User Centred Design Approach To Visualise Opportunities For Collaboration In Large Organisations

Master's Thesis at the
Media Computing Group
Prof. Dr. Jan Borchers
Computer Science Department
RWTH Aachen University

by
Mohammad Amin Yazdi

Thesis advisor:
Prof. Dr. Jan Borchers

Second Examiner:
Prof. Dr. Martina Ziefle

Registration date: 11.09.2014
Submission date: 10.03.2015

I hereby declare that I have created this work completely on my own and used no other sources or tools than the ones listed, and that I have marked any citations accordingly.

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie Zitate kenntlich gemacht habe.

Aachen, 10th March 2015
Mohammad Amin Yazdi

Contents

List of Figures

List of Tables

Abstract

While innovation in former days may have been mainly single efforts of researchers and inventors, today innovation often stems from collaborating of different disciplines, methods and approaches with both, existing and novel partners. At RWTH Aachen University Excellence Cluster "Integrative Production Technology for High-Wage Countries" [Brecher et al., 2012] researchers from 40 institutes collaborate on extending both practice and theory on production by addressing the polylemma of production. In order to ensure collaboration in cluster, so called cross-sectional processes (CSP) are established. These research groups address sustainability from a theory, human resources and technology transfer point of view. As one measure from the CSPs [Jooß et al., 2012] a social portal is developed for collaboration support and to integrate communication in a single source of truth.

In this thesis, we present qualitative research on a design study. We investigated a collaboration suggestion tool for social portal as a means of steering a research cluster. We evaluated the tool using a task-based analysis in regard to suitability for finding possible collaborators from both a researchers point of view and from the perspective of the COO of the research cluster. In total 15 participants were involved in a participatory design study. We then present the prototype and both qualitative and quantitative evaluations of prototypes. Three dimensions of validation were evaluated and highly ranked: *Discovery of new knowledge, Knowledge confirmation* and *Problem solving*. Overall, our visualisation was able to inform researchers about valid collaboration opportunities while at the same time effectively conveying organisational information. The prototype was evaluated using the System-Usability-Scale [Brooke, 1996] (84.5) and the Net-Promoter-Score [Reichheld, 2003] (80%) and received high ratings.

Überblick

Früher mag Innovation vor allem den Mühen einzelner Forscher(innen) zu verdanken gewesen sein. Heute hingegen bedeutet Innovation häufig die Kooperation verschiedener Disziplinen und Methoden und die Einbindung sowohl bestehender, als auch neuer Partner. Im Exzellenzcluster der RWTH Aachen zum Thema "integrative Produktionstechnologien für Hochlohnländer" [Brecher et al., 2012] arbeiten Wissenschaftler(innen) aus 40 Instituten mit dem Ziel zusammen, sowohl Praxis als auch Theorie zur Produktion zu verbessern, indem sie sich mit dem Polylemma der Produktion auseinandersetzen. Um die Kollaboration innerhalb des Clusters zu garantieren, wurden sogenannte Cross-Sektionale Prozesse etabliert. Diese Forschungsgruppen betrachten Nachhaltigkeit sowohl von einem theoretischen Standpunkt, als auch dem der Human Ressources und aus dem Blickwinkel des Technologietransfers. Zur Unterstützung der Zusammenarbeit und um die Kommunikation in eine "*single source of truth*" zu integrieren, wurde durch die CSPs eine gemeinsame Online-Plattform entwickelt [Jooß et al., 2012].
In dieser Arbeit, präsentieren wir die qualitative Untersuchung einer Designstudie. Wir untersuchen ein Tool, welches innerhalb eines sozialen Netzwerkes mögliche Kooperationspartner(innen) vorschlägt und ebenfalls dazu dient ein Forschungscluster zu steuern. Wir haben das Tool durch eine aufgabenbasierte Analyse im Hinblick auf seine Eignung bewertet, mögliche Kooperationspartner(innen) zu finden, wobei wir sowohl den Standpunkt eines Forschers, als auch die Perspektive des Geschäftsführers (COO) des Forschungsclusters mit einbezogen haben. Insgesamt 15 Personen haben an einer partizipatorischen Designstudie teilgenommen. Im Folgenden präsentieren wir den (endgültigen) Prototypen und sowohl die qualitativen als auch die quantitativen Ergebnisse der Bewertungen der Prototypen. Dabei wurde der Fokus auf drei Kernpunkte gelegt: *der Gewinn neuen Wissens, die Bestätigung von Wissen* und *auf die Problemlösung*. Unsere Visualisierung war grundsätzlich dazu geeignet, Forscher(innen) über aktuelle Möglichkeiten der Zusammenarbeit zu informieren, während gleichzeitig effektiv Informationen über Organisationsstruktur übermittelt wurden. Der Prototyp wurde aufgrund der System Usability Scale [Brooke, 1996] (84,5) und des Net Promoter Scores [Reichheld, 2003] (80%) beurteilt und durchweg positiv bewertet.

Acknowledgements

This work is dedicated to my parents, Dr. Mohammad Yazdi and Mrs. Farah Azima for their endless support and love.

I would like to express my special appreciation and thanks to my advisor Dr. André Calero Valdez, for without his guidance this work would not have been possible. I am also immensely grateful to Mr. Leonhard Lichtschlag, M. Sc., for his advices during the course of this thesis.

I would also like to thank my thesis advisor, Prof. Dr. Jan Borchers, and Prof. Dr. Martina Ziefle, my second examiner. I would like to thank wonderful people at the HCIC and Media Computing Group, for giving me valuable feedback, ideas, criticism and for readily participating in my user studies.

Furthermore, I would like to thank Susanne for her support morally and mentally.

Conventions

Throughout this thesis we use the following conventions.

The whole thesis is written in British English. Throughout the thesis the singular pronoun "they" is used to refer to both genders. In some cases the gender-specific personal pronoun is used, when a specific person of a specific gender is referred to or in direct speech citations. No gender bias is intended by this

This project was also submitted to INTERACT 2015 conference [Yazdi et al., 2015] and it is also partially used in this report.

Chapter 1

Introduction

Interdisciplinary collaboration has been considered both a boon and a bane of scientific advancement in recent years. Funding organisations like the NSF have shifted capacities to interdisciplinary research efforts [Jacobs and Frickel, 2009]. Interdisciplinary research is considered to be an effective solution for large-scale complex problems overarching the limits of disciplinary boundaries. In spite of its promise, interdisciplinary teams face several challenges in their collaboration [Repko, 2011]. Differences between disciplinary cultures (e.g., language, methodology, scientific performance evaluation) and individuals, in combination with shorter project runtimes, inhibit effective collaboration, which requires a mutual understanding of the topics and the team itself [Marzano et al., 2006]. The more experienced researchers are in interdisciplinary research, the more successfully they collaborate [Cummings and Kiesler, 2008]. However, the problem with having many scientists from different disciplines lies in how to bring them together to collaborate in an interdisciplinary environment. It is a crucial challenge to lead scientists away from their traditional means of research [Miller et al., 2014].

Interdisciplinary researches are very effective but comes with great challenges.

In a large-scale research-cluster (Cluster of Excellence "Integrative Production Technologies for High Wage Countries"at the RWTH Aachen University in Germany) a cybernetic management approach is applied, and in order to address the staff volatility and sheer size of a research cluster a "Scientific Cooperation Portal" (SCP) is created [Cummings and Kiesler, 2008]. The SCP is a web-based social portal that serves as a means to centralise communication, exchange files, list member profiles, offer interdisciplinary collaboration support and output tracking of the individual researchers. One part of the SCP is tracking of publications generated in the cluster to enable steering. When analysing co-author relationships for the reason of their successful collaboration, two types of relationships are dominant. Successful researchers are either similar ("birds of a feather flock together") in their co-authorship network, or similar in their publication output or complementary ("opposites attract") [Kretschmer, 1999].

SCP bridge the gap between researchers from different disciplines.

The first goal of this thesis is a design study of an ideal visualisation for interdisciplinary tasks. A visualisation that can resemble a mental model of users in the organisational structure. Secondly, investigate the metrics that our users trust in a visualisation. Answers to the following questions have assisted us to develop a supreme design. What kind of visualisation can help users solve problems and discover new information? How can a visualisation actually get scientists together? How can one identify possible collaborators using visualisations? Can a visualisation inspire available opportunities? Can a visualisation motivate users to produce more publications?

To achieve these goals, we undertook a literature review and utilised user studies to analysis the prototypes that we created. The literature review gave us with insights into previous research results in the field and guided us in the direction of our goals. The analysis showed that visualisation of publications can assist researchers as well as cluster administration to assess the interdisciplinary collaboration [Calero Valdez et al., 2012].

We analysed our prototypes using qualitative and quantitative studies.

We evaluated our design using a user-centred design and participatory design [Johnson, 1990]. Initially through interviews, we identified the context of use and user requirements. Later, we created paper prototypes and then software prototypes that covered all our fundamental tasks and user needs. Throughout the user studies, slight improvements were integrated between trials to incorporate user feedback. Our user studies provided us with enough evidence that not only are the initial requirements of users fulfilled, but in addition our visualisation can be used for other applications other than the main goals. Finally, we were able to confirm that our presented approach can enhance interdisciplinary collaboration in the Cluster of Excellence.

1.1 Chapter Overview

Chapter 1: First chapter discuss our motivation and overview of our work.

Chapter 2: In this chapter we discuss the infrastructure of interdisciplinary collaboration. We also offer an overview of papers which researched visualising interdisciplinary publications.

Chapter 3: In this chapter we distinguish different related works on information visualisation research projects and introduce their approaches that have improved visualisations of data. Throughout this chapter we explore different categorised information on visualisation research.

Chapter 4: Using literature reviews we derived five research questions and discussed the methodologies used throughout the whole design process. These methodologies assisted us in the evaluation and development of our prototypes with respect to our target user domains.

Chapter 5: In chapter 5 we discuss starting with user requirements and following that the use of them to develop our low fidelity prototypes, and finally, the high fidelity prototypes. We explain our prototypes and discuss the results of two user studies.

Chapter 6: In this chapter we summarise our findings and discuss our contribution to the field. We then explain possible future works and the limitations that we faced during this project.

Chapter 2

Interdisciplinary Collaboration

In this chapter we discuss the infrastructure of interdisciplinary collaboration and their challenges. We also offer an overview of papers which researched visualising interdisciplinary publications.

2.1 Infrastructure

Interdisciplinary research collaboration is a combination and integration of different data perspectives. People in such collaboration settings should gather from multiple disciplines with diverse individual experiences and different cognitive research models. Interdisciplinary teams address research questions that do not fit a specific domain of knowledge, hence the need to bring skills from other disciplines that lie beyond our own educational background. *"interdisciplinary collaboration helps in answering questions that cannot be met using a single strategy or method"* [Brewer, 1999].

Marzano et al. [2006] stated that success in interdisciplinary collaboration comes from the time and effort put in the order to generate the necessary interpersonal relationships.

Long-term collaboration also does not resolve communication barriers.

Moreover, the authors added: *"Building effective relationships within an interdisciplinary team requires good leadership, trust, receptiveness and a willingness to learn"* [Marzano et al., 2006]. Moreover, overcoming the communication barrier between disciplines in understanding each others' esoteric terminologies plays a major role in interdisciplinary challenges. Besides, long-term collaboration between researchers also does not necessarily resolve understanding the knowledge systems between counterparts.

CoE-Portal is available for all researchers at Cluster of Excellence.

As a solution to support and steer interdisciplinary collaboration, RWTH-Aachen University has initiated a local social networking website for the Cluster of Excellence[1] (CoE-Portal[2]). This portal is available to about 300 researchers of about 25 disciplines and every user has their own profile page. This common portal enhances the process of collaboration by giving information on one another and allowing sharing of data. However, researchers still lack enough information in order to encourage the discovery of collaborators for interdisciplinary research [Schaar et al., 2013a].

2.2 Relevant Work and Challenges

2.2.1 Bibliometrics

List-based structures such as Google Scholar or Scopus give users an insight into their work. As list-based approach has also been discussed in Garfield [2006]. They illustrate how well a paper is cited and how popular it is among the world of scholars. In such list-based representations, the results are presented as lists and that list is ordered according to certain criteria like most recent, most cited, etc. However, a complete listing of publications covered would also be helpful in evaluating the significance and impact of a particular work on the literature and related works. In such database approaches, Harzing and Van der Wal [2007] point out the challenges in *database coverage*.

[1]Integrative Production Technologies for High-Wage Countries http://www.production-research.de

[2]http://www.coe-portal.com

Database search engines have to scan vast number of documents to track citations and assign them to each respective paper. In addition, authors also make mistakes in writing their bibliographies, every researcher might have their own citation style and another typical challenge in bibliometrics is similarities in author names. One cannot relate the number of citations to the performance of a researcher, while the number of citations does not represent agreement or disagreement of other scholars regarding their research. Even beyond these difficulties, extraction of self-citations, honourable mentions and their meanings are computationally expensive and difficult. Moreover, Redner [1998] discussed power law distribution of citations. In other words, typical statistics such as means or variance are also not meaningful as most papers get very few citations and a limited number of papers get many citations.

Bibliometrics is statistical analysis of written publications.

2.2.2 Mixed-Node Publication Visualisation

Here we try to give an insight into publication visualisation to understand the data mathematically as well as visually from different perspectives.

Calero Valdez et al. [2012] used visualisation of publications as a means to analyse success in interdisciplinary teams. The authors used mixed-node publication network graphs to get insights into the social structure of research groups. The approach of this visualisation tool was to depict cooperation based on co-authorship. As Figure 2.1 illustrates, publications are depicted by little nodes, authors by medium-sized nodes and the authors' disciplines by large nodes.

Mixed-node representation was used to illustrate co-authorship in a research community.

The authors created an animated graph of co-authorship in a small cluster, the resulting graph is shown in Figure 2.2. It utilised Gephi[3] graph analysis to discover graph density, network diameter, average path length between nodes, average degree to depict an average of publication per author,

[3]Gephi is an interactive visualisation tool for all kinds of complex dynamic networks and hierarchical graphs.

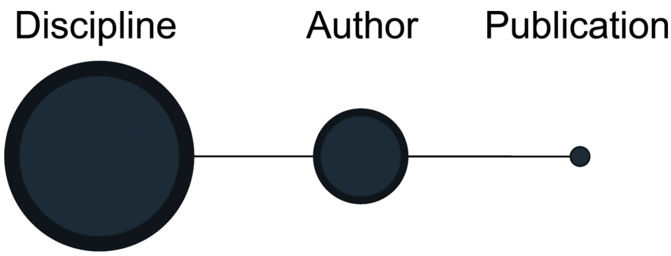


Figure 2.1: Mixed node publication graph with different types of nodes [Calero Valdez et al., 2012]

.

centrality to measure the importance of each node and community detection to identify groups of connected people.

This visualisation was successful to reveal qualitative data of researchers.

This graph can reveal qualitative data such as who talks to whom and the possibilities for changes to influence social network behaviour as a whole. Interdisciplinary team cooperations can be evoked using this visualisation as it can cause a hedonic gaming attitude on how to increase interdisciplinary success.

Later, Schaar et al. [2013a] stated that there is still a lack of knowledge on how to measure, steer, support and manage interdisciplinary success. Hence, they performed user studies on purposed visualisations of interdisciplinary collaboration and concluded that:

Visualisation of publications can address challenges in interdisciplinary collaboration.

- Mixed-node publication network graphs are able to visualise interdisciplinary structures.
- Mixed-node publication graphs are suitable to steer huge research clusters or groups.
- There is a positive impact of this graph on interdisciplinary work.

Moreover, Calero Valdez et al. [2014] conducted 22 interviews to evaluate mixed-node visualisation (Figure 2.2). By analysis and transcription of their interviews, they identified a number of agreement and disagreement concepts on mixed-node graphs. The results showed that visualisation of collaboration regarding publications can be used for *retrospective analysis, information regarding the team* and *planning*. On the other hand, there is *missing information, negative influence on work flow* and *bad legibility*.

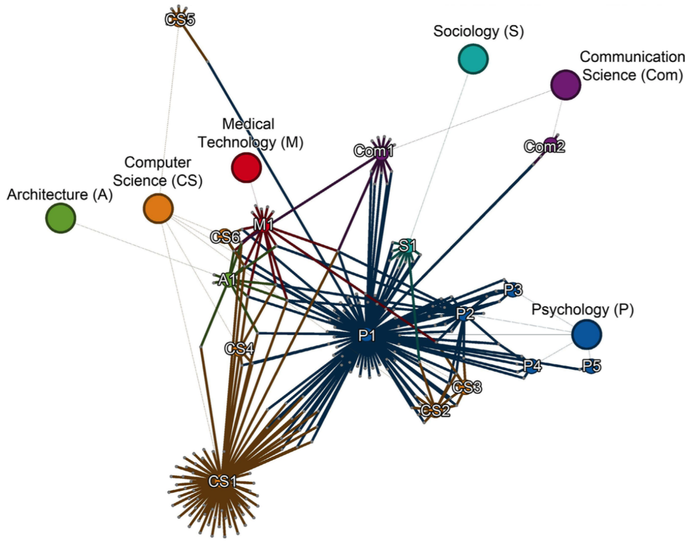


Figure 2.2: Sample of mixed node publication graph. The graph is animated using force-layout to show moving of nodes and giving the impression of birds-eye-view of moving people that group together.[Calero Valdez et al., 2012]

Additionally the authors introduced new requirements for future publication visualisations. These were assessed to be: increasing usability, a time axis, additional information such as impact-factors, journal names, filter functions such as hiding professors or institutes, sociometric data and profile pages for each author.

Chapter 3

Related Work

In this chapter we distinguish different related works on information visualisation research projects and introduce their approaches that have improved visualisations of data. Throughout this chapter we explore different categorised information on visualisation research.

As also mentioned by Munzner [2008], there are three related main domains, namely: algorithms, summative user studies and design studies. Each of these domains was described with relevant examples to give readers a precise understanding of the domain. Using these literature reviews we have driven our research questions and research methodology.

We explored previous related information visualisations.

3.1 Algorithms

In this section a new algorithm is the contribution of the research. Often, proposed techniques refine or improve previous developed algorithms. Usually the authors of these types of information visualisation papers claim that their algorithm is more efficient than previous works.

3.1.1 A Million Items Visualisation

Fekete and Plaisant [2002] described a technique to visualise a million items which rely on hardware acceleration to achieve necessary high-density interactiveness and stereovision or synthetic overlap count to enhance visualisation.

User zoom into areas of interests to investigate that area.

A coarse representation of million-node networks allows users to see clusters without aggregation, compare their sizes and comprehend their connectivity. Users are also able to zoom in on areas of their interest for more exploration. This atomic pixel-based visualisation uses one marker per data record with filtering to show subsets and zooming to focus on areas of interest.

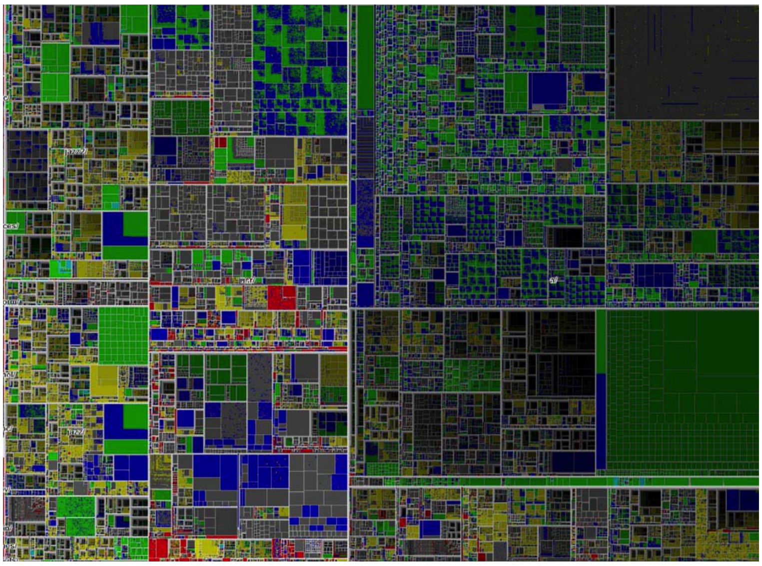


Figure 3.1: This treemap gives an overview of 970,000 files of a file system containing 1 million files on a 1600x1200 display. The size of each rectangle is determined by the file's size; colour represents file type and deeply nested directories appear darker.[Fekete and Plaisant, 2002]

3.1.2 Extreme Visualisation

Shneiderman [2008] introduceed two algorithms (**Aggregated** and **Density Plots**) to facilitate interactive visual exploration of big data sets.

Aggregated: Represent a sense-making scalable algorithm for carrying out billions of records. The work suggests using one marker per thousand data records by squeezing the information into a million pixels. This visualisation uses a semantic substrate to layout the nodes in a grid plot, then the contents of each grid cell are represented by a meta node with a size proportional to the number of nodes in that cell. Aggregation markers are then used to organise data and to suggest where and when users should click. After clicking on each cell, the contents of that cell can be represented in different ways as needed.

Aggregated algorithm can reveal communities in a network.

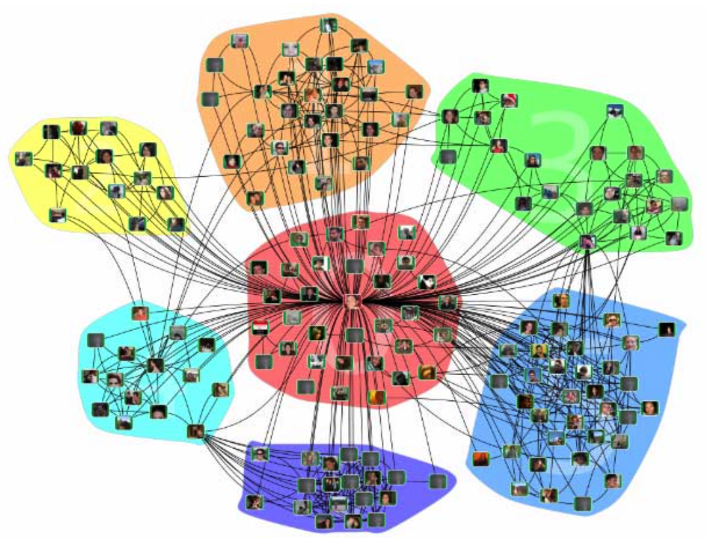


Figure 3.2: Aggregated: Grouping nodes into community structures based on link relationships in Facebook network. Each community can be replaced by a single aggregate node, enabling scaling up to large databases.[Shneiderman, 2008]

Density Plots: Uses colour codes to suggest potential areas to explore. Density plots take advantage of clustering strategies and colour codes to organise nodes into aggregate markers. A special form of aggregation is the density plot two-dimensional histogram which uses a spatial substrate organising principle and shows concentrations of markers.

Pixel concentration suggest area of more investigation.

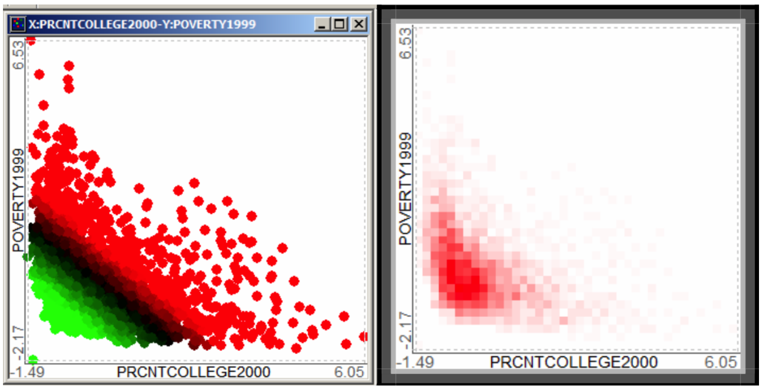


Figure 3.3: Density Plots: Representation of a scattergram (on the left) which is heavily over plotted, but converting to a 40 by 40 grid plot (on the right), enables users to see the distribution density. Clicking on a grid cell brings up the records in that cell.[Shneiderman, 2008]

3.1.3 Hierarchical Edge Bundles

Holten [2006] focused on construction of a generic technique for visualisation of compound directed graphs. It comprised a tree and an additional adjacency graph. The compound graph is stated to be a frequently encountered type of data set. They introduced hierarchical edge bundling as a flexible and generic technique for the visualisation of compound (di)graphs.

The author alleviates the problem of intertwined inclusion and adjacency edges by treating the tree and the adjacency graph as a single clustered graph. Moreover, they claim that hierarchical edge bundling provides an intuitive and continuous way to control the strength of bundling.

Parent-child relations are henceforth called inclusion relations and non-hierarchical relations are called adjacency relations.

The approach is to use the path along the hierarchy between two nodes having an adjacency relation as the control polygon of a spline curve; the resulting curve is subsequently used to visualise the relation. They also use bundling strength ß as a variable to change in the balloon layout visualisation. Low bundling strength mainly provides low-level, node-to-node connectivity information, whereas high bundling strength provides high-level information.

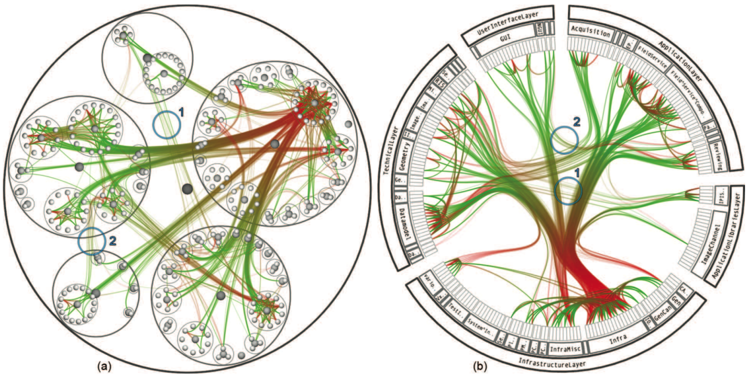


Figure 3.4: A software system and its associated call graph (caller = green, callee = red). (a) using balloon layout and (b) suing a radial layout. Both have bundling strength of ß= 0.85. [Holten, 2006]

Bundling reduces visual clutter, making it easier to perceive the actual connections as when compared to the non-bundled versions. Bundled visualisations also show relations between sparsely connected systems more clearly (encircled regions); these are almost completely obscured in the non-bundled versions.

Bundling visualisations are easier to show actual connections with less visual confusion.

Participants in this research have quickly gained an impression of the high-level connectivity information while still being able to inspect the low-level relations that were responsible for the bundles. With respect to the user study, the author suggests that their technique is useful for quickly gaining insight into the adjacency relations present in hierarchically organised systems.

3.2 Summative User Studies

Summative user studies focus on evaluation of an information visualisation system by some target users in the target domain. Typically this domain conducts formal user studies using tasks and then measures the statistically significant main effect of the experimental factors in order to examine the strengths of a system or technique. The scope of

Scope is, discussing evaluation strategies in information visualisation with respect to a user group.

such information visualisation research is to evaluate a system, quantitatively and qualitatively, in target user groups.

3.2.1 Analysts in Visualisation

Slingsby and Dykes [2012] argued for the need to distinguish between domain experts and visualisation designers. The authors found the need for involving analysts in visualisation designs and bridging the gap between designers and analysts by advancing mutual contributions and understandings between them. As the authors suggested, this is done by (a) meetings or workshops, and (b) iterative and short bursts of prototyping.

The authors then relied on Koh et al. [2011], Floyd et al. [2007] and claimed that the key aspect of their method is visualisation awareness and rapid design iterations over a short period of time, to reduce the comprehension gap between analysts and designers.

Rapid prototyping in short design cycles are encouraged.

The authors recommended that designers and analysts perform the following procedure for information visualisation tools: *"The core of the process is an iterative loop intended to last about five days, where intensive design and prototyping are followed by a feedback session where ideas and prototypes are presented and discussed. These are either endorsed, modified or discarded and requirements and priorities updated for subsequent iteration. Influenced by ideas from Agile software development[1], requirements and priorities are modified as ideas evolve, ensuring that requirements and designs remain relevant."* [Slingsby and Dykes, 2012]

The process of implementation starts with simple paper print-outs, web-based interfaces or professional applications, as long as the analysts can evaluate the design using identified tasks. Nevertheless, implementation phases might not be simple due to time limitations and design difficulties, hence Slingsby and Dykes suggested breaking down the process to multiple feedback iterations with analysts. The analysts are promoted to explore their own data

[1] http://agilemanifesto.org

without intervention by designers and the tool is examined and used in a realistic work context without being observed or monitored.

3.2.2 Quantitative Empirical Studies

Lam and Munzner [2008] posited that the benefits of studying information visualisations can be enhanced by quantitative empirical studies of meta-analysis. This paper suggested that researchers should deliver improvements to information visualisation by implementing the following instructions:

1. Use comparable interfaces in terms of levels of data, interaction complexity and information content displayed. This is done on typical head-to-head system comparison studies.

2. Capture usage patterns in addition to overall performance measurements (e.g., time and accuracy) which support us with interface efficiency and user acceptance, but may not be sensitive enough elaborate design trade-offs and interface use. Interaction recording, such as eye-tracking records or navigation action logs, can be used for this purpose.

Authors believe longer term studies provide more complete pictures of system use.

3. Isolate interface factors instead of overall interface efficiency. This is because showing the whole interface in a single view can cause more complex interactions with higher visual memory costs, hence it is suggested to divide the interface to multiple views before evaluations.

4. Report more study details using supplementary materials.

Therefore, the authors argued that by using this approach, the consistency and utility of these studies would be increased.

3.2.3 MILCs

Ethnographic process can increase trustworthiness and credibility of a system

Shneiderman and Plaisant [2006] presented Multi-dimensional In-depth Long-term Case Studies (MILCs) as a research strategy to assess the development of information visualisation. Multi-dimensional refers to evaluation of user performance and user interface efficiency through interviews, surveys, automated logging and observations. In-depth refers to the ability of the researcher to be close enough to the end-users to explore all users' needs. Long-term refers to weeks or months of observations to fully understand a system's user domain. Case studies refer to detailed reporting of system usage in normal environment settings.

To evaluate information visualisations using the MILCs method, the following procedure is proposed by the authors:

- Specify research questions and goals.
- Identify three to five users who are willing to participate in system evaluations for a long period of time.
- Document the current method which is used and the current version of tool being tested.
- Determine what would constitute success for the users. Every type of user has different success values.
- Establish a schedule of observation and interviews.
- Generated reports at each visit and record usage data.
- Encourage the users to record difficulties and frustration, as well as successes.
- Observe the learning curve of the users.
- Conduct interviews.
- Modify tools as needed.
- Document successes and failures. Immediately after each visit or interview, reflect upon the lessons learned.

Outcomes for MILCs fall generally into two categories:

1. The refinement of the tool over time and an understanding of the general principles for the design of such tools.
2. Achieving end-users' goals by employing the tool as needed.

However, with respect to the time constraints incumbent upon researchers, the outcome may be only specific suggestions for tool improvements and a better understanding of design guidelines.

3.3 Design Study

Here, a target problem is identified and their requirements are determined through task analysis, then design choices are justified in terms of how well these interaction mechanisms will fulfil the users requirements. A case study analysis is a typical approach for these projects along with an iterative design for refinements. The contribution of papers which employ these methods is a well-designed interface using a combination of existing techniques. The authors also give the readers enough evidence proving that their design study has met and fulfilled the target users' requirements.

A visual representation as a suitable solution for a specific domain problem.

3.3.1 Vizster

Heer and Boyd [2005] undertook an ethnographic case study on an online social network. The authors developed a visualisation representation of online communities which facilitated end-users to explore and navigate through their articulated social network in a playful manner. The developers of this visualisation utilised a node-link network layout, search functions and community structure algorithms.

Heer and Boyd built their design decisions upon Friendsters' ethnographic studies [Boyd, 2004].

A visual environment for the exploration of online social networks, including both topological and profile data.

Vizster used a node-link illustration, where nodes represented members and links represented the friendship or connection between members. The authors used an egocentric network presentation, consisting of a person and their immediate friends and friends of friends. This visualisation used a force directed algorithm which repelled nodes and edges away from each other.The Newman [2004] community identification algorithm was used for automatic determination of communities.

Vizster used a so-called "X-ray" mode to filter the visualisation by using specific attributes such as number of friends, gender, relationships or age. By switching to X-ray mode, the background turns black and member profile photos are removed and the node's colour changes with respect to that particular selected filter, then a legend is displayed at the top-right corner to guide the user for interpretation of the visualisation.

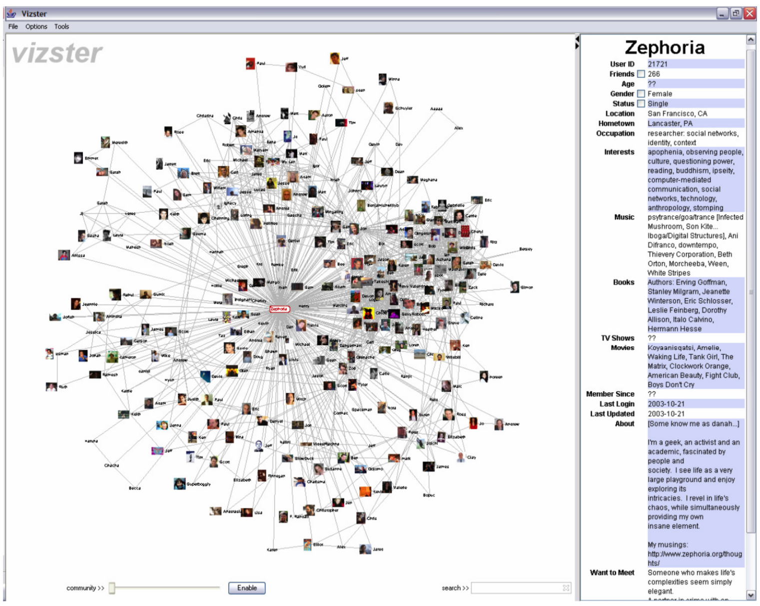


Figure 3.5: Vizster: The left side presents a network display with controls for community analysis and keyword search. The right side consists of a panel displaying a selected member's profile information. [Heer and Boyd, 2005]

Hovering on a node, it will highlight that individual node and its connections (friends-of-friends) in the larger network context. Clicking on a node will show corresponding members' profiles. Panning and zooming help with the navigation of users throughout the network.

As the authors mentioned, the goal of Vizster is to build a visualisation that can facilitate discovery and increase awareness of their online network.

3.4 Conclusion of Related Work

We have conducted a literature review on the information visualisation field. We explored research approaches used in the three main domains of information visualisation (algorithms, summative user studies and design studies) [Munzner, 2008]. These related works gave us insights into the development of our prototypes. Fekete and Plaisant [2002] used filtering functions, allowing users to explore more areas of interests. Colour codes and clustering were used to suggest an impression of the high-level data, while at same time being able to inspect lower-level information [Shneiderman, 2008, Holten, 2006].

Through reviewing summative user studies we acquired enough information on our methodology to use for our project. We identified our target users and we concluded that we need iterative and short bursts of prototyping along with multi-dimensional in-depth case studies. This can be achieved by specifying research questions and goals [Slingsby and Dykes, 2012, Shneiderman and Plaisant, 2006]. Moreover, to our knowledge no other comparable interfaces existed, hence we have emphasised on usability of our design rater than comparative studies.

As we aimed for a design study of a visualisation, we did not focus on algorithms but rather a justification of our design choices for their specific use. In the next chapter we derive five research questions and discuss in-depth the methods used to find solutions to our questions.

Chapter 4

Experimental Design

Using literature reviews we derived five research questions and discussed the methodologies used throughout the whole design process. These methodologies assisted us in the evaluation and development of our prototypes with respect to our target user domains.

4.1 Research Questions

By utilising findings in related works to visualise possibilities in publication collaboration we investigated the following research questions:

Research questions derived from literature reviews and our target user domain.

RQ1 What are users' expectations of a visualisation tool to enhance collaboration?

RQ2 Can a visualisation approach be used to suggest collaborators?

RQ3 Can the visualisation at the same time inform members about how the organisation is structured?

RQ4 What other applications can a visualisation can have besides its main goals?

RQ5 How well can our visualisation steer interdisciplinary success?

4.2 Meta-Method

Meta-Method is the general approach used throughout the whole design process. The minor changes at every iteration of the design have been discussed in each respective method subsection.

4.2.1 DIA and User-centred Design

[Slocum et al., 2003, Nielsen, 1994b, Sedlmair et al., 2012] described the process of the DIA (Design, Implement, Analysis) cycle and user-centred design to increase the usability of the created artefacts for the end-users (Figure 4.1).

It generally involves establishing context of use, requirement gathering, design, deploy and evaluation. Each of these stages is repeated in periodic iterations until it satisfies users needs. This design principle is facilitated using rapid prototyping starting with low fidelity prototypes (e.g., paper) to high fidelity prototypes (e.g., HTML, JavaScript) [Floyd et al., 2007]. To ensure high usability of a design for end-users, at every iteration, the functionalities and design of the prototype are evaluated with respect to users' feedback. Users feedback and evaluations can uncover needs and fundamental problems that the designer might not have thought of. Each design cycle starts again based on previous evaluations.

User-centred design approach promises for high usability and a better user acceptance.

Paper prototypes are widely used in rapid prototyping because of fast implementation time and less cost when it comes to modifications. They can ease the process of evaluation of initial ideas and user interfaces at the early stages [Snyder, 2003]. After creating user interfaces using papers, the designer himself simulates algorithms of interface change by exchanging the papers (user interfaces) with respect to user interaction.

By utilising such an approach in this thesis, we go through the evolution of an artefact starting with a rough design to a sophisticated design. We have conducted two DIA cycles:

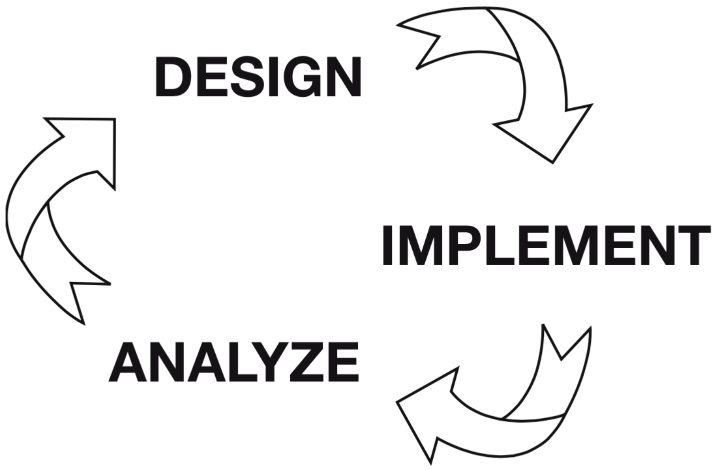


Figure 4.1: The DIA cycle

initially we started studying users to identify the context of use and the requirements through literature reviews and interviews (Section 5.1), then we created and suggested paper prototypes (Section 5.2), out of which two prototypes were selected. Based on our findings and constraints we moved to high fidelity prototypes (Section 5.3) and continued the cycle for a second time (Section 5.4). Finally, by analysing user evaluations at the third iteration, we have delivered a final prototype for our context of use which is discussed in Section 5.5.

We have obtained user needs through literature reviews and interviews.

4.2.2 Participatory Design

In order to not just fulfil the needs of stakeholders to deliver the requested functionalities, the designers emphasise the participation of users in early stages of implementation [Johnson, 1990]. In other words, real users participate throughout the design process and shape the final outcome. Hence, continuous participation of users in the Participatory Design (PD) is encouraged.

Prototypes were improved at each iteration by immediate feedback.

In this thesis, we have empowered our design by applying continuous changes with every single user study. So every design was evaluated by users, then that design was altered by analysing user feedback. In this fashion prototypes were improved at each iteration by immediate feedback evaluation from the previous users.

4.2.3 Evaluation Methods

Results from quantitative studies are typically presented in numerical data format.

Quantitative usability tests typically deal with some sort of evaluations in which the results are presented in numerical data format and used to argue for a certain degree of generalisation. This can be done by assessing task completion time, performance or questionnaires such as Net-Promoter Score (NPS) and System Usability Score (SUS). We have employed NPS and SUS as two variables to evaluate our prototype.

NPS is a questionnaire in which the loyalty of users towards a system is examined [Reichheld, 2003]. This is a survey with only one question, asking if users are willing to recommend the evaluated system to a colleague. The answer is scaled between 0 to 10, unlikely to extremely likely respectively. Answers between 0 to 6 are detractors and are unhappy with the system, between 7 and 8 are called passives and they are dropped as they are not sure about their decision, 9 and 10 indicate users who are willing to promote the system and they are very happy with it. The overall result is then displayed as a percentage.

NPS and SUS were used together to empower our quantitative studies.

SUS is a reliable tool for evaluating usability [Brooke, 1996]. It is consists of 10 questions with a scale of 0 (strongly disagree) to 5 (strongly agree). It can easily indicate usability in small sample sizes. SUS gives results between 0 to 100, but it is not a percentage ranking. SUS scores are interpreted as above average if the score is more than 68. NPS and SUS were used together in this thesis as indicators of prototype user acceptance and usability.

Qualitative analysing can enable researchers to construct a theoretical framework that can describe research findings [Hazzan et al., 2006]. The main tools used in our qualitative research were iterative interviews and observations (videotapes of interactions and sound recordings). This method assisted the researchers in investigating a system from the user perspective and it is focused on a rather small number of end-users participating in this method. This can enable researchers to obtain a very detailed documentation of actions, behaviours and reactions, or even additional environmental information. Hazzan et al. [2006] stated that qualitative research can highlight new and even unpredicted research directions which were not thought of at the beginning of the research. In our qualitative study, besides open-ended questions in interviews, we have asked users to solve problems in a think-aloud fashion. At the end, every user was given the chance to freely explore the prototypes and discuss how they feel about certain things.

video tapes of interactions and sound recordings assisted us for qualitative analysis.

One important aspect in evaluation of a system is triangulation [Rester et al., 2007, Tory and Staub-French, 2008, Kaplan and Duchon, 1988]. Meaning, collecting of information from different resources. We have used triangulation by employing qualitative and quantitative analysis side-by-side. As quantitative research does not explore all aspects of a complex system, qualitative evaluation was also encouraged [Shneiderman and Plaisant, 2006]. Moreover, we will see in the requirement study (section 5.1), the qualitative study also gave us highlights on users' mental models of social processes and disciplines.

Due to our small user sample, we combined qualitative and quantitative studies.

4.3 Participants

There are currently over 200 researchers from about 20 disciplines working in the Cluster of Excellence at RWTH-Aachen University. Interdisciplinary collaboration is important and highly encouraged. We have distinguished three different user categories for our studies, namely: newbie (with two or fewer publications), regular users (more than two publications) and management (coordinators of the Cluster of Excellence). From this population we selected about 70 participants for our studies by randomly selecting researchers. Fifteen participants from seven different institutes agreed to take part in this study (see Table 4.1).

Category	Sample	Req. study	1st Study	2nd study
Newbie	4	2	3	1
Regular User	10	3	7	3
Management	1	0	0	1
Total	15	5	10	5

Table 4.1: Selection of participants for all studies

4.4 Network Analysis

Gephi was used for network analysis of our data sample.

Beside the focus of this thesis being a design study, we needed a simple technique to recommend collaborators to one another. Therefore, in order to better understand how our network is structured and how users are connected to one another, we have used Gephi[1]. Gephi is an open source software for graph and network analysis. It assisted us in identifying some main requirements for our network visualisation and recommendation system. Gephi outputs can deliver interpretations of a network in high quality layout algorithms, clustering and data filtering through interactive exploration [Bastian et al., 2009].

For the purpose of a suitable recommendation system, we utilised statistics on our available network database. We needed to know how well our users knew each other in this network. Therefore, we acquired and used a real database of publications between 2012 and early 2014 from the Cluster of Excellence at RWTH-Aachen University. Initially co-authorships were extracted and then this database was used in Gephi to obtain the needed statistics and visualisation. The following parameters were investigated from our database:

Graph diameter: Maximal distance between any two nodes.

Average path length: Average length from any node to all other nodes.

[1]Available online: `http://gephi.github.io`

Graph density: How connected a network graph is.

Clustering coefficient: Degree to which nodes in a graph tend to cluster together.

In the Figure 4.2, we can see the Gephi visualisation of our raw database using a force-vector algorithm available in Gephi.

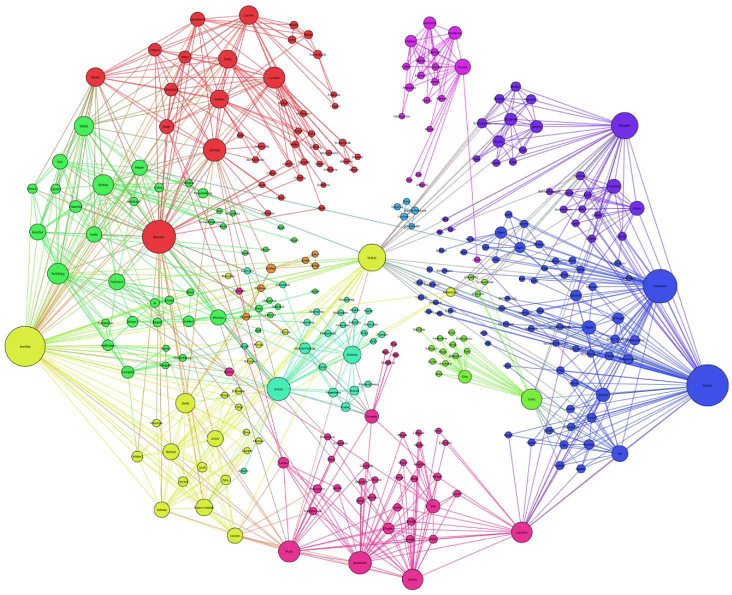


Figure 4.2: Gephi visualisation of our network including professors.

From the graph statistics (Figure 4.2), we found a network diameter of 8, with an average path length of 3.43. Graph density was 0.014 and the clustering coefficient was 0.617. These results mean that our network was very dense and there were influential nodes (Professors) which interconnected all nodes together. The bureaucracy of co-authorship in our target domain of users was that the professor's name was always published as co-authors, but they were not necessarily involved with the process of research in the field. Additionally, we discovered that all researchers employed by a professor do not necessary know each other either. Thus, we ran our network analysis one more time, but this time we removed all known professors from our database. The resulting visualisation is shown in Figure 4.3.

Professors at Cluster of Excellence created a dense network of collaborators.

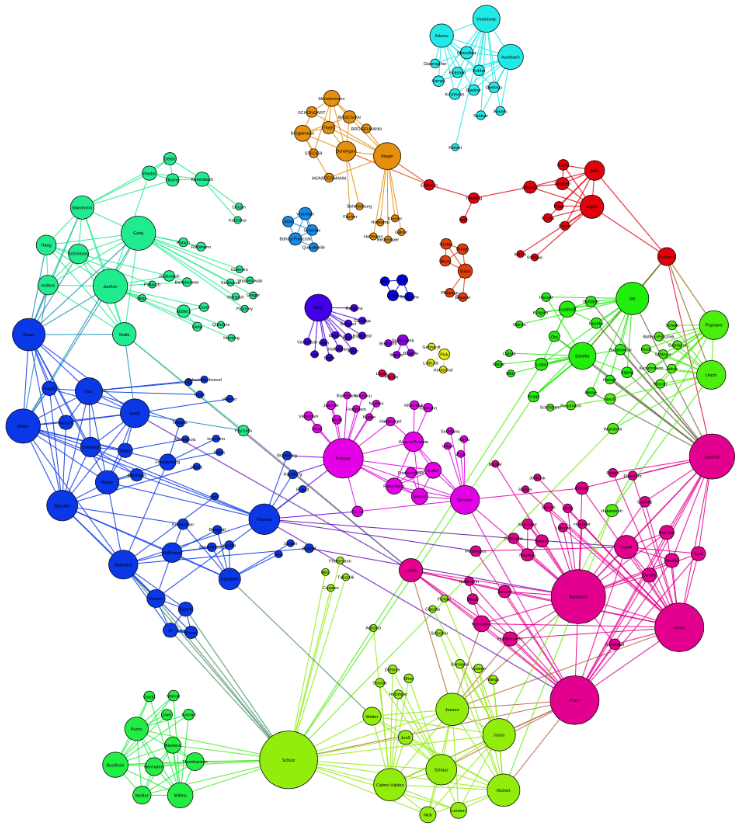


Figure 4.3: Gephi visualisation of our network excluding professors.

The analysis revealed the following statistics on our graph without professors. We found a network diameter of 9, the average path length was 3.88. Graph density was 0.01 with a clustering coefficient of 0.394. The statistics showed that our network density decreased rapidly. Therefore we concluded that a path length of 2 can be used for our recommendation algorithm. In other words, we could recommend co-authors by using path length as a parameter for a co-author recommendation system. This could be achieved by excluding professors from network analysis and exploring a path length of 2 for each single node. However, a path length of 2 could not be used alone for recommending co-authors as it did not imply similar interests. Hence, we used a path length of 2 with similarities of terminologies together to indicate a more accurate recommendation.

4.5 Conclusion of Experimental Design

Empirical methods are the main approach used in evaluating user interfaces [Nielsen, 1994a]. Nielsen proved that the best results can be achieved by combining several methods. Hence in our project, we initially started with literature reviews and later we undertook a requirement analysis by interviewing users. Then as the DIA suggested, we created multiple paper prototypes, and two of these prototypes were selected for a data-driven evaluation (Section 5.2). This was due to less intuitiveness of static visualisations. Thus, we moved on to a higher fidelity of publication visualisation using real data (Section 5.3). We then used Nielsen' heuristic evaluations as the first and most informal method to discover usability issues before performing user studies [Nielsen and Molich, 1990]. One should note that the suggested prototypes are proposed to be used on social networking platform (CoE-portal[2]). Therefore, we can take advantage of the meta-data given in profile pages.

The participatory design method assisted us in discovering new insights into our design. The prototypes were improved at each iteration by immediate feedback evaluation from the researchers. We then aggregated usability issues as one bigger DIA cycle and used it for the second iteration (Section 5.4). The same approach was used in the second iteration and this resulted in a final prototype as discussed in Section 5.5. RQ4 was addressed throughout both the first and second iterations. The amount of change request reduced with respect to repetition toward the end, and at the end only a few changes remained and agreement with the type increased over time.

Moreover, with respect to our network analysis, we used simple algorithms to recommend collaborators. Keyword similarities beside network path length of 2 were used as parameters to suggest collaborators to researchers. Validity of this technique was investigated during user studies.

In addition, taking into account the limitations of our publication database, we obtained the following variables that

[2] http://www.coe-portal.com

were later used in our prototypes.

1. Overall number of papers per person.
2. Number of papers per year, per person.
3. Year of publication.
4. Expertise keywords extracted from paper title.
5. Number of previous co-authors.
6. Name of previous co-authors.

We also discovered that there is a scoring system based on the number of authors per paper and order of their name as co-author. Due to the fact that such a scoring system was believed to be inaccurate, we dropped it from our reliable variables.

Due to the fact that we had a limitation regarding the number of participants, we improved the validity of our system by combining user-centred design and participatory design methodologies and then applied both qualitative and quantitative evaluation methods to our user studies.

Chapter 5

User Studies

Findings and research methods discussed in previous chapter assisted us in orientating our research approach and design process. Hence, in this chapter we discuss starting with user requirements and following that the use of them to develop our low fidelity prototypes, and finally, the high fidelity prototypes. We explain our prototypes and discuss the results of two user studies.

5.1 User Requirements

Besides the literature review, we found a user requirement study to be necessary for our domain of users. In this we have explored and identified issues that researchers face during collaboration with other disciplines and co-authors.

5.1.1 Method

users needs were obtained from interview studies.

For the requirement analysis we conducted five semi-structured interviews (see Table 5.1). The interviews were divided into three sections. First, questions regarding the participants' background knowledge were asked (i.e., role, level of expertise, self-evaluation, interdisciplinary experience, software usage, interdisciplinary motivation).

The second part dealt with the process of publication and their practical experiences (i.e., track record, publishing frequency, interdisciplinary publications, favourite publications, literature study process, collaboration and publication practice, joys and frustrations). In this section we focused on the internal perspectives of the researchers [Shneiderman and Plaisant, 2006]; in other words, this included questions that directly addressed the process of writing and finding co-authors that possibly have required knowledge. It also included the perceived importance of choosing good keywords. The last part of the interview related to publishing in the cluster specifically and their context of use, in particular whether finding co-authors from within the cluster is necessary and whether other members of the cluster show a willingness to collaborate. Questions were designed to investigate understanding of and concern about interdisciplinary research. The questions are attached in the Appendix of this thesis.

The interviews were divided into three sections.

We sent out 20 invitations for the interview study and we got five researchers willing to participate (see Table 5.1). Each interview took about one hour and audio was recorded during interviews for later analysis and scripting of interviews.

Category	Requirement study
Newbie	2
Regular User	3
Management	0
Total	5

Table 5.1: Selection of participants from different experience levels for both studies

5.1.2 Results

From the transcription and sorting of the semi-structured interviews we derived a total of six requirements by categorisation (given in italics). For this purpose interviews were transcribed and evaluated according to Mayring

[2011]. We determined that researchers would like to *form a mental model* (i.e., a structural representation) of the cluster, the institutes, and the connections between researchers to improve the understanding of the main organisational research interests and orientation of the cluster as a whole. Members are willing to *present their own research interests* to others through keywords in order to identify each researcher's expertise and skills. Here they referred to *similarities of keywords* between two researchers as a satisfying indication of relatedness between two researchers. We found that members of the cluster often face the challenge of *discovering new co-authors or experts* in a specific field from another discipline that also match their research interests. Some authors have left the cluster but are still considered for consultation, but they should be identifiable. Interviewees referred to *willingness to collaborate and motivation* as key factors for identifying possible candidates that want to get involved in interdisciplinary collaboration. It is necessary to *acknowledge current and preceding research interests* to evaluate a possible collaboration. Lastly, every participant was asked to tell us what would assist them to enhance their publication process in general, this gave them the opportunity to think outside of the box and give us insights into their internal processes and wishes.

We acquired six fundamental requirements by categorisation.

From our audio transcription we extracted and sorted all statements related to the scope of our project. Table 5.2 shows extracted transcripts based on our categories. The results from the requirements analysis adequately address RQ1 (see Section 4.1).

During the interview studies, we also encountered a common research method that all our participants use to initiate research. However, language differences were a barrier to understanding each others' methodology. This common research method is described as follows :

1. Find keywords and synonyms.
2. Aggregate data using Google Scholar.
3. Read different literature and their references.
4. Check and focus on top authors and experts.

Category	Transcripts
Form a mental model	*"There are meetings in this regard but I can only imagine my own group of researchers."*
Present their own research interests	*"I was talking to some of my friends in cluster and then they came to know my topic."*
Similarities of keywords	*"I believe other experts in my field also use similar terms and keywords."*
Discovering new co-authors or experts	*"I don't have access to expert or I don't know anybody with experience in my topic. "*
Willingness to collaborate and motivation	*"Everyone wants to do so, but not every one comes to the desk."*
Acknowledge current and preceding research interests	*"Someone asked me to be his co-author for a topic, but I was working on that topic many years ago."*

Table 5.2: Example transcripts from requirement study

5. Get possible external help.
6. Sort out irrelevant information.
7. Reading papers until no information is new.
8. Design and run an experiment.
9. Write results.
10. Get feedback from experts before publishing.

As our project is focused on a design study, the research method was explored no further, but more investigation is encouraged.

5.1.3 Discussion

Using semi-structured interviews, we asked and explored the basics of users context of use and their requirements. Besides huge inconsistency in interdisciplinary research, communities of researchers were usually also shaped by coordinators of Clusters of Excellence. Pre-organised communities have caused researchers to ignore their full capabilities and interests in research. With respect to our interviews, it was clear that researchers needed another organisational method of interdisciplinary research. The explicit structural representation of a cluster was an issue, whereas some believed that they have a mental model for organisational representation. Authors used social collaboration as a means by which to inform others of their field of research and interests. It was also used for finding an expert to contact in the case of the need for assistance and guidance. Participants found keyword similarity as a good indication of closeness in the world of expertise.

Researchers needed another organisational method of interdisciplinary research.

Moreover, researchers would liked to expend as little effort as possible on the formation of groups, often caused by previous negative co-authorship experiences. Hence, we needed to find a way to address this issue, encouraging researchers to approach each other with renewed confidence in publication success and with the least amount of effort. Additionally, participants stated that their mental image of the cluster was indeed bubble shaped (instead of hierarchical as in a triangle for instance).

Keyword similarity could be a good indication of similar research interests.

These context of use and requirements guided us in generating paper prototypes for our visualisations which are described in the next section.

5.2 Paper Prototypes

We started our design process using low fidelity prototypes [Snyder, 2003]. Keeping in mind the requirements acquired as detailed in the last section, we started creating prototypes. The following section will describe the evolution of our paper prototypes. We started with chart prototypes and later moved onto cluster shaped prototypes. Here we will give a description of each paper prototype.

5.2.1 Simple Line Chart Prototype

Number of papers per year for each researcher.

In order to support users to acknowledge current and preceding research interests, we initially created two line charts to indicate each user's efforts in publication. Figure 5.1 shows years on X-axis and the number of total papers per year on the Y-axis. This chart was capable of showing if a researcher is no longer active in our cluster or how active they were.

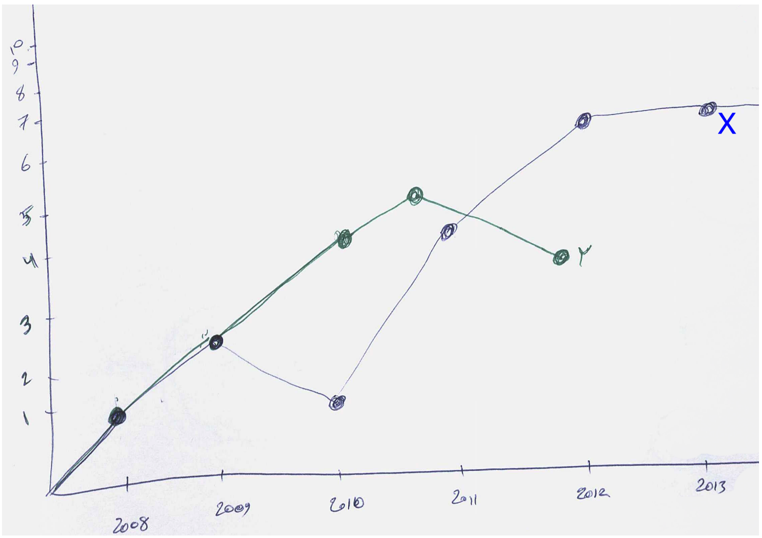


Figure 5.1: Simple Line Chart: X-axis represent years and Y-axis represent number of papers. Every line depict each persons' effort in publication of papers per year.

Active and inactive researchers were easily distinguished.

Such a graph could be used by coordinators of a Cluster of Excellence to get a broad view of active and inactive researchers, but this prototype on the other hand was unable to depict keywords of users and did not fulfil our fundamental requirements. Thus, we moved on to the next prototype as shown in Figure 5.2.

5.2.2 Individual Keywords Prototype

The individual keywords prototype (see Figure 5.2) illustrated more information on each researcher. This prototype was suggested in order to obtain each researcher's current and preceding research interests and to enable the discovery of the level of expertise of each person in a field. Overall this histogram prototype was a good suggestion allowing us to show the research life cycle of a user.

Current and preceding researches were elaborated for each researcher

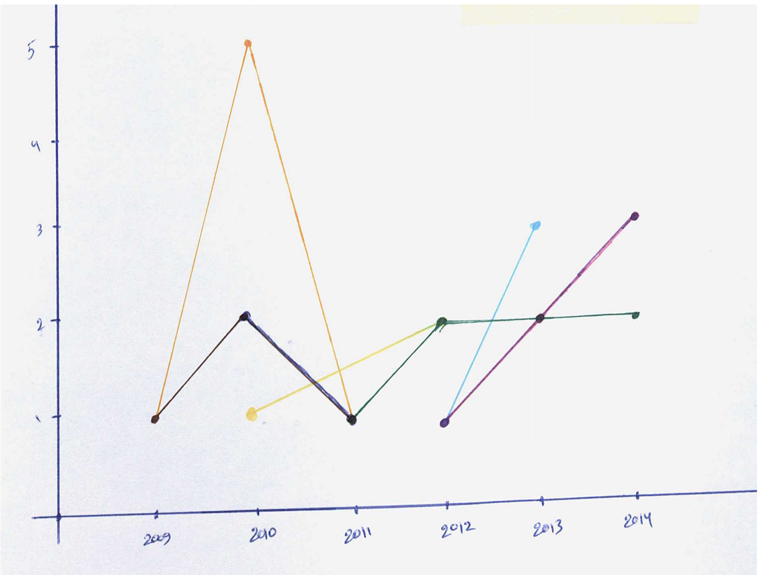


Figure 5.2: keywords Chart: X-axis represent years and Y-axis represent number of repetition of a Keywords. Every line depict each keyword.

Nevertheless, users could only see a single user at a time and the user was not able to present a willingness to collaborate. The data shown in this prototype did not depict any new information to users. Hence, such a graph was not interactive enough to make researchers use it.

5.2.3 Single Keyword Prototype

We focused on investigating the relation of researchers to a single keyword. We utilised bubble shapes because of their natural capability of showing multiple dimensions.

We tried to discover the relation of each researcher to each keyword.

Bubbles are furthermore spatially efficient and their shape naturally encodes the behaviour of transient groupings [Watanabe et al., 2007]. Here we tried to discover the relation between each researcher to each keyword. A star represented a chosen keyword and every bubble represented a user who had a similar keyword. The size of the bubble varied with respect to the overall number of papers and length(distance from a node to the keyword) represented the number of times that a keyword was used per person. A node was closer to the keyword if they have used the same keyword more in their previous publications.

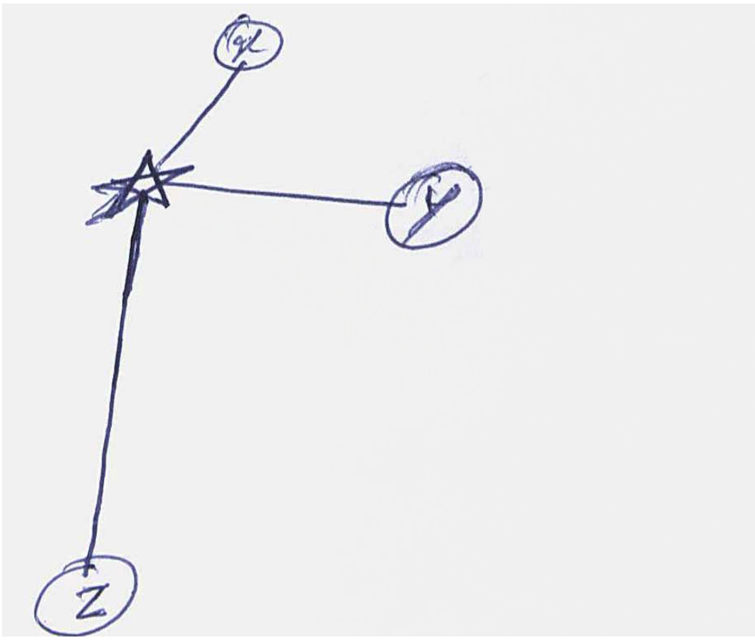


Figure 5.3: Single keyword: Star was a chosen keyword, size of bubble was equivalent to number of papers for a researcher and length of edges was equivalent to same keyword previous repetition for that researcher.

Gave insights on self-awareness on who else worked on same field.

This prototype could give users insight into who else was working on the same field. But it did not show their interest toward that specific keyword. It was also important to note that in interdisciplinary collaboration, a researcher might have used a keyword in their paper, but they might not be an expert in it. We needed a better prototype in which we can also form a mental model of collaboration teams. We touched upon this issue in our next prototype.

5.2.4 Publication Mixed-Node Prototype

Using our previous prototype properties, we created a publication mixed-node visualisation. By taking advantage of our reliable data, we visualised a network of a researcher based on their publications and their co-authors' publications. Such an approach was also used in Calero Valdez et al. [2012] visualisation of a cluster of researchers. However, here we have added other variables to our prototype. As the number of publications with a similar co-author increased, the length of edge between those two nodes decreased.

Formed a mental model of collaboration teams.

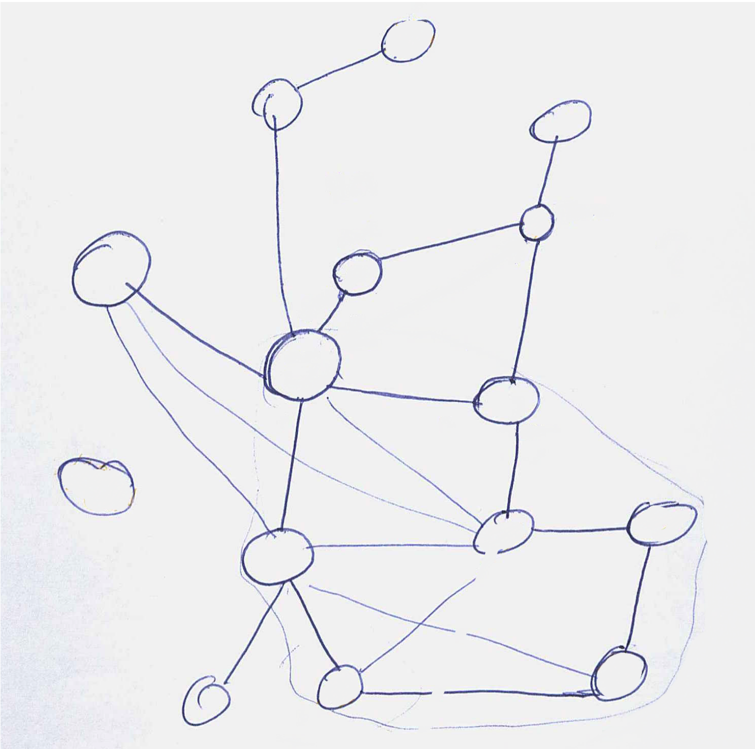


Figure 5.4: Publication Mixed-Node: Bubble size was equivalent to number of papers. Length of edges was equivalent to number publication with the other node.

From this prototype we put forward a hypothesis that later during user studies we examined and explicitly asked users about its validity:

H1 If my co-author's co-author has similar keywords to me, they are likely to be willing to collaborate with me also.

Publication mixed-node visualisation was missing a representation of keyword similarities. Additionally, we needed a visualisation that can show all needed aspects of a user in one visual in order to suggest to them a good collaborator.

5.2.5 Cluster View Prototype

A structural mental model of Cluster of Excellence was illustrated.

A cluster view prototype was created with respect to the properties of the bubble structure [Watanabe et al., 2007]. We attempted to give users a structural mental model from the Cluster of Excellence using a visualisation (see Figure 5.5). Bubble size was equivalent to the number of papers per user. Each user's bubble was located inside their institute and the bubble size of the institute was equivalent to the number of researchers and the number of papers published. Clicking on each node triggered multiple actions, (a) lines are drawn between co-authors, displaying connections between an author to other researchers and institutes. (b) It also opened a small box showing relevant information for other researchers.

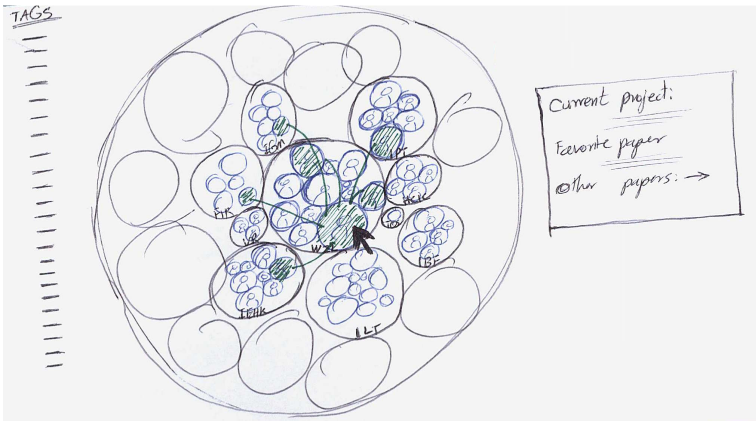


Figure 5.5: Cluster View on a Node: Clicking on a node showed their connections with previous co-authors and opened up a information page of that node.

Our visualisation was supposed to be used inside a social networking portal. Hence, here we imagined a user who has logged in to their account. Then their keywords (tags) were shown as a column on the left side. Clicking on a tag highlighted who else had a similar keyword.

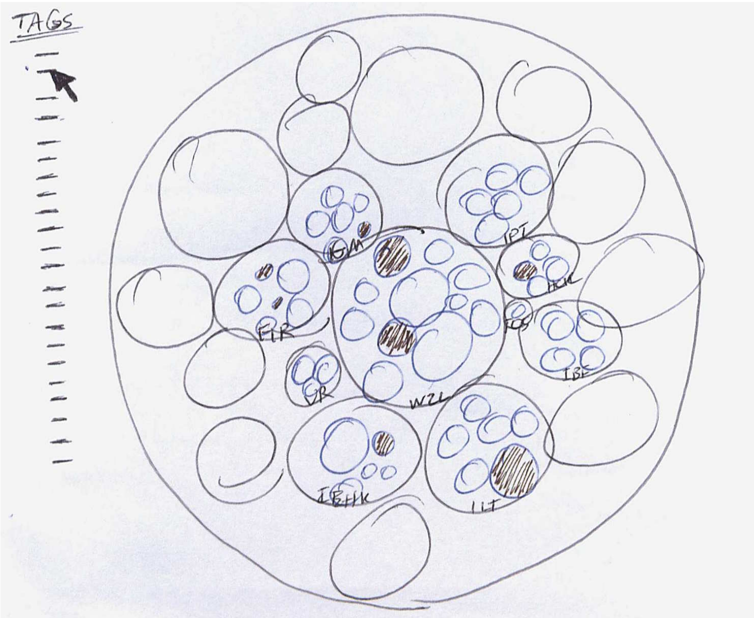


Figure 5.6: Cluster View on Tag: Clicking on a tag showed who else had that similar keyword in whole cluster.

With this visualisation we have attempted to fulfil all requirements. Using colours to show connections between nodes decreased distraction (see Figure 5.5). Moreover, there were people who leaved a research cluster after a few years but they were still available, however by using this approach we were missing these possible collaborators.

Such a visualisation could cause information overload.

5.2.6 Orbit View Prototype

To resolve the problem of high information load in the previous prototype, we created an orbit view representation (Figure 5.7). This prototype inherited the properties of the cluster view prototype, but here we used orbits to distinguish between different categories of people in the Cluster of Excellence. In first orbit showed the user himself, sec-

Orbits separated users' relevant connections.

ond orbit showed the previous co-authors (in green), third the orbit contained suggested collaborators, fourth the orbit depicted all other researchers inside the Cluster of Excellence and the outer orbit showed either previous co-authors or suggested collaborators who have left the cluster.

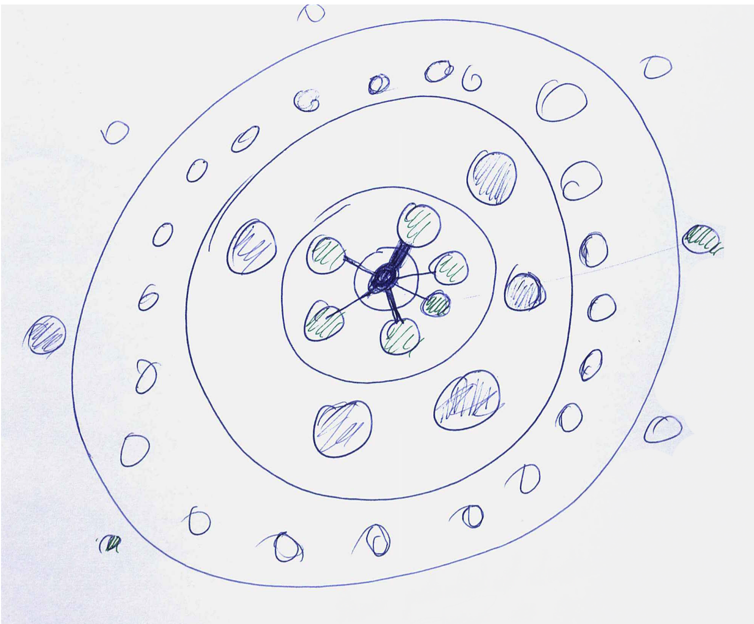


Figure 5.7: Orbit View: Relevant collaborators were sorted in different orbits.

Of course there was a trade-off between the amount of information and information overload. Here we were missing the structural mental model of the Cluster of Excellence, but on the other hand we could easily discover relevant people.

5.2.7 Conclusion

Orbit view and cluster view prototypes were selected for more investigation.

We saw the evolution of our paper prototypes and our interest in the discovery of unknown people, connections and patterns of communities. Furthermore, the orbit view prototype and cluster view prototype were selected for further investigation. The design choices made in these prototypes allowed users to visually explore and interrogate the structure of the cluster by visualising the relevant dimensions of data. Interactive bubble-bag visualisations allowed encoding of multiple dimensions (e.g., numbers of papers, keywords, institute, previous/possible connections, etc.), which were indicated as relevant information. Bubbles are furthermore spatially efficient and their shape naturally encodes the behaviour of transient groupings [Watanabe et al., 2007]. Additionally, during user requirement study participants also mentioned that their mental image of their cluster was bubble shaped.

Bubble-bag representation allowed encoding of multiple dimensions.

We observed that our participants were struggling to comprehend the functionality of our prototype using low fidelity prototypes with imaginary data, hence we decided to take our prototype into high fidelity using real data.

5.3 First Iteration: High Fidelity Prototypes

5.3.1 Method

In our first user study, we acquired a database of publications from the research cluster from 2012 to early 2014. Our database included 294 publications of 213 researchers. Furthermore, we extracted authors and keywords from the titles of the papers. Additionally, we identified authors that were no longer in the cluster. Slight improvements were integrated between trials to incorporate user feedback. The following Figures 5.8 and 5.9 are from the last iteration of our participatory design process in first iteration.

sample size included 294 publications from 213 researchers over almost two years.

We tested the developed prototypes, which were based on our requirements analysis, on two participants from the interview study and eight additional users ($N = 10$, see also Table 5.3). We evaluated them using a scenario-based speak-aloud procedure. Both final visualisations were tested in all trials. We randomised the ordering of the visualisation between subjects.

Participants were first asked to interpret the visualisation without any interaction. In a second step participants were

Category	Sample	Requirement study	1st User Study
Newbie	4	2	3
Regular User	9	3	7
Management	0	0	0
Total	13	5	10

Table 5.3: Participants of first design iteration.

We used semi-structured interviews with speak-aloud procedure.

asked to interact with the visualisation and speak about the changes in the visualisation. In a third step, finding a possible co-author was given as a task and an evaluation of the suggestion was requested. Lastly, the participants were asked to comment freely on the visualisations and compare both for suitability. The visualisations were then assessed using the System Usability Scale (SUS) [Brooke, 1996] and the Net Promoter Score (NPS) [Reichheld, 2003].

Researchers were represented as bubbles and their institute as bubble-bag.

The authors were represented as bubbles. Institutes were represented as bubble bags, containing all authors from the respective institute. Bubble size was determined by publication output and increased linearly with increasing publications (see Figures 5.8 and 5.9). The position of the each author was fixed to a relative location by using the name as a hash for its positioning within its institute. Institute bubbles contained the acronym of the institute. These design choices were made to allow users to visually explore and interrogate the structure of the cluster by visualising the relevant dimensions of data. Bubbles were furthermore spatially efficient and their shape naturally encoded the behaviour of transient groupings [Watanabe et al., 2007].

In both prototypes clicking on a bubble triggered a panel that revealed the authors name, picture and email-address. Additionally, the list of keywords and publications were shown, which could be filtered according to the year of publication. In the cluster view prototype, hovering over a bubble highlighted its connections and suggestions, whereas in the orbit view prototype these connections were predefined. Orange bubbles were used to indicate previous co-authors, green bubbles indicated having at least two similar keywords, and blue bubbles imply two common co-

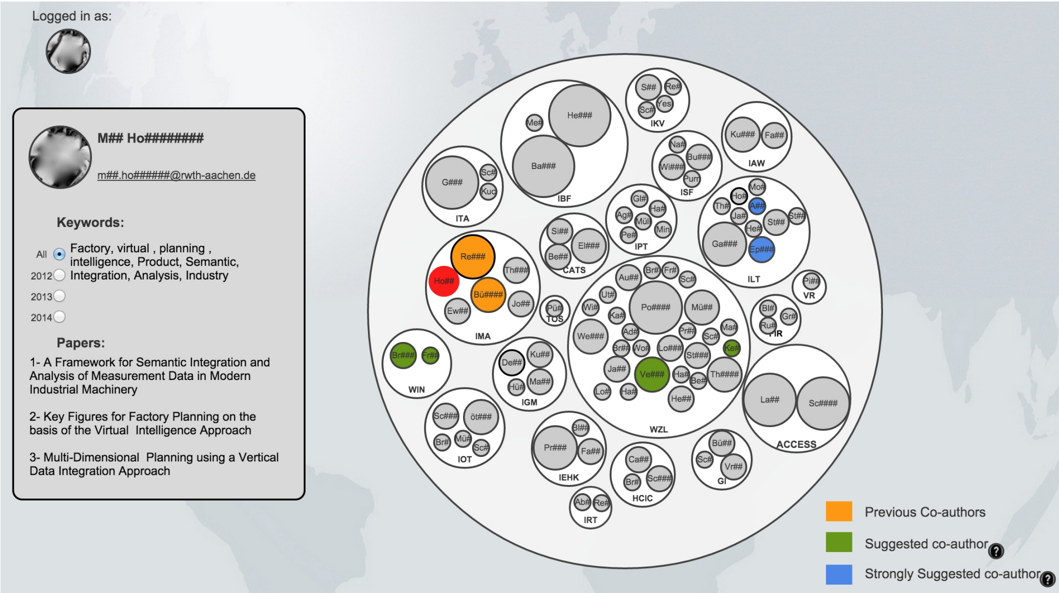


Figure 5.8: Cluster View : Showing all members of the cluster. Green and Blue indicated level of suggestions for collaboration. Orange showed previous co-authors and user itself was highlighted in red. Names were blurred for privacy reasons.

authors that also had at least two similar keywords. The user himself was highlighted in red. Moreover, we asked our participants to validate our hypothesis for strongly suggested (blue bubbles) co-authors. Both prototypes can be seen in a short video online[1].

Similarities of keywords and a shared co-authorship used as two parameters for suggestions.

We used two types of parameters to find new collaborators. We used heuristics to determine possible co-authors according to the "birds of a feather flock together" rationale [Settles and Dow, 2013]. Similarity according to keywords and a shared co-authorship network were used to find suggestions for new collaborators. The validity of extracted keywords was assessed by asking the respective interviewees.

Our second prototype focused on highlighting only the recommendations by leaving out all non-suggested co-authors (see also Figure 5.9). Suggestions were placed in orbits according to their suggestion as a co-author. Previous co-

[1] A short video demonstration can be found at: `https://vimeo.com/116338338`

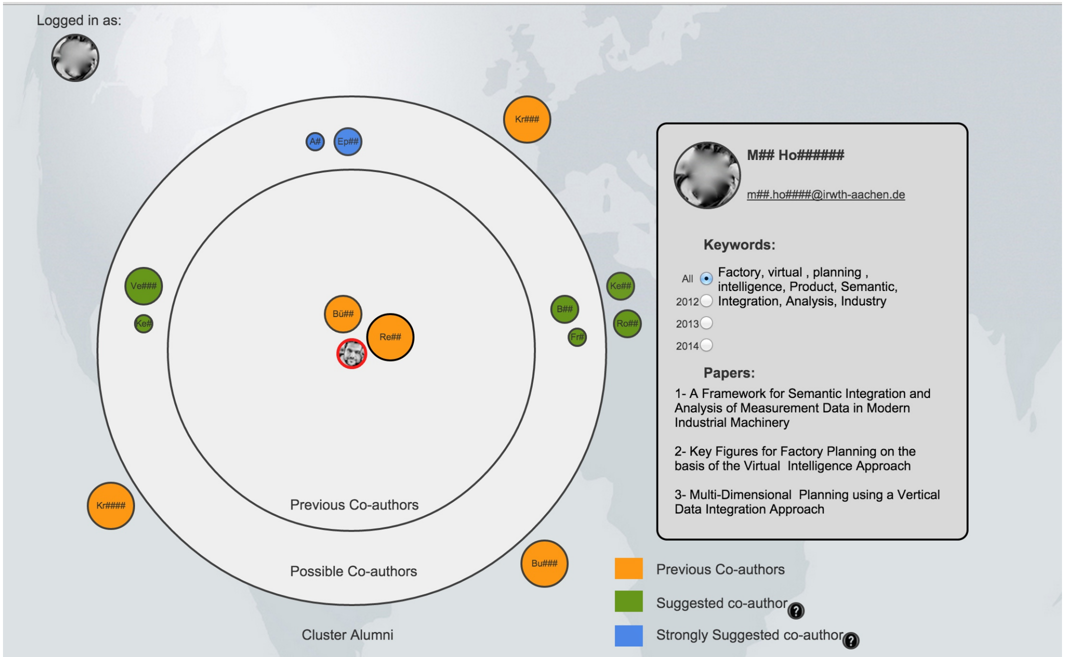


Figure 5.9: Orbit View: Showing only recommended co-authors. Green and Blue indicated level of suggestions for collaboration. Orange showed previous co-authors and user itself was highlighted in red.

authors that were not in the cluster are placed outside the bubble, addressing the requirement of also showing but at the same time identifying external collaborators. Suggested co-authors were placed in the medium orbit. Placement of bubbles within orbits was done using a force-based layout. Authors from the same institute attracted each other, while others repelled.

5.3.2 Results

Users agreed with keywords as their research interest indicator.

At the initial stage of our prototype we found that having only one similar keyword was an insufficient indicator of similarities in research interest according to the users. Moreover, the users agreed with our extracted keywords but they would liked to be able to add and edit associated keywords.

As the interviewees compared the publication efforts of their colleagues to the size of the bubble, all immediately concluded that the size of the bubble was proportional to the number of papers per person and that it represented more active and experienced researchers. Interactive bubble-bag visualisations allowed encoding of multiple dimensions (e.g., numbers of papers, keywords, institute, previous/possible connections, etc.), which were indicated as relevant by the users. The users tried to understand our suggestion system by analysing and comparing their own work, keywords and papers with previous co-authors to those of each suggested person from the visualisation. All participants agreed with our hypothesis H1 for the suggestion method used (see Section 5.2.4). Moreover, the interviewees found both prototypes complementary for each other.

Bigger bubble indicated more active researcher.

Both visualisations believed to be complementary.

All users understood the meaning of the colours by hovering over the legend which explained the reasoning for the different colours. Users found the notification system that informed them about changes in their graph helpful and necessary for long-term use. Moreover, even though the size of bubbles did not give enough space for the full name of the researchers, they could easily orientate and found colleagues by guessing abbreviations used for each user. Additionally, the participants were able to investigate the currentness of keywords and papers using the filter function inside the profile section.

Colour coding was well received by users.

Quantitatively the SUS showed a mean of *M=82.5* indicating a high acceptance of the prototype. The NPS yields four promoters, six passives and zero detractors. The overall NPS is 40 indicating good usability [Brooke, 1996, Reichheld, 2003].

Reflections on the Cluster View Prototype: This prototype supported the process of decision making by locating key players, their publication efforts and connections at the institutional level. Self-awareness, which is another key issue in large organisations, was now resolved by being able to consciously track who did what, when and where. Our visualisation also gave an opportunity for exploring the possibilities of collaboration between researchers who already knew each other. Some participants mentioned that the

Self-awareness and organisational structure were key aspects of this visualisation.

visualisation contained more information about their colleagues than they previously knew. Over all, it became clear that the users did not follow a specific pattern to rate or rank suggested collaborators. All preferred to use their own instinct and background knowledge to investigate and choose between suggestions.

Caused less cognitive load for our users.

Reflections on the Orbit View Prototype: This type of visualisation enhanced information delivery by removing all unrelated researchers. This reduced cognitive load and direct the users attention. Participants were much quicker in finding possible co-authors but lacked insights on organisational structure. The closeness of authors, caused by the force-based layout, was understood by all users. The benefit of showing external collaborators was well received by the participants. Moreover, the usefulness of being able to see users who have left the cluster but still might be a relevant researcher to contact was appreciated.

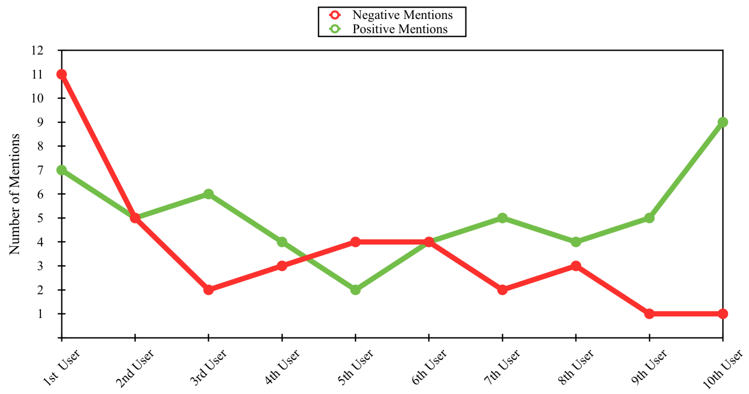


Figure 5.10: Distribution of usability mentions: Slight improvements were integrated between trials to incorporate user feedback.

Using the participator design process we have encountered a number of positive mentions (in the favour of the prototype) and negative mentions (criticisms). Slight improvements were integrated between trials to incorporate user feedback. By sorting the feedback by positive (green) and negative (red) we can indicate the effectiveness of our approach using the participatory design (see Chart 5.10). The amount of change request reduced with respect to repetition toward the end, and at the end only a few changes remained and agreement with the type increased over time. However, since not a same user was evaluating our visualisation over each trials, further investigation in this regard is needed.

Requested Features

All users were given the chance to express their wishes regarding our visualisations (see Table 5.4). We distinguished between major changes and minor changes. Minor changes were applied after each participatory design study. Here we have collected the major requested features that were mentioned repeatedly. We applied these features in our next design iteration.

Additional requirements by users feedback.

Number of mentions	Requested Feature
7	Combination of both prototypes.
6	Filtering of publication and keywords with respect to year.
6	Need for more contact information for each user.
4	Clicking on a keyword highlight bubbles which has similar keyword.
3	Getting full paper by clicking on title of each paper.
3	Clicking on a paper title highlight co-authors of that paper.

Table 5.4: Requested features in first design iteration.

Validity

From our video transcription we extracted all statements that were related to the usefulness of our system. We grouped them into three categories: *confirmation*, *discovery of new knowledge*, and *problem solving* (see Table 5.5). Each had six, six and four distinct statements respectively. From these statements we derived that our approach successfully addressed RQ2 and RQ3.

Category	**Number of Mentions**	**Sample Transcripts**
Confirmation	6	*"Oh, I have met this person at a conference recently and we have agreed to write a paper together."*
Discovery of new knowledge	6	*"I do not know the person but it seems like what he does really fits good to my work. I think I can work with him together."*
Problem solving	4	User hovers over a suggested co-author: *"This visualisation could help us having a publication from multiple disciplines."*

Table 5.5: Validity statements for first design iteration.

Possible Applications

Additional applications were obtained from interviews.

In addition to finding co-authors through our visualisation, interviewees suggested that they could also apply the system to solve other challenges such as finding literature ($N = 2$), discovering experts ($N = 3$), locating people with access to particular facilities or hardware ($N = 1$) and also simplifying the process of developing proposals for research grants ($N = 1$). From these statements we have successfully addressed RQ4.

5.3.3 Discussion

In our design rationale, similarity according to keywords and a shared co-authorship network were used to find suggestions for collaborators. Design choice of bubble shape was valid as users immediately perceived bubble size proportional to the number of papers, subsequently it represented the publication effort of each researcher. Qualitative and quantitative evaluations were used, this insured usability and high acceptance of our visualisation. Moreover, as was mentioned by Heller and Borchers [2011], colour coding conventions are at the highest level of abstraction and we have resolved this by using a legend.

RQ2, RQ3 and RQ4 were successfully addressed.

By extracting statements from video transcriptions we successfully addressed RQ2 and RQ3. Hypothesis H1 was securely supported by the study results and participant feedback. Furthermore, we discovered additional applications for our system which addressed RQ4.

Overall, feedback from the ten users was evaluated again and we decided to drop some functionalities (such as the notification system) and some were marked for improvements (keyword and paper filtering based on year). The interviewees mentioned that they preferred to have both visualisations side-by-side with the map in order to access the necessary information more easily and quickly. Hence, we performed a second design iteration to address the issues and requests.

5.4 Second Iteration: Summative Prototype

5.4.1 Method

The methodology used in this design iteration was identical to the first iteration. Major improvements were made to fulfil users requirements and needs. We also combined both previous prototypes into a single prototype to incor-

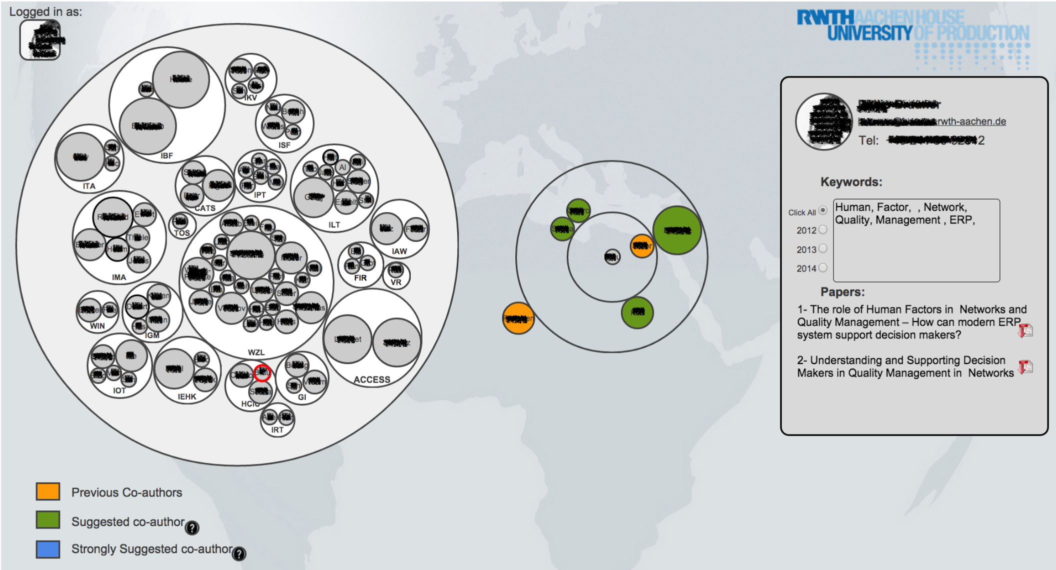


Figure 5.11: Second user study prototype: Combination of both prototypes with minor interaction changes.

We incorporated users feedback in second design iteration.

porate user feedback. We investigated our visualisation with the help of three users from previous studies and two additional users ($N = 5$). The management of the Cluster of Excellence were also involved in this study to give us more insights into our prototype from the cluster coordinator perspective. Similarity according to keywords and a shared co-authorship network were here also used to find suggestions for new collaborators.

5.4.2 Results

We integrated our users' requested features and developed another prototype, as shown in Figure 5.11. The properties of our prototype (bubble size interpretation, recommendation system, colour codings and self-awareness) received similar complements to those at the last iteration.

We obtained better SUS and NPS results.

Quantitatively the SUS showed a mean of *M=84.5* indicating good usability of the prototype. The NPS yields four promoters, one passive and zero detractors. The overall NPS was 80 indicating a high acceptance of the prototype.

The visualisation enhanced information delivery as well as giving a mental model of organisational structure to our users. All requirements mentioned in Section 5.1 were fulfilled.

Reflections on the Second User Study Prototype: This visualisation enhanced information delivery by allocating an orbit view and a cluster view visualisation side-by-side. Participants were much quicker in extracting information needed to judge suggestions by looking at the cluster view as well as the orbit view together, this reduced the cognitive load necessary to orientate themselves and their suggestions. On the cluster view users were able to easily locate themselves and had a bird's eye view of the whole organisational structure. Additionally, participants were able to discover more about the suggested collaborators by hovering over the respective bubble in the orbit view and it then highlighted each researcher's suggestions and previous collaborations in the cluster view. Even though this prototype was tested without the labelling of the orbit view, the meaning of each orbit was well received by our users without any previous acknowledgement.

Number of mentions	Requested Feature
4	View full outlet of each paper.
3	Switching between Orbit views by clicking on each bubble, within orbit view.
3	Multiple visualisation metrics to chose by user.
2	Use colour brightness to indicate value of similarities in keywords.

Table 5.6: Requested features in second design iteration.

Requested Features

During the course of the evaluation, the following requested features were extracted (see Table 5.6). Later these features were developed within the proposed design which was described in Section 5.5.

Validity

Similar to the last iteration study, we extracted all statements that related to the validity of our system. The statements were categorised into *confirmation*, *discovery of new knowledge*, and *problem solving*.

Category	**Number of Mentions**	**Sample Transcripts**
Confirmation	3	*"He is suggested to me, that make sense, we have recently published a paper together outside of cluster."*
Discovery of new knowledge	3	*"I did not know that Mr.X has done so many papers inside cluster and they have been so active in cluster."*
Problem solving	1	*"I can easily find someone to ask my questions."*

Table 5.7: Validity statements for second design iteration.

Possible Applications

In addition to finding possible collaborators through our visualisation, interviewees suggested that they could also apply this system to solve other challenges such as developing new topics for research between two institutes that had not collaborate previously, providing a representation of researchers' activeness and providing a motivation tool to encourage researchers competing with each other to produce more publications.

5.4.3 Discussion

The participatory design gave us the opportunity to enhance the usability of our system with each trial. Moreover, besides the main goal of visualisation we discovered multiple applications, hence we changed our legend from *suggested co-authors* to *suggested collaborators*.

Overall feedback from the five participants was evaluated and transcribed. As with our previous design study, we addressed RQ2, RQ3 and RQ4 using the qualitative study. Finally, all feedback and critiques were accumulated and evaluated for developing the final visualisation proposal. The design rationale of the final prototype is discussed in the next section.

We addressed RQ2, RQ3 and RQ4 again.

5.5 Final Proposed Design

In this section we discuss in detail the design and interaction of our visualisation system. A video of our proposed design is also available online[2] for further inspection.

5.5.1 Design Criteria

This prototype was initially created using D3.js Javascript library. However, for rapid prototyping purposes and focus on design study, we switched to Justinmind prototyping software solution. Our prototype (see Figure 5.12) consisted of three main parts. A cluster view visualisation, an orbit view visualisation and a profile section.

D3.js library uses HTML, SVG, and CSS to create interactive graphical interfaces.

[2] To view online demonstration, please use password: "CollVis2015" `https://vimeo.com/121127523`

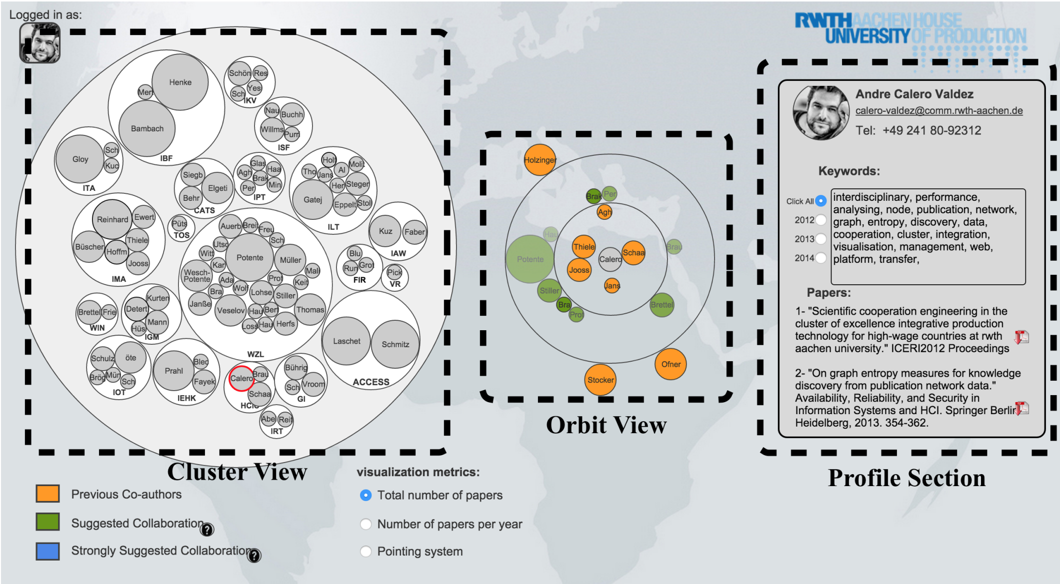


Figure 5.12: Final proposed visualisation: from left to right, Cluster view, Orbit view and profile section

Cluster View

Authors were represented as bubbles. Institutes were represented as bubble bags, containing all authors from the respective institute. Bubble size was determined by publication output and increased linearly with increasing numbers of publications. The location of bubbles in every bubble bag was randomised. The position of the each author is fixed to a relative location by using the name as a hash for its positioning within the relevant institute. Institute bubbles contained the acronym of the institute. The logged-in user was depicted as a red circle to allow fast self-orientation.

Orbit View

As the name suggests, this view contained three orbits. At the centre was the user himself and then the inner circle represented previous co-authors. The middle orbit included suggested collaborators and the outer orbit contained suggestions and previous collaborators who have left our cluster.

Using a force-directed algorithm with homogenous tension, the nodes attracted each other if they were from a similar institute. To avoid distortion in the representation of orbits, they were all fixed to the boundary of their own orbit. The tension of force applied was always similar. Except for the central node, the spatial location of nodes was randomised. Every orbit maintained a minimum distance from its inner orbit to avoid overlaps.

Profile Section

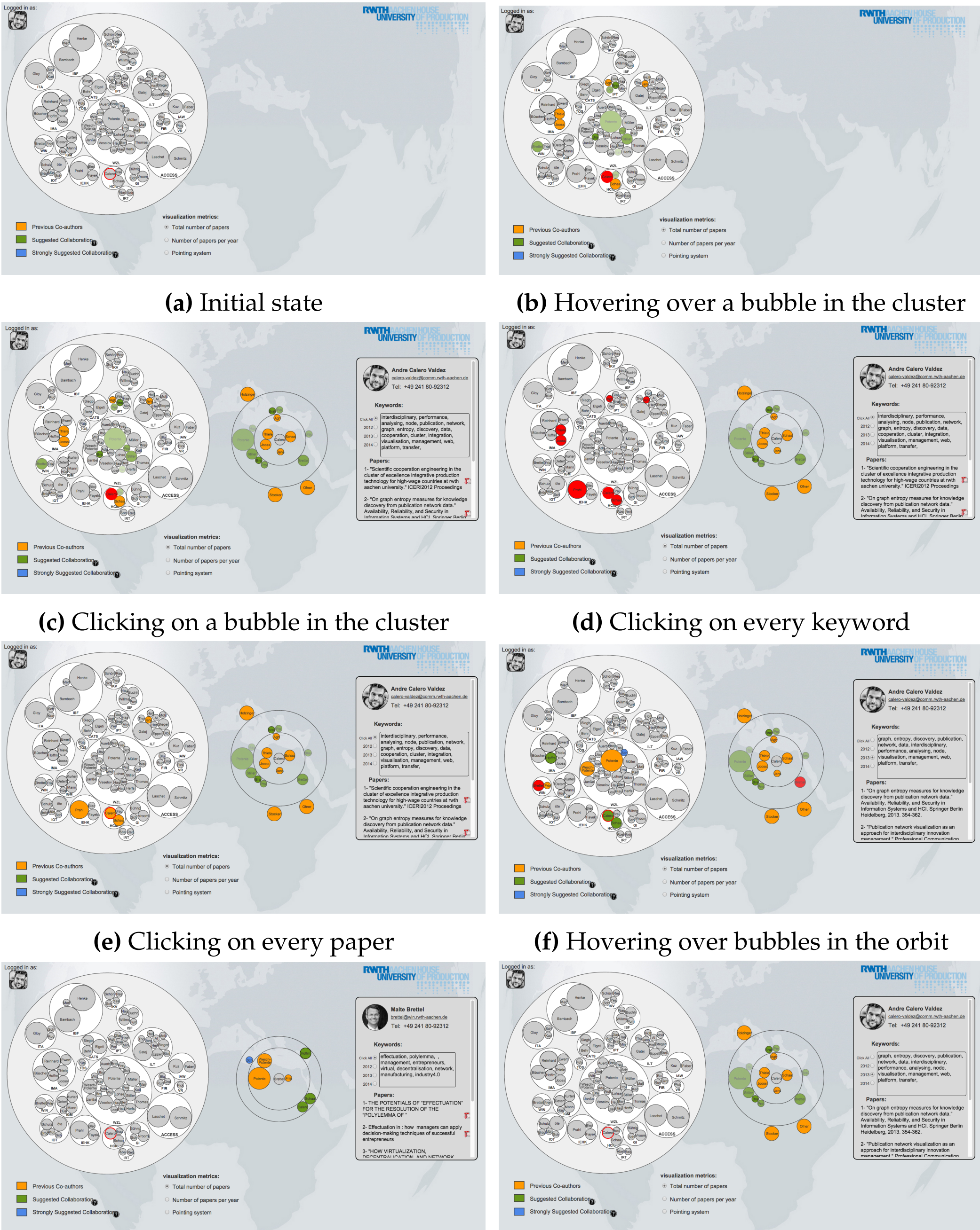
Every user had their own profile section containing a user profile photo, contact information, keywords and papers with respective year. Keywords were extracted from paper titles. The papers segment contained every paper's full outlet and a download button. Radio boxes were used to filter keywords and papers with respect to year. Every profile section was activated by clicking on the respective bubble representation in the visualisation.

5.5.2 Interaction and Navigation

We used a scenario-based description of our system for better clarification of context of use.

Logged-in State:
This visualisation was designed to be used inside a social networking portal. Figure 5.13a shows the initial state. The other views were disabled initially and then by clicking on any bubble it activated the associated orbit view and profile section.

Hovering Over a Bubble in Cluster View:
Hovering over every bubble highlighted that bubble in red and this indicated its connections throughout network (see Figure 5.13b). Orange was used to illustrate previous co-authors, green and blue represented suggested collabora-



(a) Initial state

(b) Hovering over a bubble in the cluster

(c) Clicking on a bubble in the cluster

(d) Clicking on every keyword

(e) Clicking on every paper

(f) Hovering over bubbles in the orbit

(g) Clicking on a bubble in the orbit

(h) Filtering papers and keywords

Figure 5.13: Navigation Through Final Prototype.

tors based on the level of suggestion strength. The brightness of colours changed with respect to the number of similar keywords with the logged-in user's keywords. Higher similarities were proportional to darker colours and vice versa.

Clicking on a Bubble in Cluster View:
The orbit view and the associated profile section were activated by clicking on a bubble in the cluster view. This reduced cognitive load when evaluating suggested collaborators (see Figure 5.13c).

Clicking on Every Keyword in Profile Section:
In Figure 5.13d it is illustrated that on clicking any keyword in the profile section, bubbles with similar keywords were highlighted in red.

Clicking on Every Paper in Profile Section:
Clicking on every paper outlet highlighted the co-authors of that paper in cluster view (see Figure 5.13e). To download that paper, a small download icon was used.

Hovering Over Bubbles in Orbit View:
To discover more, hovering over the suggested bubble in orbit view indicated their connections and suggestions in cluster view (see Figure 5.13f). The same could be done in both the cluster view as well as the orbit view.

Clicking on Bubbles in Orbit View:
Clicking on a bubble in the orbit view, switched between the respective bubble's orbit views. It also activated the respective bubble's profile section (see Figure 5.13g).

Filter Function:
Every profile section came with a year filter. Selecting on a year filtered keywords and papers with respect to that year(see Figure 5.13h). Initially, all accumulated keywords and papers were shown.

Legend and Metrics:
The legend helped users to understand the colour coding. This increased trust in our suggestion algorithms. Metrics were introduced to give users opportunity to change

visualisation structure with respect to the selected metric. However, validity of the proposed metrics has to be investigated.

Chapter 6

Summary and Future Work

In this chapter we summarise our findings and discuss our contribution to the field. We then explain possible future works and the limitations that we faced during this project.

6.1 Summary and Contribution

Visualisations and suggestions for collaboration in interdisciplinary settings is a complex task. Initially we studied the field of interdisciplinary collaboration and investigated basic needs and the factors that hinder collaboration. Furthermore, we conducted a literature review of related works. Ultimately we were able to produce the most acceptable and applicable visualisation for our user groups.

Target user groups were satisfied with our visualisation.

With respect to our literature review, scientific social networks and analytic sites like ResearchGate, Academia.edu, ArnetMinter, ResearcherId [Yu et al., 2007, Chaiwanarom et al., 2010] address finding collaborators using researcher profiles. Nonetheless, they do not address the task of finding collaborators with a specialised visualisation. Moreover, suggestions of collaborators were not based on prior collaboration but only on shared research interests.

In this thesis we presented a design study of a visualisation tool to convey organisational structure as well as valid possible collaborators in interdisciplinary settings. It allowed users to freely lay out the information needed to develop successful co-action. Our studies provided quantitative and qualitative evaluations for the user interface and functionalities. We started with user requirements and a context of use study, this assisted us in developing multiple low fidelity paper prototypes. Low fidelity prototypes were not sophisticated enough to achieve our goals, therefore we moved on to high fidelity prototypes using real data. Our user studies provided enough evidence that users not only lack such visualisation tools, but also such tools assist them in discovering new information about possible collaborators.

Using real data in prototyping produced impactful feedback.

With reference to Munzner [2008], the research contribution of this thesis is not a complex algorithm to discover collaborators but rather a justification of how existing techniques can be usefully combined with a visualisation to deliver users goals. However, our design study does not necessarily lead to a better quality of interdisciplinary research outcome. Yet, by employing this visualisation it will help the research community toward their goals and help break down barriers.

6.2 Future Work and Limitations

For our visualisations we performed both a requirements analysis and user studies in an iterative participatory design process. For future work we would like to include some more functionalities that were suggested in order to optimise user fit. As an example, we want to give users the ability to accept or reject a suggested collaborator after evaluation of their relevance. This feedback should be integrated into the recommendation algorithm. Furthermore, recommendations could be generated by using text-mining procedures instead of keyword analysis.

Still, more functionalities can be integrated into this tool.

Another example is to display the keyword similarities between the user and the suggested co-authors or allowing the user to modify associated keywords. By extending the scope to suggesting particular papers instead of authors, we could allow the user to judge the relative importance of a certain keyword for the researcher in question. Furthermore, the approach should be extended to include collaborators that have not published yet. This would require new researchers to fill in a profile indicating research interests using keywords. Due to the short research cycle in a Cluster of Excellence, a notification system will need to be integrated. Users would be informed about changes in their graph and this would be necessary for long-term use. However, we need to investigate the impact of our tool in the long run to address the success ratio (RQ5) of our visualisation.

RQ5 can be answered after long-term evaluation of our system.

A limitation of this work is that the specific sample came from just one research cluster. To generalise our approach we could map our visualisation to other contexts. The bubbles could also reflect institutes from an entire department or school in order to understand collaboration in a university as a whole. Whether the visualisation will effectively scale is yet to be answered, and whether the approach can be used in non-academic scenarios also warrants investigation.

Culture [Alagoz et al., 2010] and in particular differences in organisational or disciplinary culture should also be considered when visualising performance and thus sensitive data of employees. In our approach we assumed a relative homogenous user group. Since regional, organisational and disciplinary cultural differences can lead to a very heterogeneous user group, factors of user diversity must be considered when dealing with employee data [Schaar et al., 2013b].

We assumed a relative homogenous user group.

In addition, only titles of papers were used for the extraction of keywords. Using full texts or abstracts should reveal better keywords in the long run as would manual keyword selection by users. Furthermore, no disambiguation of keywords or synonym detection was applied. In interdisciplinary settings this is a particularly important requirement. Thus, in this regard our system does not help overcome disciplinary language barriers.

Synonym detection can resolve a major keyword disambiguation.

Finally, the sample for this study was relatively small (approx. 5% of the research cluster). For a better quantitative evaluation and better picture of system use, more participants should be considered over long-term evaluations [Barkhuus and Rode, 2007]. Publication data were only selected from 2012 to early 2014, limiting the insights from senior researchers and very recent publications.

Appendix A

Appendix: Forms and Interview Questions

The following questionaries were used for requirement study, first and second user studies.

Dear Mr/s.......

The topic of the interview is "Publication visualisation in an interdisciplinary workflow as a tool for self-evaluation and analysing interdisciplinary team success". What this means is, I would like to ask you certain questions about your process of daily work, when you publish, so that I can find better ways of publication visualisations, and This interview should help me to find the initial requirements of my visualisation tool in an interdisciplinary setting.

The interview would be recorded with your permission. However, we can stop the interview at any time, just let me know. This interview will remain anonymous.

This interview has 3 parts,

1st part: Your background knowledge
2nd part: Your process of publication.
3rd part : Your cooperation within cluster.

1st Part : Background knowledge

1. Can you introduce yourself in brief (your role at cluster, what you do)?
2. What are your expertise?
3. What are your skills that you want others to know?
4. How do you measure yourself in terms of Experience in research work (in publishing in cluster, very less-very much)?
5. Do you usually do disciplinary or interdisciplinary research?
6. What disciplines you work with?
7. what sort of Hardware/Softwares you use to produce a interdisciplinary work ?(Latex, MS Word , library management, Matlab , large screens , tablets, etc.)
8. What do you think can be achieved by interdisciplinary research?

2nd Part: Process of Publication

1. How many publication you had so far ? How many per year?
2. How often you produce a paper? Whether your publish within cluster of excellence or outside?

Figure A.1: Interview questions of user requirement study 1/2

3. How long does it takes you to publish a paper?

4. How much time do you usually spend for each interdisciplinary publication? (Compare to single disciplinary)

5. What is the process when u are looking for new literature(new topic)? (research methods)

6. Do you first start searching for a topic or do you start by collaborating with a person ?

7. How do you know that you have achieved enough information to write a paper? (background knowledge)

8. Is there a common (theme/relation/connection) among your publications ? if so, what are they ? (Same topic, same technology)

9. If someone (a non-expert) were to look for an expert in your "theme" what would they have to look for (Keywords)? (Investigating, if the term used is different from an expert and a non-expert view)

10. Do you have a favourite publication that you want people to find? How do you know it's your favourite (quality, citation, journal, etc.)

11. What do you find frustrating during the process of interdisciplinary publication ?

3rd Part: Cooperation in cluster

1. Do you have regular co-author(s)?

2. Have you ever had a need to find an author for your publication? If so, How do you find one? if no, why?

3. Is there a way to identify the willingness of people in cluster to write a publication for specific topic with you? (What are the aspects that you would like to know before starting a collaboration?)

4. Do you feel any sort of competition with your co-authors?, Is it different when they come form different disciplines ?

5. What is impact of your research (in Cluster of Excellence, industry or academia)? How do you measure your impact ?
6. What do you wish for during the process of publishing in interdisciplinary outlets?

Figure A.2: Interview questions of user requirement study 2/2

Hello Mr/s.......
My name is Amin Yazdi
The scope of this user study is to create a visualisation for collaboration in large-scale organisations. For this, we have created two prototypes and I would like you, to test them for me. We have obtained database of publications in Cluster of Excellence from 2012 till early 2014. Our visualisations were created based on these data. For the purpose of collaborator suggestion we have extracted keywords from titles of papers.

In order to evaluate the interview, I would like to record the audio and the computer screen for later analysis of your interactions. The data will not be shared and will never connect to your name. This interview will be anonymous.

This interview is semi-structured. There will be few tasks to perform and then you can interact with the system freely. Kindly think-aloud throughout this interview so that I can know the reason for your actions. In this user study we would like to evaluate two visualisations separately.

On your request we can stop the interview at any time.

Cluster View Prototype:

1. Please tell me what do you see ? (First sight)
2. Please locate yourself in this Visualisation.
3. Please hover over yourself and tell me what king of information you get ?
4. Can you guess why the person is suggested for you?
5. How do you rank the suggested co-authors ? Which one would be the first to contact ? why?
6. If you are searching for new co-author, What would you do next after spotting someone on Vis. ? (Calling to friend, emailing , double checking with others)
7. Does it happen for you to write paper with Profs. of other institutes directly?
8. Is there anybody that you have planned to have paper with before, that you can find him here ?
9. Is there any more information that you would like to add?

Figure A.3: First user study interview questions 1/2

Orbit View Prototype:

1. Please tell me what do you see ? (First sight)
2. Please locate yourself in this Vis.
3. Please hover over yourself and tell me what king of information you get ?
4. You can now freely interact with the visualisation.
5. What information do you get? (What are new informations?)
6. Is there any more information that you would like to have?

Figure A.4: First user study interview questions 2/2

Hello Mr/s.......
My name is Amin Yazdi
The scope of this user study is to create a visualisation for collaboration in large-scale organisations. For this, we have created a prototype and I would like you, to test it for me. We have obtained database of publications in Cluster of Excellence from 2012 till early 2014. Our visualisations were created based on these data. For the purpose of collaborator suggestion we have extracted keywords from titles of papers.

In order to evaluate the interview, I would like to record the audio and the computer screen for later analysis of your interactions. The data will not be shared and will never connect to your name. This interview will be anonymous.

This interview is semi-structured. There will be few tasks to perform and then you can interact with the system freely. Kindly think-aloud throughout this interview so that I can know the reason for your actions.

On your request we can stop the interview at any time.

1. Locate yourself and click on your bubble. Please tell me what do you see ? (First sight)
2. Now without any further interaction, can you describe what do you perceive from middle diagram ? (distance, layers, etc..)
3. Please find Who else has similar keyword as you for the word "Cluster".
4. In which year Mr/s........ has published "*Title of a paper*" ?
5. Please find and download paper "*Title of a paper*" from Mr/s ....
6. Please identify co-authors of Mr/s. .... for the paper "*Title of a paper*".
7. Who are co-authors of Mr/s. .... ?
8. Can you guess why that person is suggested to you?
9. How does this Visualisation can help you beside finding co-authors.
10. Is there any more information that you would like to have or add?

Figure A.5: Second user study interview questions

Usability Questionary

Visualisation of Collaboration in Large-scaled Organisations.

* Required

I think that I would like to use this system frequently *

1 2 3 4 5

Strongly disagree Strongly agree

I found the system unnecessarily complex *

1 2 3 4 5

Strongly disagree Strongly agree

I thought the system was easy to use *

1 2 3 4 5

Strongly disagree Strongly agree

I think that I would need the support of a technical person to be able to use this system *

1 2 3 4 5

Strongly disagree Strongly agree

I found the various functions in this system were well integrated *

1 2 3 4 5

Strongly disagree Strongly agree

I thought there was too much inconsistency in this system *

1 2 3 4 5

Strongly disagree Strongly agree

I would imagine that most people would learn to use this system very quickly *

1 2 3 4 5

Figure A.6: SUS and NPS questionaries used for first and second user studies. 1/2

Strongly disagree Strongly agree

I found the system very complex to use *

1 2 3 4 5

Strongly disagree Strongly agree

I felt very confident using the system *

1 2 3 4 5

Strongly disagree Strongly agree

I needed to learn a lot of things before I could get going with this system *

1 2 3 4 5

Strongly disagree Strongly agree

How likely is it that you would recommend this system to a friend or colleague? *

0 1 2 3 4 5 6 7 8 9 10

Extremely Unlikely Extremely Likely

Submit

Powered by

This content is neither created nor endorsed by Google.

Report Abuse - Terms of Service - Additional Terms

Figure A.7: SUS and NPS questionaries used for first and second user studies. 2/2

Bibliography

Firat Alagoz, André Calero Valdez, Wiktoria Wilkowska, Martina Ziefle, Stefan Dorner, and Andreas Holzinger. From cloud computing to mobile internet, from user focus to culture and hedonism: the crucible of mobile health care and wellness applications. In *Pervasive Computing and Applications (ICPCA), 2010 5th International Conference on*, pages 38–45. IEEE, 2010.

Louise Barkhuus and Jennifer A Rode. From mice to men-24 years of evaluation in chi. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007.

Mathieu Bastian, Sebastien Heymann, Mathieu Jacomy, et al. Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 8:361–362, 2009.

Danah M Boyd. Friendster and publicly articulated social networking. 2004.

Christian Brecher, Sabina Jeschke, Günther Schuh, Susanne Aghassi, Jens Arnoscht, Fabian Bauhoff, Sascha Fuchs, Claudia Jooß, Wilhelm Oliver Karmann, Stefan Kozielski, et al. Integrative production technology for high-wage countries. *Integrative Production Technology for High-Wage Countries*, pages 17–76, 2012.

Garry D Brewer. The challenges of interdisciplinarity. *Policy Sciences*, 32(4):327–337, 1999.

John Brooke. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

André Calero Valdez, Anne Kathrin Schaar, Martina Ziefle, Andreas Holzinger, Sabina Jeschke, and Christian

Brecher. Using mixed node publication network graphs for analyzing success in interdisciplinary teams. *Active Media Technology*, pages 606–617, 2012.

André Calero Valdez, Anne Kathrin Schaar, Martina Ziefle, and Andreas Holzinger. Enhancing interdisciplinary cooperation by social platforms. In Sakae Yamamoto, editor, *Human Interface and the Management of Information. Information and Knowledge Design and Evaluation*, volume 8521 of *Lecture Notes in Computer Science*, pages 298–309. Springer International Publishing, 2014.

Paweena Chaiwanarom, Ryutaro Ichise, and Chidchanok Lursinsap. Finding potential research collaborators in four degrees of separation. In *Advanced Data Mining and Applications*, pages 399–410. Springer, 2010.

Jonathon N Cummings and Sara Kiesler. Who collaborates successfully?: prior experience reduces collaboration barriers in distributed interdisciplinary research. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 437–446. ACM, 2008.

Jean-Daniel Fekete and Catherine Plaisant. Interactive information visualization of a million items. In *Information Visualization, 2002. INFOVIS 2002. IEEE Symposium on*, pages 117–124. IEEE, 2002.

Ingbert R Floyd, M Cameron Jones, Dinesh Rathi, and Michael B Twidale. Web mash-ups and patchwork prototyping: User-driven technological innovation with web 2.0 and open source software. In *System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on*, pages 86–86. IEEE, 2007.

Eugene Garfield. Citation indexes for science. a new dimension in documentation through association of ideas. *International journal of epidemiology*, 35(5):1123–1127, 2006.

Anne-Wil Harzing and Ron Van der Wal. Google scholar: the democratization of citation analysis. *Ethics in science and environmental politics*, 8(1):61–73, 2007.

Orit Hazzan, Yael Dubinsky, Larisa Eidelman, Victoria Sakhnini, and Mariana Teif. Qualitative research in computer science education. *ACM SIGCSE Bulletin*, 38(1): 408–412, 2006.

Jeffrey Heer and Danah Boyd. Vizster: Visualizing online social networks. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*, pages 32–39. IEEE, 2005.

Florian Heller and Jan Borchers. Powersocket: towards on-outlet power consumption visualization. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, pages 1981–1986. ACM, 2011.

Danny Holten. Hierarchical edge bundles: Visualization of adjacency relations in hierarchical data. *Visualization and Computer Graphics, IEEE Transactions on*, 12(5):741–748, 2006.

Jerry A Jacobs and Scott Frickel. Interdisciplinarity: A critical assessment. *Annual Review of Sociology*, 35:43–65, 2009.

Jeff Johnson. Participatory design of computer systems (panel). In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 141–144. ACM, 1990.

Claudia Jooß, Florian Welter, Ingo Leisten, Anja Richert, A Schaar, A Calero Valdez, Eva-Maria Nick, Ulrich Prahl, Ulrich Jansen, Wolfgang Schulz, et al. Scientific cooperation engineering in the cluster of excellence integrative production technology for high-wage countries at rwth aachen university. *ICERI2012 Proceedings*, pages 3842–3846, 2012.

Bonnie Kaplan and Dennis Duchon. Combining qualitative and quantitative methods in information systems research: a case study. *MIS quarterly*, pages 571–586, 1988.

Lian Chee Koh, Aidan Slingsby, Jason Dykes, and Tin Seong Kam. Developing and applying a user-centered model for the design and implementation of information visualization tools. In *Information Visualisation (IV), 2011 15th International Conference on*, pages 90–95. IEEE, 2011.

Hildrun Kretschmer. A new model of scientific colloboration part 1. theoretical approach. *Scientometrics*, 46(3), 1999.

Heidi Lam and Tamara Munzner. Increasing the utility of quantitative empirical studies for meta-analysis. In

Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization, page 2. ACM, 2008.

Mariella Marzano, David N Carss, and Sandra Bell. Working to make interdisciplinarity work: investing in communication and interpersonal relationships. *Journal of Agricultural Economics*, 57(2):185–197, 2006.

Philipp Mayring. *Qualitative inhaltsanalyse*. 2011.

Lisa J Miller, Rich Gazan, and Susanne Still. Unsupervised classification and visualization of unstructured text for the support of interdisciplinary collaboration. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1033–1042. ACM, 2014.

Tamara Munzner. Process and pitfalls in writing information visualization research papers. In *Information visualization*, pages 134–153. Springer, 2008.

Mark EJ Newman. Fast algorithm for detecting community structure in networks. *Physical review E*, 69(6):066133, 2004.

Jakob Nielsen. Usability inspection methods. In *Conference companion on Human factors in computing systems*, pages 413–414. ACM, 1994a.

Jakob Nielsen. *Usability engineering*. Elsevier, 1994b.

Jakob Nielsen and Rolf Molich. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256. ACM, 1990.

Sidney Redner. How popular is your paper? an empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, 4(2):131–134, 1998.

Frederick F Reichheld. The one number you need to grow. *Harvard business review*, 81(12):46–55, 2003.

Allen F Repko. Interdisciplinary research: Process and theory. 2011.

Markus Rester, Margit Pohl, Sylvia Wiltner, Klaus Hinum, Silvia Miksch, Christian Popow, and Susanne Ohmann. Mixing evaluation methods for assessing the utility of an interactive infovis technique. In *Human-Computer Interaction. Interaction Design and Usability*, pages 604–613. Springer, 2007.

Anne Kathrin Schaar, André Calero Valdez, and Martina Ziefle. Publication network visualization as an approach for interdisciplinary innovation management. In *Professional Communication Conference (IPCC), 2013 IEEE International*, pages 1–8. IEEE, 2013a.

Anne Kathrin Schaar, André Calero Valdez, and Martina Ziefle. The impact of user diversity on the willingness to disclose personal information in social network services. In *Human Factors in Computing and Informatics*, pages 174–193. Springer, 2013b.

Michael Sedlmair, Miriah Meyer, and Tamara Munzner. Design study methodology: Reflections from the trenches and the stacks. *Visualization and Computer Graphics, IEEE Transactions on*, 18(12):2431–2440, 2012.

Burr Settles and Steven Dow. Let's Get Together : The Formation and Success of Online Creative Collaborations. In *Proc. CHI 2013*, pages 2009–2018, 2013. ISBN 9781450318990. doi: 10.1145/2470654.2466266.

Ben Shneiderman. Extreme visualization: squeezing a billion records into a million pixels. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pages 3–12. ACM, 2008.

Ben Shneiderman and Catherine Plaisant. Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. In *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*, pages 1–7. ACM, 2006.

Aidan Slingsby and Jason Dykes. Experiences in involving analysts in visualisation design. In *Proceedings of the 2012 BELIV Workshop: Beyond Time and Errors-Novel Evaluation Methods for Visualization*, page 1. ACM, 2012.

Terry A Slocum, Daniel C Cliburn, Johannes J Feddema, and James R Miller. Evaluating the usability of a tool for visualizing the uncertainty of the future global water balance. *Cartography and Geographic Information Science*, 30(4):299–317, 2003.

Carolyn Snyder. *Paper prototyping: The fast and easy way to design and refine user interfaces*. Newnes, 2003.

Melanie Tory and Sheryl Staub-French. Qualitative analysis of visualization: a building design field study. In *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization*, page 7. ACM, 2008.

Nayuko Watanabe, Motoi Washida, and Takeo Igarashi. Bubble clusters: an interface for manipulating spatial aggregation of graphical objects. In *Proceedings of the 20th annual ACM symposium on User interface software and technology*, pages 173–182. ACM, 2007.

Mohammad Amin Yazdi, André Calero Valdez, Leonhard Lichtschlag, Martina Ziefle, and Jan Borchers. Visualizing opportunities of collaboration in large organizations. In *Submitted for publication at INTERACT '15: International Conference on Human-Computer Interaction*, 2015.

Wei Yu, Ajay Yesupriya, Anja Wulf, Junfeng Qu, Muin J Khoury, and Marta Gwinn. An open source infrastructure for managing knowledge and finding potential collaborators in a domain-specific subset of pubmed, with an example from human genome epidemiology. *BMC bioinformatics*, 8(1):436, 2007.

Index

Typeset March 8, 2015