bibliography to Support Structuring Scientific Literature

selbstständig und ohne unzulässige fremde Hilfe (insbes. akademisches Ghostwriting) erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

independently and without illegitimate assistance from third parties (such as academic ghostwriters). I have used no other than the specified sources and aids. In case that the thesis is additionally submitted in an electronic format, I declare that the written and electronic versions are fully identical. The thesis has not been submitted to any examination body in this, or similar, form.

Ort, Datum/City, Date

Unterschrift/Signature

*Nichtzutreffendes bitte streichen

*Please delete as appropriate

Belehrung:

Official Notification:

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

Para. 156 StGB (German Criminal Code): False Statutory Declarations

Whoever before a public authority competent to administer statutory declarations falsely makes such a declaration or falsely testifies while referring to such a declaration shall be liable to imprisonment not exceeding three years or a fine.

§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt

(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

(2) Strafflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.

Para. 161 StGB (German Criminal Code): False Statutory Declarations Due to Negligence

(1) If a person commits one of the offences listed in sections 154 through 156 negligently the penalty shall be imprisonment not exceeding one year or a fine.

(2) The offender shall be exempt from liability if he or she corrects their false testimony in time. The provisions of section 158 (2) and (3) shall apply accordingly.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:

I have read and understood the above official notification:

Ort, Datum/City, Date

Unterschrift/Signature

Contents

Abstract	xi
Überblick	xiii
Acknowledgements	xv
Conventions	xvii
1 Introduction	1
1.1 Outline	4
2 Related work	5
3 User Research Study	13
3.1 Context & Aim	13
3.2 Setup	16
3.3 Results & Evaluation	18
3.3.1 Background Information	18
3.3.2 Research Workflow Approaches	20

3.3.3	Observations	23
3.4	Interpretation & Consequences for the tool PaperOrganizer	28
4	The Tool PaperOrganizer	33
4.1	Description of the Tool and Design Ideas . .	33
4.2	Important features & their benefits	40
4.3	Limitations	44
5	Summary and future work	47
5.1	Summary and contributions	47
5.2	Future work	48
A	User study consent form	51
B	User study sheet	53
	Bibliography	59
	Index	63

List of Figures

2.1	Interface showing the semantic structure in digital libraries	6
2.2	Visualized faceted information spaces by PivotPaths	9
2.3	Interface of JigSaw	9
2.4	Interface of the Action Science Explorer	10
2.5	Interface of VisualBib	11
3.1	Evaluation of the frequency of literature research	19
3.2	Evaluation of the amount of performed literature research	19
3.3	Evaluation of the reasons for literature research	20
3.4	Evaluation of the amount of published scientific papers	20
3.5	Evaluation of how many scientific papers are used during literature research	20
4.1	Overview of the interface of the tool Paper-Organizer	34

4.2	Overview of different paper formats	35
4.3	Concept of different paper element states . .	35
4.4	Related work visualized with PaperOrganizer	41
4.5	Close up view of a custom research cluster .	42
4.6	Network view with the tool PaperOrganizer	42
A.1	Informed Consent Form	52
B.1	User study sheet general question and re- search task	54
B.2	User study sheet research task part two . . .	55
B.3	User study sheet with further questioning . .	56
B.4	User study sheet attention points	57

List of Tables

Abstract

Literature research is an important part of the scientific community and with a rapidly increasing amount of literature, it is helpful to maintain a good structured overview. Getting to know the own bibliography can often be a time-consuming challenge, where a clear overview and a well-organized collection of documents can have a major impact.

In this work, we tried to understand in-dept problems of organizing and working with scientific paper collections during literature research and built, based on our findings the tool PaperOrganizer. The here presented study has thereby shown that not only is there a great need for free interaction in literature research, but it also provides important insights into the process of understanding scientific paper and working with a bibliography. After this study we came up with the tool PaperOrganizer, which helps organizing and structuring local literature collection interactively. The aim is to get a better overview about the own subject area in order to ensure more efficient literature research, as well as to improve familiarizing with old bibliographies.

The PaperOrganizer tool is the first approach that supports individually interactive and personalized organization of literature, while displaying relevant information from and between papers.

Überblick

Literaturrecherche ist ein wichtiger Bestandteil wissenschaftlichen community und mit einer stark ansteigenden Menge an Literatur ist es hilfreich, einen gut strukturierten Überblick zu behalten. Das Kennenlernen der eigenen Bibliographie kann oft eine zeitaufwändige Herausforderung sein, bei der ein klarer Überblick und eine gut organisierte Sammlung von Papern einen großen Einfluss haben kann.

In dieser Arbeit haben wir versucht, tiefgreifende Probleme zu verstehen, die bei der Organisation und der Arbeit mit wissenschaftlichen Papern während einer Literaturrecherche auftreten, und basierend auf unseren Erkenntnissen das Tool PaperOrganizer entwickelt. Die hier vorgestellte Studie hat dabei gezeigt, dass nicht nur ein großer Bedarf an freier Interaktion in der Literaturforschung besteht, sondern auch wichtige Erkenntnisse zum Verständniskern und zur Einarbeitung von Literatur präsentiert. Nach dieser Studie haben wir das Tool PaperOrganizer entwickelt, mit dessen Hilfe die lokale Literatursammlung interaktiv organisiert und strukturiert werden kann. Das Tool bietet dabei einen besseren Überblick über den eigenen Themenbereich, um eine effizientere Literaturrecherche zu gewährleisten und erneutes Einarbeiten in alte Bibliografien zu erleichtern.

Der PaperOrganizer ist das erste Tool, das die individuell interaktive und personalisierte Organisation von Literatur unterstützt und relevante Informationen von und zwischen Papern anzeigt.

Acknowledgements

I would like to thank Prof. Dr. Borchers and Prof. Dr. Schroeder for examining this thesis.

I also want to thank my supervisors Oliver Nowak and Adrian Wagner for supporting me during the entire time of my bachelor thesis, giving me great advises. During the bachelor thesis I was able to learn a lot, for which I am very grateful. Finally, I would like to thank all of the participants who took part in the time-intensive study.

Conventions

Throughout this thesis we use the following conventions.

Text conventions

The whole thesis is written in American English. The first person is written in the plural form. Unidentified third persons are described in female form.

Chapter 1

Introduction

Literature research is anchored in the process of acquiring knowledge and is an important part of scientific work. With a permanent increase in literature publications, however, it becomes more difficult to keep an overview. The research areas are becoming more and more diverse or dense, which urges the manageability of literature research to meet these demands.

According to research from the University of Ottawa, the total number of 50 million published scientific papers was passed in 2009 and approximately 2.5 million new papers are published each year [Jinha, 2010] [Ware and Mabe, 2015]. This amount of knowledge can be exhaustive or even overwhelming.

Fully comprehending papers poses several challenges for the reader. Not only is there a large variety of papers, some papers require prior knowledge of a topic area, so that the reader first has to comprehend the knowledge of other papers of the area [Ware and Mabe, 2015]. Accordingly, these processes turn out to be very opaque and sometimes difficult to view [Davidson et al., 1998].

In the following we roughly divide the process of a literature research into three main areas.

Literature search and
bibliography
extension

The literature search in particular is a very broad field and includes divers methods to find and compare important papers [Ware and Mabe, 2015]. Common tools are Google scholar ¹, the ACM Digital Library ³, Connected Papers ⁵, Scopus ⁷, OpenCitations ⁹ and Web Of Science ¹¹. Many visual analytical literature methods also provide assistance for paper search, but overlook that sorting and organizing papers has a huge importance after literature has been found. Some approaches can be found in the related work in chapter 2.

Management of
references

Besides research tools and search engines, there is also a wide variety of tools that help organizing or working with literature by providing reference management, like Cross-Ref ¹³, Zotero ¹⁵, Qiqqa ¹⁷, Jabref ¹⁹, and many others. These programs provide a good overview, especially for individual papers and their references, and make it easier to work with citations Ahmed and Al Dhubaib [2011].

The own
organization and
familiarization with
individual papers

After Literature search and reference management the process follows with structuring and understanding the general context up to understanding the literature in-depth. This state is often time-consuming and cognitive demanding.

Therefore it is an important and necessary step to find efficient approaches to help users. The focus is on optimizing literature organization, to keep this process as simplified and time-saving as possible, in order to sustainably optimize literature research.

Even in spite of these clearly defined shaping, these three main areas of literature research are often difficult to separate from one another and have smooth transitions, which is shown by software tools for these areas. In addition,

¹<https://scholar.google.com/>²

³<https://dl.acm.org/>⁴

⁵<https://www.connectedpapers.com/>⁶

⁷<https://www.scopus.com/>⁸

⁹<https://opencitations.net/>¹⁰

¹¹<https://www.webofknowledge.com/>¹²

¹³<https://www.crossref.org/>¹⁴

¹⁵<https://www.zotero.org/>¹⁶

¹⁷<http://www.qiqqa.com/>¹⁸

¹⁹<https://www.jabref.org/>²⁰

these areas are repeatedly run through in the process of expanding the bibliography, which makes literature research an iterative process [Ware and Mabe, 2015].

In the past, many approaches have tried to support users with visual exploration of literature as presented in chapter 2. But people behave differently when it comes to the actual organization and incorporation into individual papers, which also follows out of our user study.

Therefore, the set goals of this thesis is, not only to understand Literature Research in-depth to identify individual procedures and support their research process, but also to create the PaperOrganizer tool, which is based on this user study and fulfill the following tasks:

- Supporting the user after an extensive search of papers with a tool that visualizes their bibliography: One of the key benefits of using visualization is the improvement on cognitive and perceptive understanding in quantitative and verbal information [Mohamad et al., 2018]. Visual concepts lead to a better overview that is not initially available in a separate local bibliography [Synnestvedt et al., 2005].
- Organizing a local paper storage: In the past, many approaches offered special organization in search and reference management, whereas research tasks aimed at an independently organized bibliography.
- Offering a customizable tool that supports structuring research topics: Interactive designs could be a promising approach to enable sustainable personalized organization.

1.1 Outline

In the following, we describe the related work in Chapter 2, focusing on tools that help organize biographies, as well as visual approaches for analyzing scientific literature.

In Chapter 3 we will go into detail on the user study. First, we describe how the study is structured and what goals are set, especially with a view of the PaperOrganizer tool. Then, we show the findings of the evaluation and draw concrete conclusions, where we derive features for the PaperOrganizer. We also go into more detail about various results that emerged from the study.

After that, we focus specifically on the PaperOrganizer tool in Chapter 4. Here, we go into more detail of tool features but also limitations of the software.

Furthermore, we give in Chapter 5 conclusions about how far we have already achieved our goals. Finally, we conclude a summary of the thesis and show possible future work.

Chapter 2

Related work

In order to optimize the process of literature organization, we focus primarily on visual approaches for analyzing scientific literature as well as organizing and arranging found literature. Roughly nine areas emerge that have presented concepts to shape the organization and analytical visualization of literature.

A survey carried out by Federico et al. [2016] examines 109 different approaches of visually analyzing scientific literature created between 1992 and 2016. This survey distinguishes between two types of categories: Data types (text, citations, authors, metadata, and combinations thereof) and analyzing types (exploring and comparing single entities, investigating relations between entities, finding patterns on various levels, exploring the temporal dynamics, and finding complex connections between phenomena). The approaches presented here also show the underlying complexity and provide information about the possible use of this method. Some of them will be presented in more detail below.

A quite general method of exploring interactive data is given by Davidson et al. [1998], who presented an approach that discusses a set of properties that a large-scale visual data exploration tool should have. Based on that they provided VxInsight, a tool to explore data like huge bibliographies as a 3D environment by a dynamic terrain generation.

A survey on visual approaches for analyzing scientific literature & patents

Knowledge mining using VxInsight

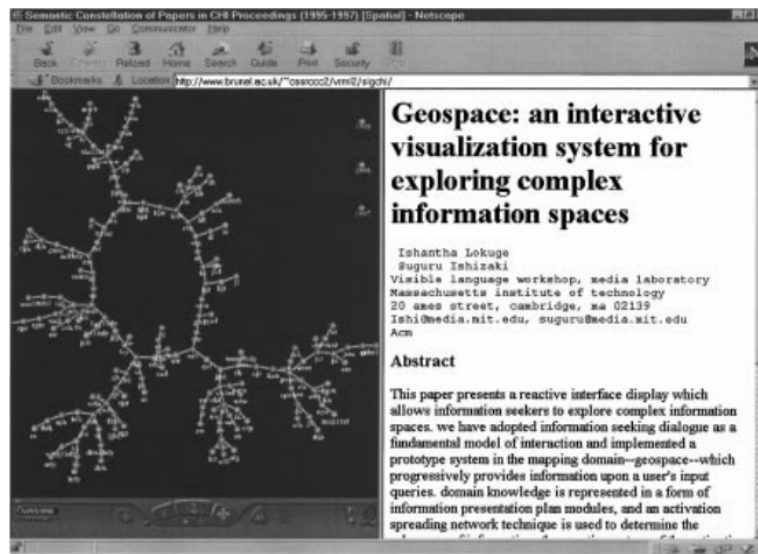


Figure 2.1: Interface by Chen [1999] showing the semantic structure extracted from the ACM SIGCHI collection 1995-1997. The graph is generated by a document similarity matrix. Clicking on a sphere displays content of the document in the right screen.

Height values are thereby based on density of similarities in the data, leading the user interface to 3 dimensional terrain, to identify structures and patterns.

Visualizing semantic spaces and author co-citation networks in digital libraries

Chen [1999] published visually broad-based semantic spaces and author co-citation networks in digital libraries, as shown in Figure 2.1. This approach mainly serves the analysis of document collections and provides revealing insights of the data, which follows from the attached study.

Visualizing citation networks with CiteWiz

In the course of visualizing citation networks, Elmqvist and Tsigas [2007] brought further techniques into this domain with the tool CiteWiz. In general it deals with authors, citations and metadata, features based on three different views. Including visual ranking within citation networks based on node sizes, timeline visualizations and influence visualization using growing polygons. To archive this it uses a central citation dataset in a simple XML-based file format.

CiteWiz is one of the tools that have been proposed in the

occasion of the InfoVis [2004] contest to visualize and analyze 7 years of the InfoVis conference proceedings. Other tools that emerged from this contest include BiblioViz [Shen et al., 2006], CiteVis [Stasko et al., 2013], PaperLens [Lee et al., 2005], CiteWis [Elmqvist and Tsigas, 2007] and PaperCube [Bergström and Atkinson, 2009], all with some unique features and visual characteristics.

BiblioViz published by Shen et al. [2006] presents different types of paper arrangement systems, like tabular metadata based comparisons between papers up to 2D and 3D network visualization views of bibliographic data. A special feature here is the dual window in which both the tabular view and the network view can be viewed at the same time. This perspective tries to clarify the context of papers. Multi-windows are particularly popular for analyzing paper collections, as it is the case with PaperLens.

Visualizing
bibliography
information with
BiblioViz

PaperLens, a tool presented by Lee et al. [2005] offers tightly couples views of topic popularity in literature research. Additionally it is capable of paper and authors rankings, visualizing the degree of separation of authors and gives meta data overview to analyze topic trends. Multiple views enable a simplified view of difficult-to-see data interconnections. Later this tool was extended to visualize 23 years of the CHI conference proceedings.

Understanding
research trends in
conferences using
PaperLens

CiteWis by Stasko et al. [2013], not to be confused with CiteVis, offers a user interface in the form of a matrix with selectable papers, in which publications which citing it and its citations are highlighted in color, to quickly understand the range of a focused paper.

Exploring conference
paper citation with
CiteVis

The last tool that emerged from the InfoVis conference that we want to highlight is PaperCube by Bergström and Atkinson [2009]. This tool deals with visual networks, hierarchy elements, and timeline structures. It also integrates switching between views without losing focus. Among them a circle view of papers in citation networks, a paper per year table with an interactive view of referenced papers and a tree map view with references as child's. Furthermore the user study of PaperCube revealed that with the right data set, PaperCube would greatly improve the daily

Web-Based
Visualizations using
PaperCube

workflow with Literature. In addition PaperCube reduced the cognitive load of a researcher by making it easy to focus and show relevant information.

Detecting and
visualizing emerging
trends with
Citespace

Citespace II is yet another tool from the field of analyzing bibliographic prepared datasets and databases. This approach by Chen [2006] goes deeper into the analysis of detecting and visualizing emerging trends and transient patterns in scientific literature. Those concepts were then further developed through various studies like the study by Synnestvedt et al. [2005] to Citespace III by Chen et al. [2010], which is a freely available trend analyzing domain visualization tool for scientific literature.

Mohamad et al. [2018] used a more complex and interactive approach to visualize Dental research data in Scopus according to co-authorship networks, Keyword co-occurrences and cited references pattern and citation burst.

Exploration of large
research areas

In the following time, various methods of analysis and visualization were used to explore large areas of subject areas and conferences: 20 Years of Four HCI Conferences by Henry et al. [2007] and Statistical Analyzes of CHI Kaye [2009] show how different methods contribute to a deeper and clearer understanding of a whole topic landscape. Another large-scale literature review was then provided by Rodrigues et al. [2014] in which 8480 publications are explored and clustered according to custom categories in order to track better affiliations and to find areas that were little explored.

Visualized faceted
information spaces
with PivotPaths

The tool PivotPaths by Dörk et al. [2012] brought yet another view of Literature Networks. The tool uses graph representations to show author relations, publications, and keywords all views combined in one graph divided by 3 layers as seen in Figure 2.2. The striking factors here are the various animations when interacting with the graph to extract details from it. Here again the visualization is limited by the explored dataset of papers in the fields of HCI, which contains publications up to 2012. A demo version can be found online.

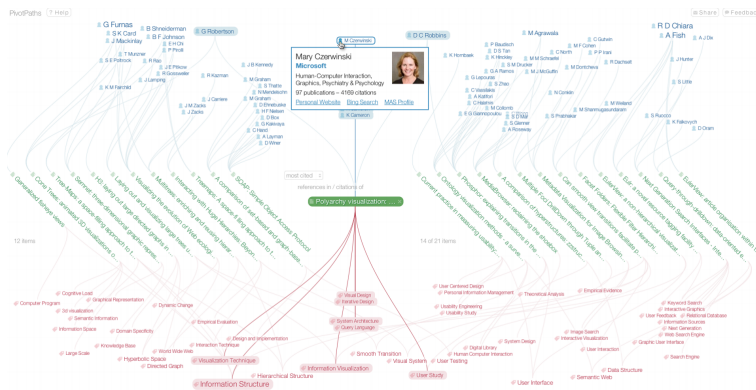


Figure 2.2: Interactive interface of PivotPaths by Dörk et al. [2012], which is separated into three main regions: People, resources, and concepts.

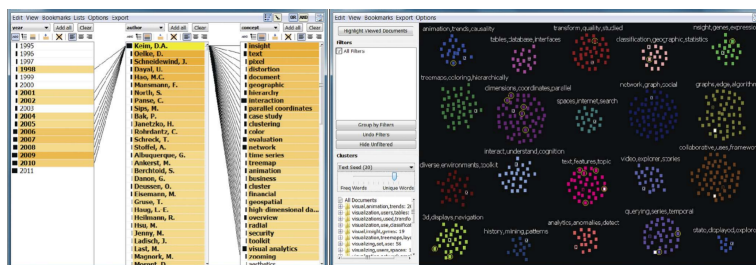


Figure 2.3: Interface of JigSaw showing a list view of years, coauthors, and concepts connected to a selected author (left) and a cluster view (right) with different clusters of related papers with highlighted selected author papers [Görg et al., 2012].

Besides this, there are also more interactive visualization concepts presented in JigSaw. The tool by Görg et al. [2012] provides multiple views of connected entities and a document cluster view of related papers as presented in Figure 2.3. Having a list view as well as a graph view is also one of the features that made it into the PaperOrganizer tool.

A drastically new way of representing citations networks was then introduced with the tool Citeology by Matejka et al. [2012]. This Tool offers a chronological pre-built interactive map of 3,500 papers, which has been extracted from

Combining computational analyses and interactive visualization in JigSaw

Visualizing paper genealogy with Citeology

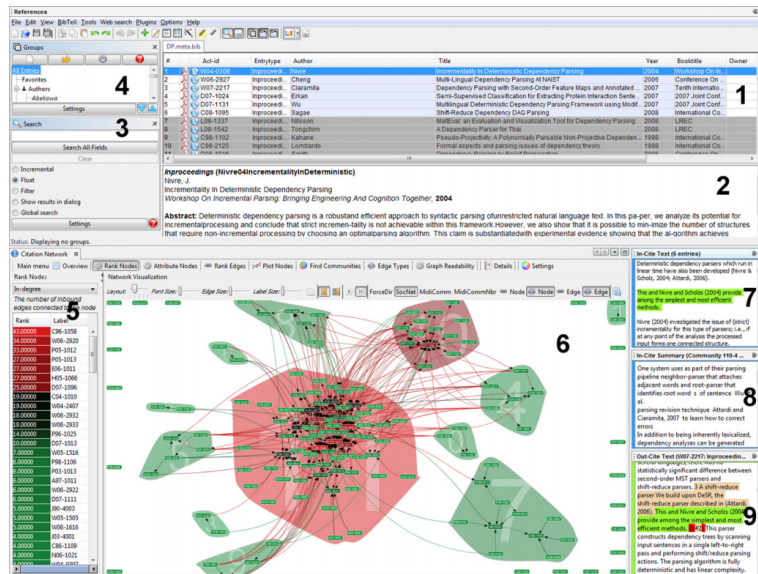


Figure 2.4: Interface of the Action Science Explorer with reference management (1–4), citation network statistics (5–6), citation text (7), multi-document summary (8), and a Full Text view [Dunne et al., 2012].

the ACM Digital Library. It provides a tree structures of paper to determine the citation level to another by selecting one paper and hovering over others.

Visualizing
information flow of
literature

Another very unique way of visualizing information from scientific Literature collections is Eigenfactor: Stefaner et al. [2009] presenting visualizations to explore emerging patterns in scientific citation networks by calculating a measure of importance for individual journals as well as measuring their citation flow and hierarchical clustering. Therefore a subset of 7000 scientific journals over the past decade was used. The data is visualized in 4 different diagrams showing citation patterns, change over time with an alluvial diagram, clustering with indicated incoming and outgoing citation flow and a graph visualization where journals, which frequently cite each other, are closer together.

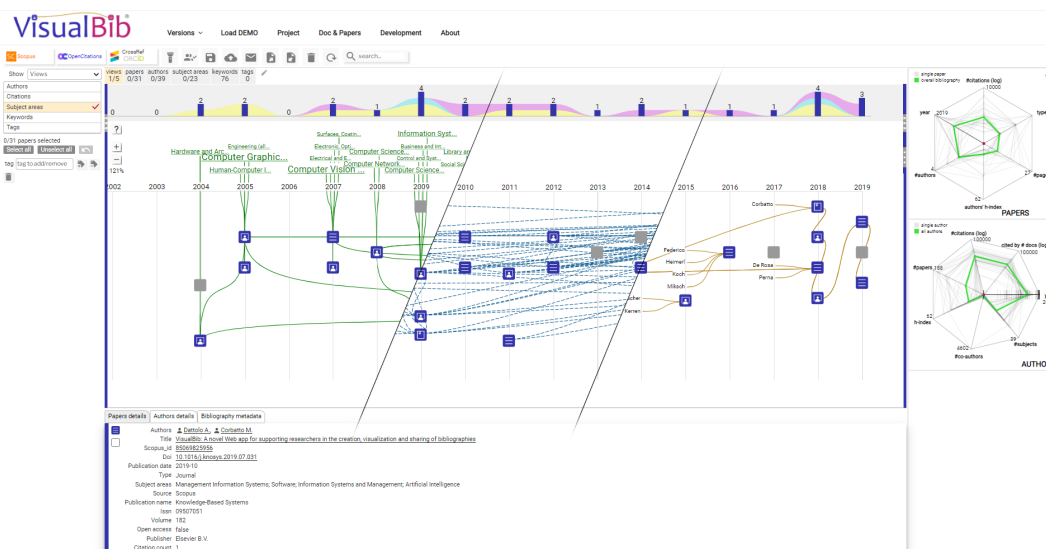


Figure 2.5: Interface of VisualBib showing 3 views of loaded papers: Keyword relations (left), citation relations (middle), author relations (right) and further information of papers at the bottom [Dattolo and Corbatta, 2018]

A more recent Java tool so-called CitNetExplorer by Van Eck and Waltman [2014], than allowed the visualization of citation networks, by expanding, reducing and clustering of publications, offering the opportunity to discover personal paper collections.

Visualizing custom citation networks with CitNetExplorer

Stepping in the direction of organizing bibliography, a popular approach is reference management such as provided by Zotero [Ahmed and Al Dhubaib, 2011]. Besides reference management it is also capable of sorting papers, offering paper notes, and an overview of provided metadata of papers.

Dunne et al. [2012] offer rapid understanding of scientific paper by presenting a tool that further combines reference management with citation network statistics and visual literature analysis. The user interface is presented in Figure 2.4 and shows a loaded bibliography. The Software is a non usable research prototype.

Rapid understanding and analysis of literature collections with the Action Science Explorer

One outstanding example of visualizing bibliographies is VisualBib by Dattolo and Corbatta [2018]. It combines many concepts that we have seen before and provides a

Narrative views for customized bibliographies with VisualBib

browser tool to present metadata and paper-relations. Papers can be added and arranged by their published year, author and title as seen in Figure 2.5. Furthermore its possible to show author dependency's, citation relations, connected subject areas and keywords. Thereby it is capable to import and export this customized view, offering an opportunity to build up a customizable and shareable bibliography view.

None of these approaches work with a local bibliography that supports the user by visualizing his personal storage in a supportive way, to get a better view of what he is working with. In addition, these tools also do not offer any free interaction, to work closer with the content of papers.

A pattern approach
to interaction design

In order to better understand the design of interactive systems, Borchers [2001] presents an approach that uses pattern languages which shapes and consolidates major milestones of interactive user interface design in software development, human-computer interaction and the application domain. The knowledge of some of the concepts also shape the design of the PaperOrganizer published in this thesis.

A survey on text
visualization
techniques

Moving over to the last research area, text visualization. A visual overview of the various text visualizations is then presented in a survey by Kucher and Kerren [2015]. Here 141 different interactive web-based browser techniques of text visualization are presented, which gives insights in the current state of the text visualization field.

Active reading and
its discontents

Last but not least, active reading also meets some of the challenges that literature research encounters. Tashman and Edwards [2011] present the current situation, problems and ideas around active reading based on observations of participants. In their study they report as one tension that comparing and integrating information between texts causes problems for participants. In addition, the study shows at which points there are sources of frustration when working with document windows. This study suggests, among other things, hybrid systems as a solution to support better different parts of the AR process.

Chapter 3

User Research Study

3.1 Context & Aim

Getting acquainted with a researched paper collection can be a very exhausting process. Often it is not necessary to have papers pre-sorted, but to have them arranged in a larger picture. In addition to some individual processes, literature research often includes inefficient, identical processes or similar recurring problems. Therefore, in order to create a tool that helps organizing literature, it is necessary to understand the process of literature research precisely and in depth.

Looking at the overall concept of research, orientation and classification of scientific literature, our focus is on how users specifically deal with literature, manage it and integrate it into their work. It is particularly important to recognize same patterns in order to help users as accurately as possible.

With a view to the PaperOrganizer, we concentrate on the target group of users who already have a little up to extensive experience with literature research. For this target group, the program should be as intuitive as possible. It should offer an opportunity to better organizing both old and new literature research and thereby supporting a more efficient view of the topic. In addition, users should to lose

a certain distance in the abstraction of papers in order to familiarize themselves with a paper collection faster and with a better precision.

For this setup, there are 6 general important questions that we want to shed more light on:

1. Where do participants lose a lot of time while working with literature?

Above all, we try to understand how and what is done with papers in order to identify and optimize inefficient procedures. More specifically, we are looking for some patterns that are made, to become more familiar with paper, but which could be solved in an efficient way, or could be optimized by minimizing the repetition of processes without losses. Therefore, with regard to the Software, we try to enable more efficient organization.

2. What general system, method or principle do most of the participants use?

General principles of organization allow clear courses of action to be identified in the course of the research to better understand intuitive processes. This also makes it possible to characterize workflows that play an important role for the design of the PaperOrganizer and promotes intuitive handling. Thereby it is also possible to infer guidelines for action in research tasks, which could be supported by the tool as well.

3. Are there more specific procedures that many participants handle the same, or are there technical processes that show similarities?

Same working procedures offering opportunities for a better understanding of deeper structures in research tasks. This is important to draw conclusions about important and common processes, but also to find practical processes that are necessary or helpful to many in the research. In addition, it is possible to identify potential for methods that can make work easier, which may also turn out to be useful features for the Paper Organizer tool.

4. What are important methods or features when it comes to organizing literature?

Determining priorities enables methods and important pro-

cesses to be recorded. This allows features to be emphasized and given greater relevance during implementation, which can then enable more intuitive handling with the PaperOrganizer. Additionally it can strengthen important workflows in the literature research, as well as avoiding feature creep.

5. Where is interaction missing?

With the help of improvements in the interaction it is possible to feel more comfortable and familiar with online documents. The aim is on losing the distance that the virtual abstraction of online documents entails. An important element here is the gain in control: If users get the feeling that they have greater influence on a bibliography, mental concepts can be adapted and integrated more easily. Conversely, there is no need to learn a concept that allows only one view of the bibliography by dragging the user away from his personal view. The focus here is particularly on interactive improvements in the process of organization: File explorer and reference management tools only offer simple and rigid lists. Even approaches that we have presented in the related work, chapter 2, presenting so far mostly only static views on literature, which severely restricts users.

The last main question of our study deals with a general view of literature: 6. How do users get overviews of their subject?

This question is probably the most central question of the user study. It is not only important to find out what activities are taken to achieve a better view, but also the tendency to achieve these measures. The aim is to recognize the needs of the user, which is based on the knowledge gained from individual papers as well as the general view of bibliographies. Experts in particular have already worked deeply into systems that may not be specifically tailored to the work, but accepted as the best solution. With reference to the tool, this is the consequence of the design of the clarity of paper collections. The knowledge hidden here decisively determined the organization of the clarity of paper collections.

In summary, the aim of the user study is to find out where advantages are bent in the organization, where important

procedures are taken, at which points intuitive processes occur that make working with the tool easier and which practical features are useful when organizing literature.

3.2 Setup

In total, we conducted 15 meetings via Zoom. On average, these interviews lasted 50-60 minutes. The participants of the study were computer science students and graduates. In addition to the audio transmission, participants screen-shared the interview for better traceability. Therefore, we divided the user interview into three sections in order to attack the broad goals and to cover most of the details in research practices, to understand them in depth:

The first section is about the data collection of the participants, to get some more general information of the participants and their research background.

This is followed by tracking the workflow of participants, in which current and old research projects of the interviewee are discussed.

Finally, the third part of the interview includes a more detailed survey in which current or old research assignments are dealt with in more detail.

The questionnaire with a brief introduction for the participants, as well as all three parts including questions and question directions can be found in appendix B.

General Data of the participants

The collection of data from the participants serves to differentiate between groups of people with regard to different empirical values, as well as the collection of basic data. The PaperOrganizer requires this data know which demands the tool has to meet.

When dealing with research assignments it is important to determine the experience of participants by categorizing them, in order to more precisely determine which groups tend towards which procedure. To differentiate research experts from regular participants, we asked participants about the number of literature researches they conducted before, their circumstances and number of published scientific papers.

Furthermore, marginal data is recorded, in order to be able

to devalue technical challenges that the tool has to face.

In the second phase an attempt is made to determine the workflow of participants in the literature research.

Participants were given a start paper published by Mlakar and Haller [2020], which deals with non wearable textile interfaces.

Then, we simulate a real literature research scenario: Participants are tasked to find one to two more papers on the topic of textile interfaces and citing each paper once in an environment in which they usually work with papers.

The start paper was chosen in such a way that it is current, not too overwhelming and generally an accessible subject area, although this may differ for some participants. Since for the most part only computer science students or assistants took part in the study, we did not include a selection of topics. In addition to the research assignment, the participants were also interviewed during this task in order to find out more precisely which reasons lead to which decisions.

Workflow
observation

In the end of the interviews, there is a survey on participants current and past literature research process.

Here, particular attention is paid to differences in the procedure with regard to the second part of the study in order to classify the workflow in a larger picture and to inquire about various details in procedures and recommendations, as well as difficulties in the process or uncovered points. This follow-up survey also serves to see how large bibliographies are handled and organized. Particular attention is paid to the following points:

Further survey on
current and past
literature

- Where do processes take place repeatedly?
- Where are similarities in workflows?
- How is literature structured?
- How is organization approached and carried out?
Where are files stored? How are these dealt with?
- At which points in the research do users feel restricted?

- Where is more interaction needed? Where is a better overview required and possible?
- What do people pay particular attention to? What is important or normal for them?

3.3 Results & Evaluation

3.3.1 Background Information

In total, we conducted our survey on five female and ten male participants. The participants were on average slightly over 23 years and ranged from 21 to 29 years.

Participants stated that they conduct literature research on average monthly to every six months. Occasionally there are also deviations in which it is stated that performed research takes place only annually or vice versa several times a month, presented in Figure 3.1.

Research that has already been carried out amounts to an average of almost five to ten papers. None of the participants has never carried out a scientific research. Two people stated that they had more than 20 researches performed. An overview is given in Figure 3.2.

Most participants cited theses or seminars as the main reason for doing the past research tasks, as shown in Figure 3.3. Either in the form of writing papers themselves or reviewing them. Publications were also given five times as a common reason for literature research.

Figure 3.4 shows the distribution for the number of papers already published. Six people stated that they had already published one to five papers and nine people stated that they never published papers before. So the background of the participants amounts to a range of zero to five paper publications.

The last question of the first part of the study is about the average number of papers that are used and more closely looked at during literature research. Figure 3.5 shows the distribution of the participants answers. Participants were allowed to specify a freely selectable value or a range which represents the average number of papers that they actively

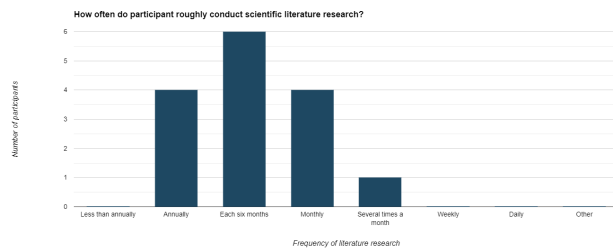


Figure 3.1: Evaluation of the frequency of literature research. The period is shown on the x-axis and the number of people is shown on the y-axis.

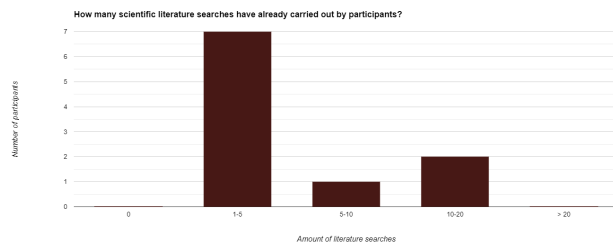


Figure 3.2: Evaluation of the amount of literature search.

work with. On average, the value is 28.75, therefore almost 29 papers that participants use. However, the arithmetic mean is strongly influenced by the spread of the distribution. Therefore, it makes sense to look at the median, which is 15 papers. In addition, the number of papers varies between five and 217 papers, with high occurrences between ten and 20 papers.

Based on those results we differentiate in the following between more and less experienced participants. More experienced participants are categorized as such if they have already been involved in publications or have carried out more than five literature searches.

Of course, the individual practical or empirical knowledge also depends on individual other factors, which we do not take into account and do not consider in detail here for the sake of simplicity.

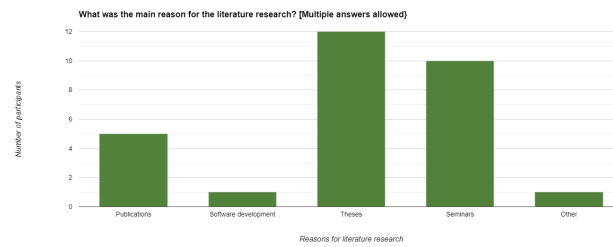


Figure 3.3: Evaluation of the reasons for literature research.

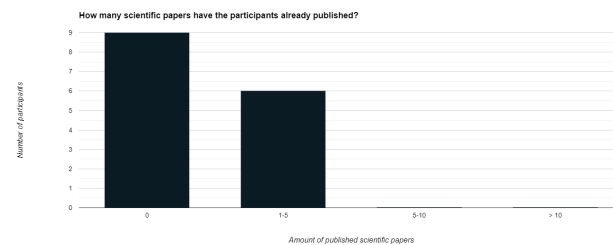


Figure 3.4: Evaluation of the amount of published scientific papers.

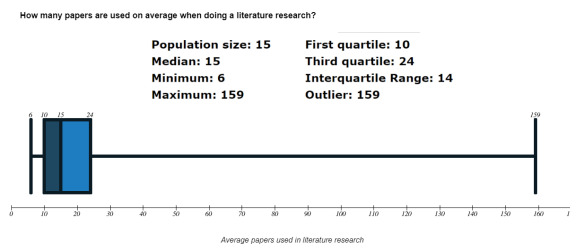


Figure 3.5: Evaluation of how many scientific papers are used during literature research. The box plot shows the distribution and a large deviation.

3.3.2 Research Workflow Approaches

In the general procedure of the literature search and organization clearly recognizable workflows emerge, which we will examine in more detail here. In most cases the ini-

tial paper is skimmed through without taking any notes. The focus here is on the title, year, keywords and abstract, which are read sometimes several times. Important or complex words are often looked up. Subsequently, people like to scroll through the paper to get an overview of the length of the paper and chapter division. Participants orientate themselves particularly on headings or figures and corresponding captions. This is where an initial estimation of the complexity, the level of abstraction and the length or depth of the approach is made. The orientation phase is followed by the familiarization phase, where more attention is paid to the introduction and conclusion. When working with their text, participants then try to quickly understand what the important results of the paper are, or what the approach or presented methods lead to. Data such as year, keywords or the conference of the paper are also checked. Additionally, the text is often combed through for individual peculiarities. Among other things, it happens that the points mentioned in the abstract are filtered more precisely in order to quickly access the main part of the paper. This entire process represents the first general overview of the paper and only takes a few minutes. The aim is not to memorize details here, but to grasp the overall context of the paper. The initial paper is, however, examined in much more detail than all others.

Once a good first impression has been made, the procedure is divided into three types:

- In-depth understanding
- Reference search
- In the wild

Many participants prefer to understand the initial paper in depth first. The paper is read and processed once from top to bottom. Sometimes the title, abstract or introduction is read again. Here people like to make extensive notes, mark lines or the like. How and in what form notes are created varies greatly. In order to keep the interviews within one hour, this phase was usually interrupted after ten minutes and participants were advised to concentrate more on the

In-depth
understanding of
literature

further search. Occasionally, this phase was also canceled by less experienced participants themselves if they had the feeling that they were not really getting any further.

Literature reference
search

If the participant has previously obtained an overview of the paper, the reference search can then be used to explore the paper and its surroundings. To do this, the paper is skim over and sometimes partially marked. Of particular interest is the abstract, introduction, conclusion, section titles, images and image captions, as well as the main parts of the paper. Quotations or important passages as well as key terms play an important role here. Notes are also common, but keep concise. Via citations, titles or important key terms, participants then search for the literature environment. Search engines, conference-specific searches, author searches of the paper or the exploration of citation networks are used to find new literature. When new documents are found, they are usually combed through with a little less detail than the initial paper and the process of literature research starts from the beginning until a first collection of papers is created.

Research in the wild

However, some people also tend to get a very broad picture of the literature first before getting deeper into individual papers. Especially people who had little experience with the topic of the initial paper tended to do a broad search. Participants stated that in this way a broader view of the topic landscape could initially be obtained in order to not overlook important areas. After the overview of the initial paper, further literature is usually directly searched. Here, too, citations, keywords or the title can influence the search, but does not have to contain any specific details of the original paper. Occasionally, participants search articles or overviews of the general topics, which provide less specific broad overviews of the related topics. In this type of research, the paper collections grow considerably faster. On the other hand, papers also fall out of the collection just as quickly if they are then examined more closely.

In general, this type of approach offers the fastest growing bibliography. However, papers fall out of the collection just as quickly, whenever they are examined more closely in retrospect.

These three types of literature searches cannot be strictly separated from one another and can therefore overlap. In some cases participants started with the approach to understand the first paper in depth, while then the workflow turned into a less focused one by switching to reference search or search in the wild. In practice it is possible that all three types are represented evenly in a normal literature search without a time limit.

After papers have been found, they usually only remain open in browser tabs. If the paper has not yet been sorted out after a long period of time, it will be downloaded together with others and saved as a PDF file. Some also save the Bibtex, URLs of the page or the DOI of documents. The exception here is the process of deep understanding, in which a document is downloaded directly after an initial overview in order to then work better with it.

Saving PDFs varies a little more and has some special features, which are discussed in more detail in the next section. In general, however, PDF files are categorized in a general folder or in several sub-folders according to subject areas that are important for research, or managed with the help of reference management systems. The most frequently used tools that are used in the course are here: Google, Google Scholar, Google Docs, Zeterno, Mendeley, ACM Digital Library, Text Files, PDF Viewer, tablets and printers.

3.3.3 Observations

In the user study research task as well as in the subsequent survey, some peculiarities in various areas emerged, which we will examine in more detail below.

Pretty much every participant takes notes. However, the underlying form of the created notes varies greatly:

This can be done, for example, as note files in the form of simple text documents, Google Docks files, or notes in Bibtex. The simple notes of information in text files is a very popular method, as this is done quickly and very easily with little effort. The disadvantage, however, is that in addition to jumping back and forth in tabs, you also have

Notes in text documents

to switch to other windows, where people quickly lose their orientation and have to briefly familiarize themselves with paper again in order to differentiate them from one another. Furthermore, the creation of notes in separate files quickly leads to confusion, which was particularly evident in the further survey.

Notes with a tablet
and pen

Another very frequently used method is the procedure of using a tablet and marking passages with the pen, as well as taking handwritten notes on the side of documents or in a collective place. Synchronize repositories are also often used or the literature research is carried out entirely on the tablet. The advantage here is that notes can be made directly on the text, which promotes improved text understanding and makes it easier to comb through the text again. The disadvantage, however, is that the tablet setup is time-consuming, the synchronization of data often offers the potential for frustration and several papers are difficult to handle on the tablet at the same time, which is why it does not encounter help for the overall overview of a bibliography.

Notes in reference
management
systems

Sometimes notes are also attached to papers in reference management systems, accordingly notes can be assigned directly. However, this can lead to a poor overview, since papers are only displayed in a list, which narrows the field of view of the bibliography.

Notes in the form of
tables

One participant preferred to compare papers according to own categories in order to understand relationships better and to compare paper easier to another. For example, functions of individual methods presented in papers, conclusions of papers, or advantages of individual methods were categorized and compared. This enables a clear and systematical comparison of papers, but is initially more complex, as this is very topic-specific and laborious.

Notes in text marks

Comments on PDF files or in marked passages have also been used several times. To do this, a line is marked and a comment is added to this marking via an editor. Similar to working with the tablet, it is possible to quickly incorporate the text here, but it also does not support clarity about relationships between papers.

Printing out papers

Last but not least, 3 people also used printers for current or old literature research: While this effort was not made in the literature research task, people reported that printing out papers is often used for more intensive literature re-

search. It was mentioned that marking, comparing as well as the general overview are drastically improved as a result. Participants report here in particular that thereby it is possible to work more freely and more closer with literature. Furthermore the distance from screens improves the focus on paper itself. It was also described that when comparing papers with this procedure, they are often placed next to each other, together with notes, in order to see them at a glance, to compare them and to switch quickly between literature. Participants particularly emphasized that this approach is very convenient for comparing paper content. But there are also some problems that this approach proposes: On the one hand, printing requires time, money and physical space. On the other hand, this clearly separates the process of literature search from the process of the familiarization with literature. Once the papers have been printed out, it is more time-consuming to briefly look for other papers that have been cited, for example. As a result, it can happen that an expandable view of the bibliography is given up.

In general, the test persons stated that they chose exactly this type of note because, in addition to the points already mentioned, the form of the note is the easiest to use for the topic, this type of note has proven to be the most time-efficient, or participants do not know or have not tried any other form. This study also showed that more experienced people do not prefer an explicit form of notes, since more experienced participants were represented in almost all types. The way in which the notes are structured or the familiarization process is additionally not significantly dependent on the experience in literature research in this study. The decisive factor for this is that two more experienced people proceeded very chaotically, four more experienced people quickly had very routine working procedures, but almost all participants had phases in which various processes were combined at will and almost all participants had phases in which a clearly identifiable course of action was observable.

If we look at the important characteristics of paper in the familiarization phase, it becomes apparent that participants are particularly focused on the following characteristics: Title, published year, abstract, visual landmarks, like sec-

General information
on the note concepts

Important
characteristics of
scientific papers

tions and chapter headings or graphics / figures and associated captions, conclusion, introduction, authors, important quotes and references, citation count of Publications and last but not least the conference of a paper. The priority of the characteristics is sorted in descending order. Higher priority features were named or considered more often in the course of this study. Roughly, the priority here also reflects the average order in the procedure for becoming acquainted with a paper topic.

Citation networks for improved literature searches

Another special observation is the approach taken by a test person who used the ConnectedPapers tool on behalf of the study. The use of this tool turned out to be extremely time-efficient for the literature search, as further papers related to the topic were found after a few minutes. With the help of the paper title, a citation network was generated with the software. The research clusters then served as orientation points and abstracts of the papers were read here and there. This enabled the user to find papers that were of interest to him within a very short time.

Usage of reference management systems

Sticking to the usage of additional software, it is also noticeable that more experienced people here four out of six (around 67 percent) are more likely to use reference management systems than less experienced people three out of nine (around 33 percent). The reference management systems used by participants in the study are Zotero and Mendeley. However, some of the users who use these reference management systems criticize the poor overview of papers that is offered by using the programs. In addition, participants criticized that there are only a few limited options for comparing papers.

Categorizing literature to get a better overview

The problem that lists do not provide a sufficient overview also arises from the general organization of local PDF files on the PC. Many people mainly work with PDF files. When it comes to saving these, six out of 15 participants try not to simply save them as lists in a folder, but to integrate categories in the storage system. The further survey shows that especially with large collections of papers, documents are often sorted according to custom categories. Some participants create folders and save each paper in exactly one category. Others, on the other hand, assign documents to

all categories that fit the paper context and thus allowing redundant storage, which can make bibliographies appear significantly larger and probably more confusing than they really are. Again, others save papers which belongs to several categories in the parent folder, which is also a storage space for, for example, notes, saved urls, or the Bibtex. Each of these approaches has certain disadvantages in terms of clarity, which participants also refer to. Nevertheless, these problems getting encountered because these results are still clearer than normal lists. Another possible solution here are colored tags, which is a method that has not been encountered once, also reference management system offering methods for this. Participants replied that tags do not particularly improve the clearer kit, but in some cases even make it worse.

The urge for an increase of free, interactive organization of bibliographies is not only evident when storing files, but runs through the entire process of becoming acquainted with literature. Users are even ready to change their working environment and use a tablet just to edit paper by hand. Some participants even print them out and accept a lot of effort only to work with them in a more interactive way. When the people who print out the paper are questioned more closely, it is reported that handwritten evidence, for example, is easier to understand. Concepts can be drawn for better understanding and markings and notes can be made more specific. But handwritten work is by far not the only reason for this very extreme procedure, which could be encountered with tablets, for example. Participants also describe in particular that they like to spread out papers on their desks, by putting them next to each other and comparing them. In addition, printed documents are often sorted by placing them in different places on the desk.

This procedure can also be found in digital form. Three participants like to work with the desktop environment before the papers are finally saved. Here, the documents are then also sorted, renamed or categorized. The desktop is also used as a clipboard for other important files that are part of literature research. Participants report that this makes the desktop look a bit cluttered, but this approach contributes to a more pleasant way of working.

Analogous workflows

Using the Desktop as an more interactive method

Front page for direct
identification of
papers

Another great peculiarity that was observed especially during the literature research is that test persons often quickly lose track of the identity or the content of a paper, but restore this overview again in a very short time through the title page of a paper. The title page itself doesn't even need figures or special graphical elements. The layout seems to be more decisive, thus the structure and arrangement of individual elements or individual keywords. This observation becomes noticeable in the study through processes in which participants switch between papers. This can happen, for example, when many papers are open in tabs and a search for a specific paper is performed, but also for example when a list of papers is scanned through, so that papers get checked individually. The title alone was also not decisive, as some participants also named the PDF files after the title of the document when they were downloaded and still only remembered the context after opening the paper. With the help of the front page, even documents with no graphics could be recognized quickly using the page layout. Nevertheless, graphical elements have still an important role during the orientation, especially when they are read for the first time.

3.4 Interpretation & Consequences for the tool PaperOrganizer

Examples in which people tried to have an desktop environment when organizing a bibliography or examples where papers are printed out to be able to work better with them show that more opportunities for interaction in the organization of bibliographies is urgently needed.

The effort of printing out literature indicates an elaborate process, which reveals many potential strengths in interaction during the process of organizing documents. Furthermore, this action shows that some participants are willing to forego further searches or quick research options and accept the use of resources to be able to use the advantage of free interaction. This basic concept should therefore also be one of the central features of the PaperOrganizer: Free sort-

ing and individual arrangement of papers. The underlying concept is the spatial orientation of people. If people look for literature, there is an idea or a perspective on the topic: An immature mental model of the bibliography. But this model is nowhere visually consolidated, which makes organizing and working with bibliographies more cognitive demanding. The study shows that participants often need these mnemonic device in order to be able to better classify their mental model by means of spatial orientation, which serves to gain a better orientation and overview in the researched topic landscape. The process of organizing papers therefore not only serves to structure literature but rather to gain orientation in a researched topic landscape.

The publications by Newcombe [2010], for example, show that spatial thinking and orientation have a decisive influence on learning processes. Their importance is also examined in more detail by Mathewson [1999], who shows that spatial orientation is deeply anchored in cognitive processes and serves as mental supports. The PaperOrganizer should therefore offer a desktop-like environment in the form of an interactive map in which papers can be arranged individually, which enables better orientation and more individual familiarization.

The observations also highlight the problem that some participants have difficulties categorizing papers. Sometimes a topic does not fit any category or vice versa to a certain extent to several at the same time. An interactive map would already make it possible to cluster papers according to categories. The additional naming of these clusters can then clarify the affiliation to a category by means of the distance of the paper. Connections could also help clarifying complex relationships between literature or promote better examination. The underlying concept is that of a mind map: Developing an own super ordinate big pictures requires decisions that combine own concepts about the topic with information about the content. The additional intensive study of papers and the requirement that a more precise vision is needed before people trust themselves to assign papers in their bibliography, provides as a by-product exactly this deeper understanding of the own bibliography. At the same time, knowledge is graphically consolidated, which promotes the memorization of important key points

or features and improves clarity. The fact that mind-maps offering important learning concept, could therefore also promote the process of getting acquainted with a bibliography through a design concept, which goes in this direction.

The observations of the study further show that the most important features that play a major role between papers are authors' commonalities and citations. Citation networks in particular offer the potential to understand and decipher deeper bibliographies, of which Van Eck and Waltman [2014] also reports. Therefore it is important that the PaperOrganizer offers the possibility to support these structures automatically. If we consider features that are particularly relevant for the understanding and recognition of individual papers, it can be assumed that the identity of a paper is reflected by the title page, as this was also described in more detail in the observations, in Section 3.3.3. If a title page view is not possible, the title, year of publication, authors and abstract should be supported and displayed in another form.

In order to enable an intuitive design, the PaperOrganizer should also connect to already known views and combine them with new features. One of the most common views in literature research is the list display of papers. Integrating this already known view could help to enable faster and more intuitive familiarization with the tool.

The last big topic is the support of notes. The study has shown that notes are created in completely different ways and that different approaches also have different advantages, which, however, can also depend on the individual topic. Therefore, notes should not be restricted in any way, but specially promoted and expanded in order not to rob users freedom or to make routine workflows more difficult. A promising approach here is to open papers with the standard PDF viewer or editor, which offers a more customized and diverse work environment and supports special intuitive as well as individual features when working with PDF files. The additional offer of a note function in the PaperOrganizer promises the potential to enable information to be collected and stored centrally in one window, thus reducing the loss of data or unstructured processes in storage man-

agement, such as avoiding the creation of individual note files.

In general, the consequence of this is that the PaperOrganizer should enable an interactive, personalized work environment that supports important features from and between papers, enables diverse categorization and promotes intuitive familiarization with paper through the support of external programs.

Chapter 4

The Tool PaperOrganizer

4.1 Description of the Tool and Design Ideas

With the help of the study, we developed the PaperOrganizer tool, which allows the user to structure local paper collections with a variety of functions and interaction options in order to facilitate clarity as well as to become acquainted with the literature. The PaperOrganizer works as a desktop application and supports Linux, Mac and Windows operating systems.

Figure 4.1 shows the interface of the tool PaperOrganizer with all marked components. The user interface consists of the top menu bar, the header with important control buttons, the context bar on the left, a extendable and retractable document viewer and the work surface on which papers can be organized and viewed.

The structure of the work surface is similar to that of the desktop, but with a few significant differences. Paper or title elements with various functional options can be displayed here. Furthermore, it is possible to zoom and arrange the elements freely and interactively on the surface.

Work surface as the main component which contains paper and title elements

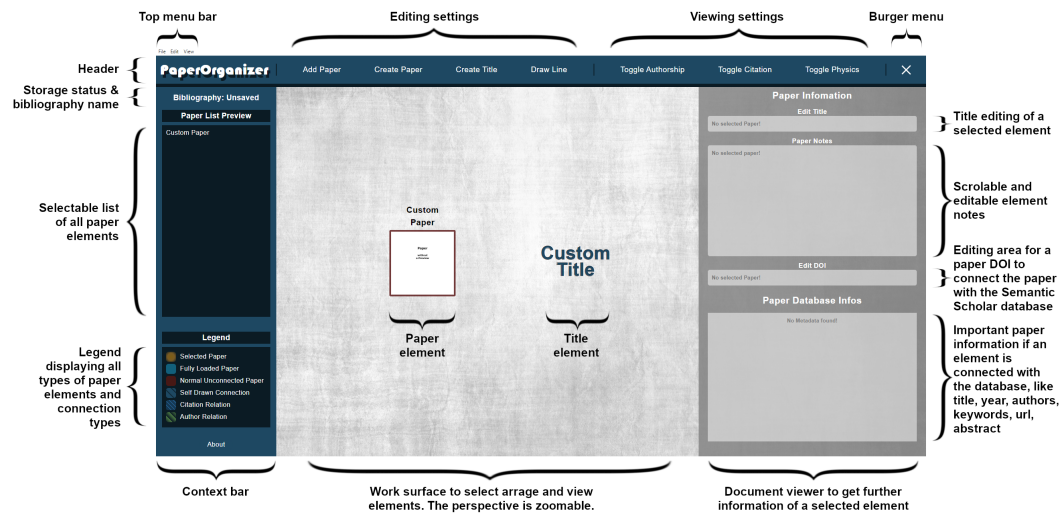


Figure 4.1: Overview of the interface of the tool PaperOrganizer

The zoom function can be controlled with the mouse wheel and has a minimum and maximum factor, so that the user is not able to choose a zoom level in which someone gets lost. The entire work surface can also be moved by holding and dragging it. The working area itself has no limits and thus allows any size of area for organizing the bibliography and its elements. To intuitively demonstrate that the user can arrange the elements, the mouse pointer changes to a hand when moving over an element and the element itself is highlighted. The hand is chosen in such a way that in addition to the association of moving the element, there is also the possibility of opening the content of the paper. Furthermore, the design should allude to the desktop surface, where the user is already used to moving elements and opening them. A simple click on the element marks it and colors it in a yellowish glow. Paper elements can be opened with a double-click if they contain a PDF reference. The title and paper of the elements differ significantly from one another in various characteristics.

Paper elements and their functionalities

A paper element can take on different states and forms, which is shown in Figure 4.2. In general a paper element consists of a rounded rectangle and a heading above it. The heading refers to the name of the represented paper and is of a size that is barely legible when it is fully zoomed out.

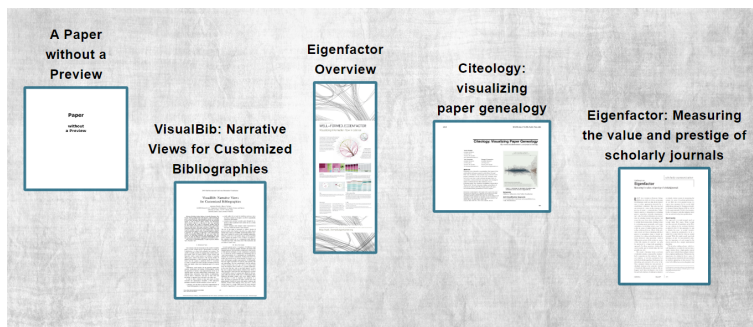


Figure 4.2: Overview of different paper formats partly with loaded title pages. All paper elements are calculated in the way that they have the same space.

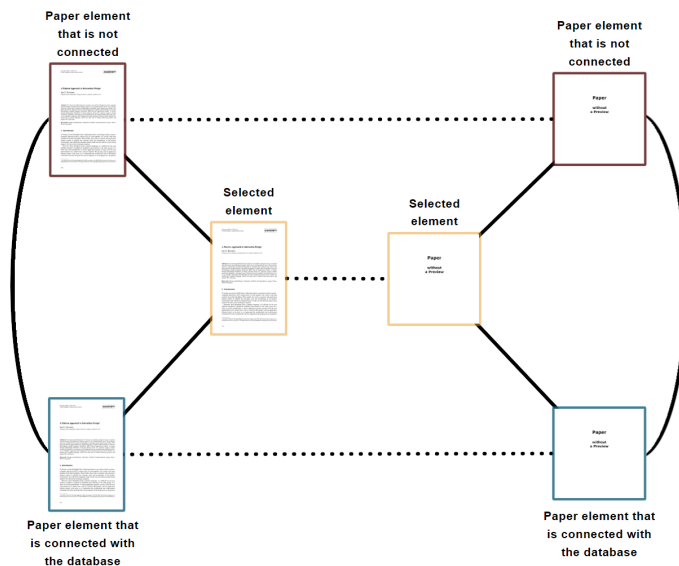


Figure 4.3: Concept of different paper element states and state relations.

The heading of the paper element, as well as the shape and color of the rectangle depends on various properties of the paper. If a paper is added to bibliography, the rectangle is filled in black with the heading *Loading* to indicate that the paper has not been completely processed. Once it is loaded, it changes to a new form, with a new containing image and a different named header. The

name in the metadata is initially selected as the heading of the element. The rectangle of the paper element is then rendered sharply with the first page of the PDF file so that the text on it is easy to read. When fully zoomed in, the rendered thumbnail image takes up the entire screen to be easily read. This view serves as a close-up view of individual papers in order to make even small graphics and texts recognizable.

The shape of the rectangle of the paper element adapts to the aspect ratio of the title page of the loaded PDF file. The clue is that all paper elements fill the same space. Therefore they have exactly the same total render area in on the working surface. With this adjustment, the user recognizes all papers equally, even though they support very different formats.

With the aid of the equations f_S and f_R , the dimensions of the width x and height y are determined by means of a constant area S , the aspect ratio R and the zoom factor z for each paper element. f_S represents the surface dependencies under respect of the zoom factor and f_R the aspect ratio dependencies. R is taken from the render process of the target title page.

$$f_S : S(z) = x(z) \cdot y(z) \quad \text{and} \quad f_R : R = \frac{y(z)}{x(z)}$$

It follows for $x(z)$ and $y(z)$:

$$x(z) = \sqrt{\frac{S(z)}{R}} \quad \text{and} \quad y(z) = \frac{S(z)}{\sqrt{\frac{S(z)}{R}}}$$

If a paper element is created without a PDF file, a standard image is loaded as the cover image, which indicates that the Paper does not provide a document preview.

But not only the form and headline can change. The paper elements also has two color states as an unselected element, which provide the viewer with information about whether metadata of a scientific paper is stored for the element. An overview of the states of elements is presented in Figure 4.3. Loaded paper elements can be connected to the open-access database of Semantic Scholar ¹ using the Digital Object Identifier (DOI) to enable further features when

¹<https://www.semanticscholar.org/>²

handling the paper. Red-brown loaded papers are normal papers that are not connected to the database. Bluish loaded papers, on the other hand, are fully connected and support all functions of the program. In addition, the title is overwritten with the title loaded from the database if it has not been changed by the user beforehand.

A title element on the work space consists of one to three rows to provide space for words. The anchor point of the title, compared to the center of a paper element, lies under the font, which is visible when it is connected to other elements. Paper and title elements however, behave in the same way with their surroundings.

In addition, elements that have been created using a PDF file can also be opened with a double-click, if it was created with a PDF file. The document is opened with the default program for PDF files of the operating system.

Above the work surface is the header, which is divided into four areas: The logo, important buttons for editing the elements in the work surface, important buttons for viewing the bibliography and the burger button to open and close the document viewer. The gestalt law common regions is used here to group the purpose of the buttons.

The *Add Paper* button opens the operating systems dialog window to search for PDF files. Thereby it is possible to add papers with a PDF file individually, which are then displayed as a paper element with cover picture in the work surface.

Create Paper creates a paper without a PDF reference, which also allows a database connection using the DOI of scientific papers.

Create Title creates a title element in the work surface.

The last button of the edit element area in the header is the *Draw Line* button. It can be activated and the program changes to the draw-line mode. In this state, connections between elements on the work surface can be drawn or deleted. The custom connections are represented as a solid blue line. Selecting two elements one after the other leads to a line between those elements. If a line already exists, it will be deleted.

Moving on to the buttons to view the bibliography, the option *Toggle Authorship* toggles author connections. If at least

Title elements for structuring the view

Operation of important functions in the header

one author has worked on two different papers, a green long-dashed line is drawn between all papers with this property. In general this view shows authors in common. However, only fully loaded papers with a DOI are linked. *Toggle Citations* works similarly to *Toggle Authorship* and toggles citation connections. Again only fully loaded papers are connected. The line is shown as a light blue short dashed line and creates a citation network when enabled. The *Toggle Physics* button puts the work surface in a physics-based animation state. As a result, all papers lose their firm contact with the surface and repel each other with a force. Two other forces regulate the animations of dragged connected papers, which, depending on their connection level, attract each other stronger or weaker. After moving an element, it is again given a fixed and force free position. If the physics mode is switched off, all papers are also attached to their new shifted positions.

Information view on
elements in the
document viewer

In order to organize papers more easily, to allow more working space, or to get a big picture of the bibliography, the document viewer can be extended and retracted. Important characteristics of papers are displayed in this viewer and element properties can be edited. The areas are designed here with the law of closure. However, to edit or view an element, it has to be first selected.

The paper title is displayed in the text field at the top and can be edited here. During editing, the title is live updated everywhere in the program to visualize the effects of the editing for the user.

The text field below allows adding notes to paper or title elements.

The following field offers editing the DOI of a paper. If a correct DOI is entered, all the metadata of the scientific paper with the DOI is loaded from the Semantic Scholar database. Since the database only allows 100 requests per minute, the DOI is checked locally first and only then sent a request to the server. Invalid entries are marked in red. DOIs that are not found by the database are highlighted in gray. The color difference is intended to prevent users from wrongly assuming that they making a mistake when entering the DOI, if the DOI cannot be found by the server.

The last box of the document viewer shows additional information for a fully loaded element and covers the most

important characteristics of papers: Title, year, authors, keywords, database, url and the abstract are presented here as a scroll-able list.

On the opposite of the document viewer is the left context bar, which provides an overview of the loaded bibliography.

On top is the custom name of the bibliography displayed. If the system state is not saved, it indicates this as well. Below is the scroll-able list of all paper elements, which enables the a different selection method of loaded paper elements. In this way, a paper can be found and selected easier, even if it cannot be found quickly on the work surface.

The legend at the bottom, separated by the Law of Closure, shows special color features in the work surface.

Context bar to navigate through the bibliography

The program also offers deleting selected papers by pressing the *Delete* key and also provides opening papers with *Enter*, which works similar to the double-click. Therefore, a selected paper has not to be searched again, if the user only wants to open it. The operation here is also based on the use of the desktop. It is also possible to use *Space* to deselect papers.

Handling techniques

The upper menu bar lists all button functions again, with some more specific ones added to better control the tool. In general, the menu consists of three tabs: *File*, *Edit* and *View*. The *File* tab offers the possibility to open, save and close a saved bibliography, with the usual naming and handling. The bibliography is saved as a JSON file. Closing a file saves the bibliography, to continue working on an empty work surface.

Upper menu bar to view all operations

The *Edit* and *View* tabs are divided according to the principle of natural mapping, analogous to the two areas of buttons that appear in the header. Here all button options are listed again. In addition, the tool offers the option *Draw Timeline* and *Clear* in the *View* tab.

Draw Timeline works in the same way as the *Draw Line* button, except that lines can be drawn thinner and in gray, in order to offer a second, less obtrusive line type when drawing lines. This is specially intended to offer the possibility of drawing overviews of the release years.

The *Clear* option in the *Edit* tab clears the entire work sur-

face by deleting all elements.

Handling techniques

But in addition to the standard functions, the PaperOrganizer offers also some methods for more comfortable system operation and to make working with the tool easier.

In order to load entire lists of documents files quickly, PDF files can be selected and dragged into the tool. Individual files or a saved bibliography can be loaded in the same way by using drag and drop. When loading a PDF file, the title page of the document is searched for a valid DOI and, if necessary, automatically linked to the database. This prevents time-consuming searching and entering of DOIs. If changes in the bibliography are not saved and the program is closed, the tool saves the view locally on the system and loads it the next time it is opened.

The layout of the tool is designed in such a way that there is as much view of the work surface as possible to make working with papers easier, but still enough space for important operating views such as list, buttons, or notes.

4.2 Important features & their benefits

The software is written with Typescript, CSS and HTML and is based on Node³ and Electron⁵ to enable multi support for operating systems. The tool serves as a desktop application to adopt personal methods during organizing the literature. By that, it is possible to open a PDF file using a default tool of the operating system.

The PaperOrganizer tool is the first known tool to us, which enables free interaction and arrangement of locally stored documents and thus offers a new perspective at bibliographies.

The tool makes it possible to visualize entire lists of PDF files in a few seconds and provides a playground to enable individual analysis and familiarization with literature. Once papers have been loaded, they can be categorized with distances, which is more appropriate to topics in sci-

³<https://nodejs.org/>⁴

⁵<https://www.electronjs.org/>⁶

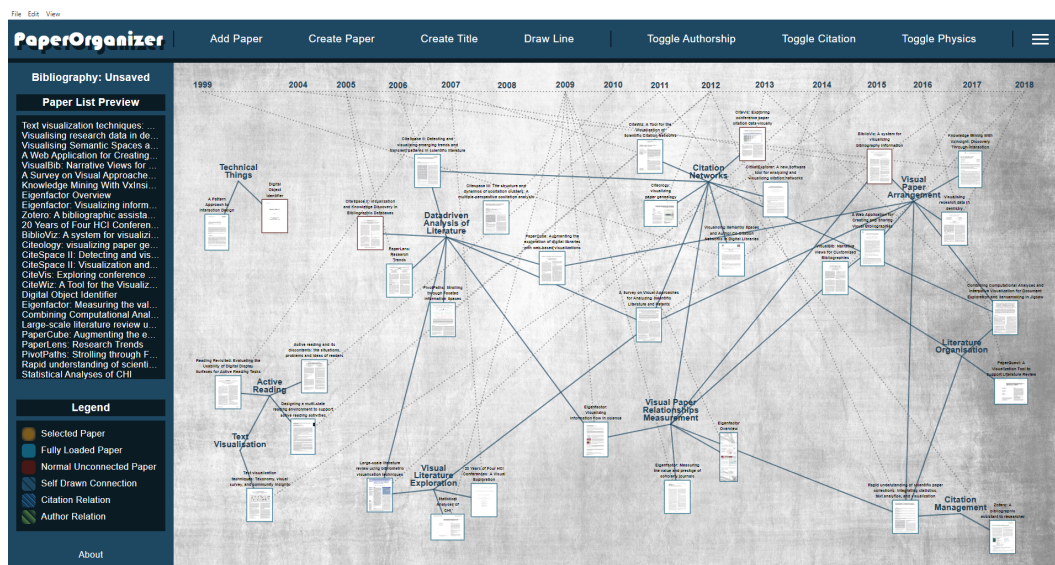


Figure 4.4: Interface of the tool PaperOrganizer, with a loaded bibliography. It provides a customized overview of the related work showing the nine clusters of research areas.

entific literature than assigning them to a specific topic. The unpleasant side effects of saving documents in lists discovered in the user study can also be avoided in this way. Figure 4.4 shows a custom overview of the related work as a loaded bibliography.

In order to facilitate intuitive handling, the list view offers a quick overview of all papers, which could prove to be a practical view if users search for a paper by the name. Also, since most of the participants in the study only worked with list views, it might be possible, that users find their way around the bibliography more quickly. A close up view of a customized cluster with the list view and document viewer is shown in Figure 4.5.

But in addition of providing the method for personal structuring of literature depending on the needs of users, the PaperOrganizer also offers a visualization for the arrangement of the citation network or the author network, which is shown in Figure 4.6

Chen [2006], Matejka et al. [2012], Van Eck and Waltman [2014], Dattolo and Corbatta [2018], all report different advantages in the clarity of the use of author and citation

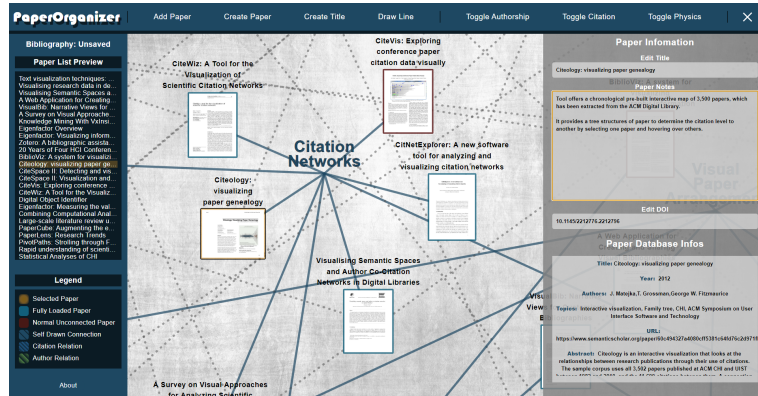


Figure 4.5: Close up view of a custom research cluster in the bibliography with a selected paper element and some information and notes about the paper in the document viewer.

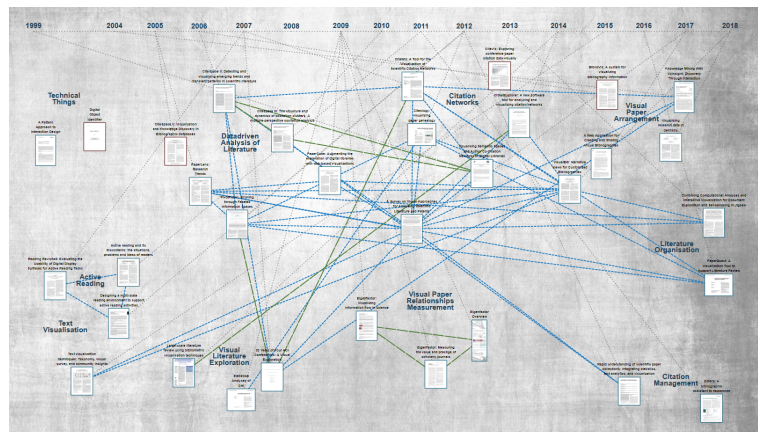


Figure 4.6: Interface of the tool PaperOrganizer, with a visualized citation network (light blue) and the author network (green).

networks, which was also partly mentioned in the related work in chapter 2. With the help of the PaperOrganizer tool, this view is retained even after the literature search. One of the key benefits of using these visualization is the improvement on cognitive and perceptive understanding in quantitative and verbal information [Mohamad et al., 2018].

Furthermore, the custom connections of papers can provide clarity and possibly increase the handling of the bibliogra-

phy. In combination with the physics mode, network focal points could be identified in a short time in a local stored paper collections, which offers the opportunity to clearly structure entire networks within a few clicks.

Zooming thereby supports different views of the bibliography, regardless of whether an overall picture, sub topic areas or just each paper.

Creating notes is another big challenge the tool faces, which was described in more detail in Section 3.4. The idea behind the implementation here is to support all forms of notes as well as to offer further methods for adding notes in the PaperOrganizer. By opening a paper with the standard PDF viewer or editor, it is possible to create notes in PDF files in a way that is the most common one for the user. In addition, notes can be written for each individual paper element in the tool. However, many participants of our study also prefer to collect notes in one place, which is achievable by creating notes in title elements. One possibility is, a main title in the middle of the bibliography, which could be used to collect notes from all papers in one place. As a side effect, this leads also to an opportunity to organize notes hierarchically, by creating categories with a title and associated paper elements. Users can add shared notes to the title element of all related papers in the category and therefore also categories notes.

The use of custom paper elements and custom lines offers users further potential, which enables working in a more creatively and unrestricted way with the bibliography in the PaperOrganizer. How exactly users create an overview of their own collection to clarify a paper context is, to a certain extent, up to themselves.

If we take a closer look at the layout of the paper element, it was designed in such a way that it shows the most important features of a scientific paper that identifies it. In this way, elements can be differentiated from other papers at a glance and probably be found faster. The title page of a paper is displayed as a preview of a paper element, which should ensure that it can be identified directly. Accordingly, it is possible to arrange and view papers as if they were printed out in front of a desk, which is sometimes done by participants in our study, as shown in Section 3.3.3. Even if the front page often covers a lot of important information, this view can be a bit overwhelming and individual infor-

mation cannot be found directly. If a paper is connected to the database, additional information is shown in the document viewer, which is even displayed if only one DOI is known, while working with a custom paper element.

The storage of the bibliography also supports sharing the personal view of literature. Dattolo and Corbatta [2018] report advantages of being able to exchange personal bibliographies and share them with others, so that teams can get uniform views of a subject area more quickly.

The study by Tashman and Edwards [2011] reports problems about active reading and some tensions that readers made. Participants in this study reported that current software do not offer enough space to support real world multi-document tasks—like comparing and integrating information between texts. At the same time, readers need efficient and flexible setups. This study also showed that switching among multiple document windows on a small screen was found to be an important source of frustration in computer-mediated active reading.

The PaperOrganizer has the potential to face some of these tasks. The view in the tool prevents switching back and forth between too many documents. Not only is it possible that the tool could improve the work with bibliographies and make organizing more efficient, but it is also conceivable that, papers could be understood more quickly, due to an easier classification of the context. Further studies could find more precisely to what extent the PaperOrganizer helps here and how far the tool fulfills its aims.

4.3 Limitations

One of the very fundamental limitations that the tool has is that of course it does not cover all areas of literature research. The PaperOrganizer tool only focuses on structuring and managing locally stored paper collections in the process of literature research. We therefore recommend using the PaperOrganizer together with a reference management system and tools for searching literature.

Another limitation results from the user study. Since only participants in the field of computer science were tested, it is difficult to assess to what extent the PaperOrganizer can

also offer help outside the domain. In addition, none of the participants is involved in more than five publications, which is why the study does not necessarily cover all areas of experienced groups of people.

The tool also has some technical limitations that can be found in paper notes. The provided note function supports normal editing of text, including copying or pasting. However, the text field does not offer any further more complex editing options, such as centering text, tables, highlighting, and so on. Our user study shows, however, that many participants do not attach great importance to this, precisely because many people use simple text documents or the Bibtex as note files. Participants who create a more specific form of notes, such as working with tables, tend to be the exception. See Section 3.3.3.

Another technical limitation is the limit of 100 fully loaded papers, which is caused by the open access database of Semantic Scholar ⁷. Further implementation steps could circumvent the problem, for example by integrating other database systems, or methods to ensure that papers remain unloaded until they are allowed to send requests to the server. However, the study also showed that mostly only ten to 20 papers are used.

Since the software still serves as a prototype, it is also recommended to always create a backup of the stored bibliography. Furthermore, the tool works with file paths and only saves the address of documents when a PDF file is loaded, which means that if a PDF file is moved to an other place in the operating system, the tool can no longer recognize the file path and thereby the associated paper element is not longer fully loaded.

In the area of intuitive handling, it might also be the case that drawing or deleting lines is not immediately understood by everyone. Selecting two papers one after the other was a pleasant step to implement, but could have disadvantages in handling.

The last point is related to the citations networks, which have no orientation. The graph is undirected, which can provide a better overview for self-organization, but leads to a loss of information in the case of citations.

Further studies might also show some more limitations, such as unfavorable color schemes or missing options that

⁷<https://www.semanticscholar.org/>⁸

could further help the system operation.

Chapter 5

Summary and future work

5.1 Summary and contributions

Our study allowed an in-depth look into the workflow of literature research. It could be shown how people search for literature collections and work with them.

Most of them first familiarize themselves with a start paper and then switch to one of the three approaches: *Deep understanding*, *Reference Search*, and *In the wild*. *Deep understanding* describes the phase in which people deal deeply with only isolated literature. It is mainly used to search for very specific questions / functions / evidence / methods in the form of literature. *Reference search* describes the search for thematically relevant papers mostly via references or authors, where the incorporation into papers is a bit more superficial. *Search in the wild* is an extremely broad and rough search for literature, mostly in order to familiarize oneself with a new subject area in which papers are only followed up on the most important features.

The results of the study showed that participants strive for more interactive approaches in organizing literature and encountered various difficulties during structuring them. In particular, own categorization of larger collections in lists can lead to confusion, which should be avoided. The

study also offered various peculiarities and priorities that have significantly influenced the PaperOrganizer tool to allow intuitive, versatile organization.

The PaperOrganizer is the first tool known to us that provides an opportunity to interactively structure a bibliography. In addition, the tool offers various functions that are important when organizing literature. This includes, finding and opening papers quickly, as an example. Furthermore, with the help of custom titles and links it is also possible to sort, categorize and compare literature interactively. But also citations networks or author connections are featured in this program, in order to get a better overview and a quicker understanding of the relations between papers. With all those functions the PaperOrganizer is capable to structure entire lists of PDF files and visualizes them in a very short time.

5.2 Future work

For the next steps, an evaluation would be useful to better assess the potential of the tool and to draw conclusions about how far the tool is achieving its goals. It is also possible to find some avoidable limitations or unintuitive operations of the tool, which could help to reduce restrictions.

If the tool turns out to be really helpful in organizing literature, individual functions could be improved or expanded in further steps. For example, when connecting a paper element with the database, the tool only accesses the metadata of scientific papers with help of their DOIs. However, sometimes it is not mandatory that every paper of the database has a DOI, which provides an opportunity for improvement.

It is also conceivable to combine different functions of the entire literature research process. That could be achieved by combining some aspects of reference management with the provided methods of the PaperOrganizer.

If the evaluation shows that the tool still has difficulties in operating, more data would be needed on how users proceed with the tool, to get a more detailed insight into the problems and to correct them. In addition, it would also be interesting to see how exactly people create overviews

with the help of the tool in order to promote these methods. Especially the process of organizing scientific literature of users could play an important role, to better combine the process of research with provided functionalities of the tool and could help to find far-reaching optimization strategies in the future.

Further work could, however, also go into the area of discovering and exploring new perspectives of research areas, which encourages opportunities to find new, yet unexplored areas of research. Overall, the tool offers promising functions and a range of interactive options, which could have a strong impact on the work and the future of literature research.

Appendix A

User study consent form

Informed Consent Form

Interview about literature research as part of the bachelor thesis

Principal investigator: Markus Horlemann
RWTH Aachen University
Email: markus.horlemann@rwth-aachen.de

Purpose: The aim of the study is to obtain a precise insight into literature research to make research easier for users.

Procedure: As a participant you are anonymized. Audio and screen broadcast will be recorded in the interview. First a paper is presented. This should then be used to search for two more papers. Then one sentence should be written about each paper with a direct quotation. If possible, you should proceed as in a normal literature research. You will then be asked to talk about current own research. The process can involve thinking out loud what to do and questions can be asked at any time.

Risks: The literature search should not take too long with an approximate maximum of one hour, but it can still lead to exhaustion or tiredness. You can take a break at any time.

Confidentiality: All information is treated confidentially and only evaluated in the context of the bachelor thesis to analyze processes in the literature research. No personal data will be published

Costs and Compensations: There are no financial costs for participating. Unfortunately, a reward is currently not possible due to the current situation.

I have read and understood the information on this form.

I have had the information on this form explained to me.

Participant's Name

Participant's Signature

Date

Principal Investigator

Date

Figure A.1: Consent Form for the user study

Appendix B

User study sheet

User Interview: Literature Research

Markus Horlemann

Introduction: As part of my bachelor thesis, it is about literature research and its organization. I would like to give you a few tasks and look over your shoulder while researching literature.

General:

What is your gender? _____

How old are you? _____

How often do you roughly conduct scientific literature research? [Annually, each six months, monthly, several times a month, weekly, daily / constant, other] _____

How many scientific literature searches have you already carried out?
[0, 1-5, 5-10, 10-20, more than 20] _____

What was the main reason for your literature research? [Publications, Software development, theses, seminars, other] _____

How many scientific papers have you already published? [0, 1-5, 5-10, more than 10] _____

When doing a literature research, how many papers do you work with on average? _____

Workflow:

User Task:

Find 1-2 other matching papers to the topic *Textile Interfaces*. Write a short text (3 sentences) in which you create an overview of the papers. Cite each paper once. Please proceed as you would in a normal literature research, in which you would work with a normal number of papers.

Questions

How would you proceed in a normal literature research, in order to better familiarize yourself with the topic?

How do you approach the research task?

How do you keep track of all papers?

Figure B.1: User study sheet with general questions and the research task

What do you pay attention to when you quote papers?

Others

Specific Questions

WHY is that? (Chop up for details)

Tools: Which ones, how do you work with it? What is important to you about the tools?

Further literature research: What do you use to find more?

Imagine that there would now be significantly more papers. How do you get an overview?

Which properties of paper are important to you? Specially to get a quick overview of a topic and how it relates to other papers?

Have you ever compared papers with each other? What is important to you there?

Figure B.2: User study sheet with the second part of the research task

Further Survey:

Would you like to show me an old or current literature research of yours?

How did you work on the topic? Where did you start?

How did you sort / save the files at the beginning?

How did you proceed then?

Which characteristics were important to you in the papers?

Would you say that some papers are significantly more important than others?

If so, how did you recognize that, especially at the beginning?

What did you do to quickly or better understand the papers?

Additional stuff

Figure B.3: User study sheet with further questioning

Important to watch carefully:

- Where do processes take place repeatedly?
- What do many do the same?
- How is literature structured?
- What is organized, where are files stored?
- In which places do users feel constricted?
- At which points do you need more interaction / clarity?
- What do people pay particular attention to?

What is important to find out / goals of the study (relation to the software)

- (1) Where do participants lose a lot of time while working with literature? (Organize efficiency)
 - (2) What general system, method or principle do most of the participants use? (Important / practical processes)
 - (3) Are there more specific procedures that many participants handle the same, or are there technical processes that show similarities? (Intuitive design)
 - (4) What are important methods or features when it comes to organizing literature? (Feature priorities / important processes)
 - (5) Where is interaction missing? (Practical processes)
 - (6) How do users get overviews of their subject? (Important processes)
- => Finding out about advantages / importance / intuitive processes / practical help

Figure B.4: User study sheet attention points

Bibliography

K Mueen Ahmed and Bandar Al Dhubaib. Zotero: A bibliographic assistant to researcher. *Journal of Pharmacology and Pharmacotherapeutics*, 2(4):303, 2011.

Peter Bergström and Darren C Atkinson. Augmenting the exploration of digital libraries with web-based visualizations. In *2009 fourth international conference on digital information management*, pages 1–7. IEEE, 2009.

Jan Borchers. *A Pattern Approach to Interaction Design*. John Wiley & Sons, Ltd, 2001.

Chaomei Chen. Visualising semantic spaces and author cocitation networks in digital libraries. *Information processing & management*, 35(3):401–420, 1999.

Chaomei Chen. Citespace ii: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for information Science and Technology*, 57(3):359–377, 2006.

Chaomei Chen, Fidelia Ibekwe-SanJuan, and Jianhua Hou. The structure and dynamics of cocitation clusters: A multiple-perspective cocitation analysis. *Journal of the American Society for information Science and Technology*, 61(7):1386–1409, 2010.

Antonina Dattolo and Marco Corbatto. Visualbib: narrative views for customized bibliographies. In *2018 22nd International Conference Information Visualisation (IV)*, pages 133–138. IEEE, 2018.

George S Davidson, Bruce Hendrickson, David K Johnson, Charles E Meyers, and Brian N Wylie. Knowledge min-

- ing with vxinsight: Discovery through interaction. *Journal of Intelligent Information Systems*, 11(3):259–285, 1998.
- Marian Dörk, Nathalie Henry Riche, Gonzalo Ramos, and Susan Dumais. Pivotpaths: Strolling through faceted information spaces. *IEEE transactions on visualization and computer graphics*, 18(12):2709–2718, 2012.
- Cody Dunne, Ben Shneiderman, Robert Gove, Judith Klavans, and Bonnie Dorr. Rapid understanding of scientific paper collections: Integrating statistics, text analytics, and visualization. *Journal of the American Society for Information Science and Technology*, 63(12):2351–2369, 2012.
- Niklas Elmqvist and Philippas Tsigas. Citewiz: a tool for the visualization of scientific citation networks. *Information Visualization*, 6(3):215–232, 2007.
- Paolo Federico, Florian Heimerl, Steffen Koch, and Silvia Miksch. A survey on visual approaches for analyzing scientific literature and patents. *IEEE transactions on visualization and computer graphics*, 23(9):2179–2198, 2016.
- Carsten Görg, Zhicheng Liu, Jaeyeon Kihm, Jaegul Choo, Haesun Park, and John Stasko. Combining computational analyses and interactive visualization for document exploration and sensemaking in jigsaw. *IEEE transactions on Visualization and Computer Graphics*, 19(10):1646–1663, 2012.
- Nathalie Henry, Howard Goodell, Niklas Elmqvist, and Jean-Daniel Fekete. 20 years of four hci conferences: A visual exploration. *International Journal of Human-Computer Interaction*, 23(3):239–285, 2007.
- InfoVis. *InfoVis 2004 contest: Dataset and tasks*, 2004.
- Arif Jinha. Article 50 million: An estimate of the number of scholarly articles in existence. *Learned Publishing*, 23:258–263, 07 2010. doi: 10.1087/20100308.
- Joseph ‘Jofish’ Kaye. Some statistical analyses of chi. In *CHI’09 Extended Abstracts on Human Factors in Computing Systems*, pages 2585–2594. Association for Computing Machinery, 2009.

- Kostiantyn Kucher and Andreas Kerren. Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific visualization symposium (PacificVis)*, pages 117–121. IEEE, 2015.
- Bongshin Lee, Mary Czerwinski, George Robertson, and Benjamin B Bederson. Understanding research trends in conferences using paperlens. In *CHI'05 extended abstracts on Human factors in computing systems*, pages 1969–1972, 2005.
- Justin Matejka, Tovi Grossman, and George Fitzmaurice. Citeology: visualizing paper genealogy. In *CHI'12 extended abstracts on human factors in computing systems*, pages 181–190. Association for Computing Machinery, 2012.
- James H Mathewson. Visual-spatial thinking: An aspect of science overlooked by educators. *Science education*, 83(1): 33–54, 1999.
- Sara Mlakar and Michael Haller. Design investigation of embroidered interactive elements on non-wearable textile interfaces. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–10, 2020.
- Ahmad Nadzri Mohamad, Noor Nazahiah Bakri, and Mas Suryalis Ahmad. Visualising research data in dentistry. In *2018 4th International Conference on Information Management (ICIM)*, pages 35–40. IEEE, 2018.
- Nora S Newcombe. Picture this: Increasing math and science learning by improving spatial thinking. *American Educator*, 34(2):29, 2010.
- SP Rodrigues, NJ Van Eck, L Waltman, and FW Jansen. Mapping patient safety: a large-scale literature review using bibliometric visualisation techniques. *BMJ open*, 4(3), 2014.
- Zeqian Shen, Michael Ogawa, Soon Tee Teoh, and Kwan-Liu Ma. Biblioviz: a system for visualizing bibliography information. In *Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation-Volume 60*, pages 93–102. Citeseer, 2006.

John Stasko, Jaegul Choo, Yi Han, Mengdie Hu, Hannah Pileggi, Ramik Sadana, and Charles D Stolper. Citevis: Exploring conference paper citation data visually. *Posters of IEEE InfoVis*, 2, 2013.

Moritz Stefaner, Martin Rosvall, and Carl Bergstrom. well-formed. eigenfactor: visualizing information flow in science. In *SIGGRAPH 2009: Talks*, pages 1–1. College & research libraries news, 2009.

Marie B Synnестvedt, Chaomei Chen, and John H Holmes. Citespace ii: visualization and knowledge discovery in bibliographic databases. In *AMIA Annual Symposium Proceedings*, volume 2005, page 724. American Medical Informatics Association, 2005.

Craig S Tashman and W Keith Edwards. Active reading and its discontents: the situations, problems and ideas of readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2927–2936, 2011.

Nees Jan Van Eck and Ludo Waltman. Citnetexplorer: A new software tool for analyzing and visualizing citation networks. *Journal of informetrics*, 8(4):802–823, 2014.

Mark Ware and Michael Mabe. The stm report: An overview of scientific and scholarly journal publishing. 2015.

