

# FacetSearch: A Faceted Information Search and Exploration Prototype

Christin Katharina Kreutz<sup>(✉)</sup> , Peter Boesten, Alex Witry, and  
Ralf Schenkel<sup>(✉)</sup> 

Trier University, 54286 Trier, DE  
{krechtzch, s4peboes, s4alwitr, schenkel}@uni-trier.de

**Abstract.** Faceted search and exploration is effective and desirable when confronted with bibliographic data sets. As this data is mostly extensive, a powerful and flexible information retrieval tool needs to be provided. We propose FacetSearch; a scalable prototype especially designed for 1) comparing and visualizing trends in corpora via a trend visualization, as well as 2) exploring data and relations between elements with a pivot operation. Metadata can be accessed in multiple dimensions with facets and filter rules on them to help refine search queries or navigate through a corpus. An evaluation showed the potential and user acceptance of our approach when compared to the textual search engine Semantic Scholar.

**Keywords:** Faceted Search · Data Exploration · Faceted Metadata · Information Visualization

## 1 Introduction

With the ever-growing amount and richness of bibliographic data, exploring and interpreting this information with help of tools is becoming increasingly important. A simple textual search is not sufficient to access the collection in its full dimension if a user has no prior knowledge of the data being looked for. Faceted search can help overcome this by providing attributes with which queries can be refined or a corpus can be explored. Some works even point out, that users prefer searching with facets instead of only textual search [24] and exploratory visualization systems improve the users understanding of collections [11].

When comparing applicants for an open position in science, multiple aspects of researchers need to be considered. Not only their number of publications but also their h-index, their domain, the conferences and journals they published in and their publication development over the years could be relevant. In literature research, examining quality and impact of citations as well as getting a brief insight into papers is essential to estimate their usefulness and centrality. For scientists exploring a corpus is important to detect interesting researchers and publications associated with certain topics. All these tasks require a powerful information retrieval system. We introduce a prototype tailored to tackle such information needs: FacetSearch, a system especially directed at visualizing and

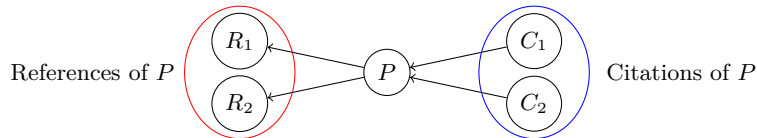


Fig. 1: Nodes symbolize publications, edges between papers symbolize citations.  $C_1$  and  $C_2$  are citations of  $P$ ,  $R_1$  and  $R_2$  are references of  $P$ .

comparing trends and relations in large bibliographic collections as well as giving the opportunity to explore them by usage of faceted search.

This work is structured as follows: in Section two, we introduce some basic definitions and our data set. Afterwards, the FacetSearch interface is described in detail and evaluated in the following chapter. We close with related work and the description of possible improvements for our preliminary system.

## 2 Terms and Definitions

### 2.1 Definition

There are different information needs when working with an information system. A simple *search* is about refining a result set from many to few elements with the goal of finding the most relevant element for a query. *Exploration* is a much more imprecise task with no certain ambition except to casually examine an information space and navigate between elements.

Metadata can be clustered in alternative hierarchical categories in the information space. These categories are called *facets*. Facets contain *elements*, which are the aggregated potential values that occur in the corresponding metadata dimension. Queries on elements of facets are called *filter rules*, they generate and refine results. A *rule set* is composed of all currently active filter rules. *Faceted search* interfaces support exploration and search by suggesting fitting facet filter options based on the respective rule set to progressively specify the results.

In the context of a publication-centered faceted search system, facets could be *author names*, *journals* or *keywords attributed to publications*, elements in a facet containing keywords could be *database*, *machine learning* and *algorithm*. A filter rule on a facet containing keywords could query for *algorithm* and retrieve publications attributed with this keyword.

For our use case of bibliographic data, we need to specify linkages between publications. We distinguish between citations and references as visualized in Figure 1. *Citations* are the publications connected by ingoing links to a paper from its citing publications, *references* are the papers connected by outgoing links from a publication to papers which are cited.

### 2.2 Data Set

Our prototype uses data from the dblp computer science bibliography. The dblp corpus contains bibliographic metadata concerning authors, publications as well

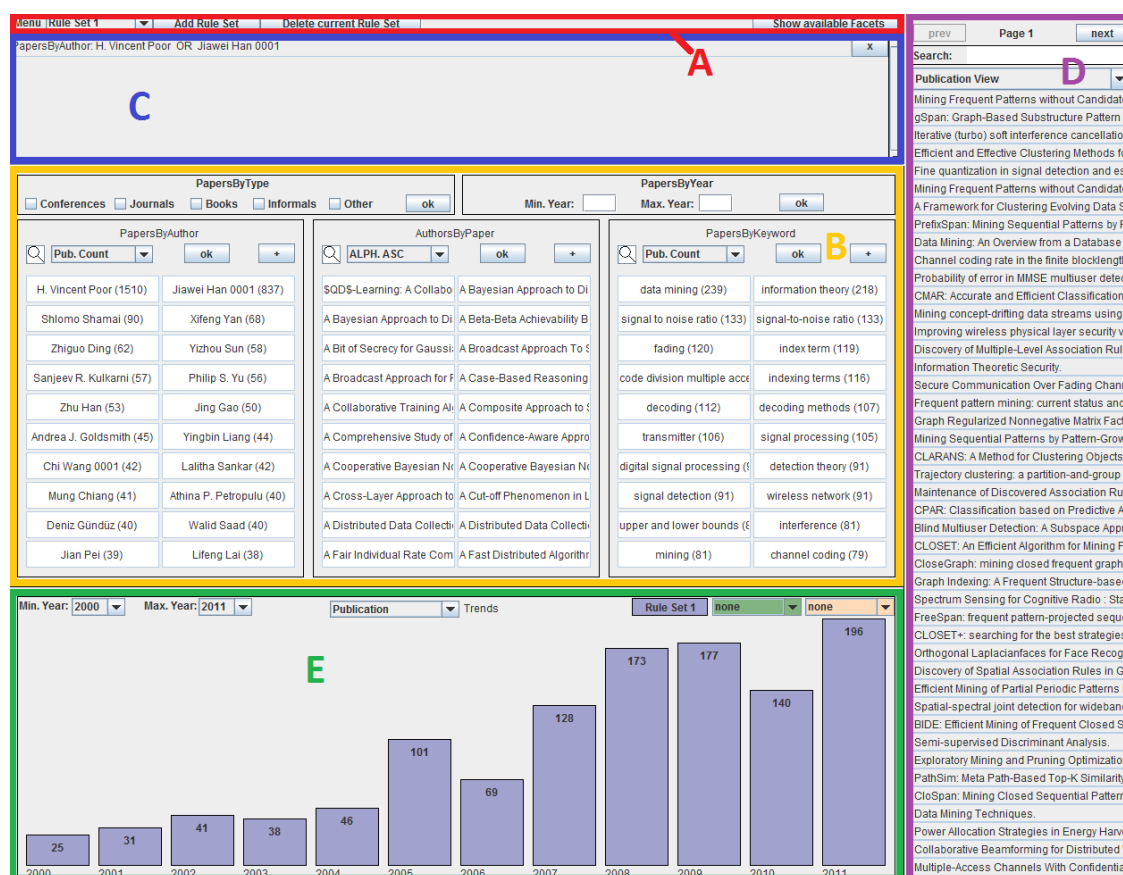


Fig. 2: FacetSearch interface with (A) menu, (B) facet area, (C) filter, (D) item view and (E) trend visualization.

as venues (journals and conferences) from the field of computer science and adjacent areas [10]. As of July 2018, it holds information of over 4.2 million publications and more than 2.1 million authors. This data set is enriched with information from Aminer [20] based on DOIs matched with dblp data or paper title and author matches where DOIs were not present. For about 100,000 papers, full texts extracted from PDFs with Science Parse [29] are included.

In dblp, author names, publication titles and related venues are stored, the extended collection additionally contains citations, references, abstracts, full texts, keywords and topics of publications. All supplementary data is partial. Except for abstracts and full text, available metadata is used as elements in facets. Topics are merged into one facet with keywords as these attributes are similar but often missing. The additional data is useful in indicating quality of papers (citations and references) as well as providing further aspects to explore (abstracts, full texts) and redefine (keywords) the data.

- |                         |                               |                                 |
|-------------------------|-------------------------------|---------------------------------|
| - Papers By Author      | - Cited Papers By Author      | - Authors Cited By Author       |
| - Authors By Paper      | - Referenced Papers By Author | - Papers Referencing Paper      |
| - Papers By Collection  | - Papers Cited By Paper       | - Papers Referencing Collection |
| - Authors By Collection | - Papers Cited By Collection  | - Authors Referencing Author    |
| - Papers By Keyword     |                               |                                 |

Fig. 3: Facets available in FacetSearch.

### 3 The FacetSearch Interface

The FacetSearch interface can be partitioned into five main components as seen in Figure 2: At the top (A), a *menu* is given with which a user can switch rule sets and select which facets are being displayed. In the *facet area* (B), currently chosen facets are depicted to explore the data with. Above in the *filter* (C), the current rule set is given. On the right in the *item view* (D), elements retrieved by the query donated by the active rule set are listed. Lastly, in the *trend visualization* (E), the temporal development of elements from rule sets is pictured.

The arrangement of areas in the interface is highly influenced by FacetLens [9]. We designed the user interface according to Shneiderman et al.'s principle of direct manipulation, where the visual representation of data, physical interaction with objects and the reversibility of actions are key elements [16].

#### 3.1 Program Components

Facets are named according to the convention *X By Y*. *Y* is the type of metadata or elements displayed in the facet, *X* is the type of results of a filter rule on this facet. Facets working on citations are an exception; they are structured as *X Cited By Y* or *X Referencing Y*. Elements in facets are dynamically queried and displayed. Figure 3 gives an overview of facets available in FacetSearch, the repertoire of facets is based on previous work [9]. In our prototype, tool tips provide short descriptions of the facets. To switch between visible facets, a menu can be accessed to pick the three displayed ones. Via search and different sorting methods on facet elements or iteration through them, elements can be selected and combined with conjunction or disjunction as new filter rules. Special focus concerning sorting options needs to be laid on facet *Papers By Keyword* as there is the possibility to sort elements according to the average percental growth of the keyword's occurrences in papers per year.

The facet *Papers By Author* contains all author names as elements. If the elements *Ralf Schenkel* and *Gerhard Weikum* are selected and conjugated in a filter rule, a query for papers which are co-authored by Ralf Schenkel and Gerhard Weikum is initiated and publications fitting all requirements of the current rule set are retrieved and listed in the item view. The elements in all facets are updated according to the restricted result set.

In addition to the traditional facets, two facets *Papers By Year* and *Papers By Type* are always visible, they do not contain dynamically generated elements. These facets refine the time interval and the desired publication type of results.

Users can switch between multiple rule sets. Filter rules contained in them are conjugated and resulting entries are listed in the item view. It can be switched

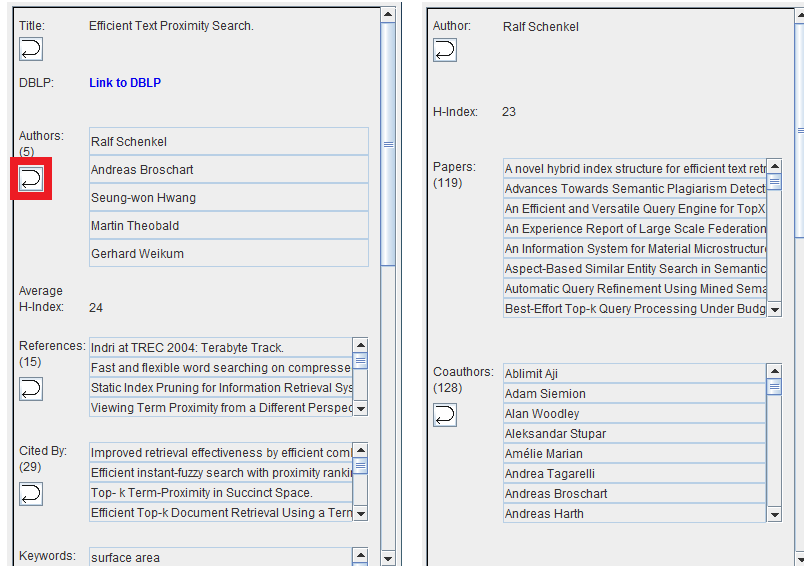


Fig. 4: On the left: Detail view for publication *Efficient Text Proximity Search*. On the right: Detail view for author *Ralf Schenkel*. One of the buttons for the pivot operation is marked red.

between a publication and an author view of the results, the default view is denoted by the *X*-component of the facet, on which the last filter rule was generated. A search is available on the item view. If an element from the item view is selected, a *detail view* opens as pictured in Figure 4. It provides all available information for an author or a publication.

Next to some attributes in the detail view, a button with an arrow symbol (marked red on the left detail view in Figure 4) offers a shortcut-functionality of our prototype, the *pivot operation*. When clicking this button, a user breaks free from the current environment and creates a new rule set without dropping the old one. The only active filter rule in the new rule set contains the elements, which a user pivoted on. By pivoting, users can quickly explore the corpus.

In the case shown in the left of Figure 4, clicking on the pivot button next to the five authors (Ralf Schenkel, Andreas Broschart, Seung-won Hwang, Martin Theobald and Gerhard Weikum) of the publication *Efficient Text Proximity Search* generates a new rule set. A user has to decide whether the authors are combined conjunctive or disjunctive. In the new rule set, the only active filter rule would be the selection of the five authors as elements from the facet *Papers By Author*. The item view then holds all publications fitting these requirements.

The development of publications over time from different rule sets can be visualized in the *trend visualization*. With two drop-down menus, the depicted interval can be determined. There are several trend types available for display: number of publications, number of citations and number of references. With a click on a trend bar, a window with publications fitting the rule set from the

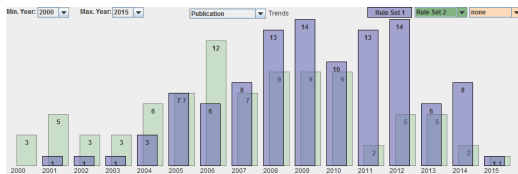


Fig. 5: Comparison of publications for researchers  $R_1$  (blue) and  $R_2$  (green) for keywords *database*, *information retrieval* and *data mining*. Height of bars indicates number of publications per year.

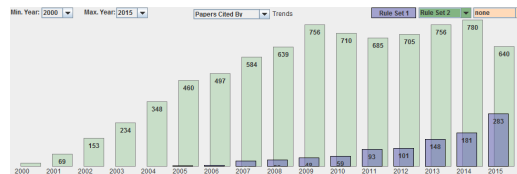


Fig. 6: Comparison of citations of publications from researchers  $R_1$  (blue) and  $R_2$  (green) for publications with keywords *database*, *information retrieval* and *data mining*. Height of bars gives number of citations per year

corresponding year opens. Trends of up to three rule sets can be compared in the trend visualization at the same time.

The trend visualization can be used for the estimation of keywords' relevance. If the keyword *logic programming* is chosen in facet *Papers By Keyword*, the trend visualization shows the decreasing number of publications attributed with this term which could hint at a vanishing importance of the topic. The selection of the keyword *smartphone* instead set could show an increasing number of publications over the years which could indicate the growing significance of this area at a current time frame.

### 3.2 Exemplary Application

One of the tasks FacetSearch is suited for is the comparison of applicants for vacancies in academia. Manning these positions is challenging, as only experts in the vacant field can fully access quality and impact of candidates research. Let's assume there are two researchers  $R_1$  and  $R_2$  who run for a professor position in database systems. In a first step, the number of publications of each applicant could be compared in this area which can be described by keywords *database*, *information retrieval* and *data mining*. When searching for publications from these authors (via facet *Papers By Author*) with OR-ed keywords (via facet *Papers By Keyword*), the comparison of results in the trend visualization is depicted in Figure 5. Based on these publication trends  $R_1$  would be the better candidate as he publishes more papers. Figure 6 shows the outcome if the number of citations of the publications of the authors is compared. Even though  $R_1$  did publish more than  $R_2$ ,  $R_2$  is cited much more often which indicates that  $R_2$  is more important than  $R_1$  and would be a more suitable population for the open position if no other aspects influence this decision.

### 3.3 Implementation Details

FacetSearch is a prototype implemented in Java 7 using local indices which store bibliographic information. Elasticsearch is used for saving and searching data.

- Q1 Who are the co-authors of author  $A$  in his most cited publication? How many publications do they have together?
- Q2 What is the most frequent keyword associated with author  $A$  between years  $Y_1$  and  $Y_2$ ?
- Q3 How many publications has author  $A$  in collection  $C$ ?
- Q4 Which of the publications citing publication  $P$  is the most cited one?
- Q5 Compare collection  $C_1$  and  $C_2$ : which one contains more publications with keyword  $K$  between years  $Y_1$  and  $Y_2$ ?
- Q6 Compare authors  $A_1$  and  $A_2$ : who was cited more between years  $Y_1$  and  $Y_2$ ?

Fig. 7: Overview of types of questions used in the evaluation.

One index stores information concerning authors (unique name, possible further names, h-index), the other one contains data regarding publications (title, unique author names, publication year, collection in which it was published, keywords, full text, citations, references, average h-index of authors) as far as available.

In Elasticsearch, the maximum result size of a query is set to 10,000 elements by default. As we did not change that value, it is the top number of elements returned in facets where  $Y = Author$  or  $Y = Paper$  if a search is performed on them. A search on authors could query for the primary unique name or a further name. Fitting names are looked up in the author index. For every author with good scoring, the unique name is used in the publication index to test if the author has at least one paper matching the current rule set and if so, the author is considered further. Search terms with high frequency such as *John* or *the* should be avoided or rendered more precisely to prevent the case of an element existing and fitting the current rule set but not being retrieved.

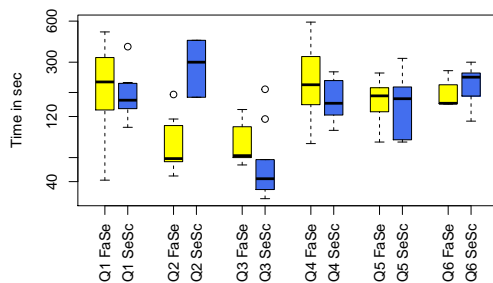
## 4 Evaluation

A direct comparison of FacetSearch with its predecessors PaperLens [8] as well as FacetLens [9] is impossible as these systems are unavailable online. In order to evaluate the usefulness of FacetSearch compared to a classical information system, we invited ten volunteers (two of them females, nine of them with a background in computer science) to participate in a user study. Our goal was to compare the usability and user acceptance of the systems by measuring speed of performance, error rate and the subjective satisfaction of users [16].

### 4.1 Experimental Set-Up

Our comparative study used FacetLens as a faceted information retrieval system and Semantic Scholar[27], an academic search engine used with textual search. Like FacetSearch, Semantic Scholar contains information on papers, authors, venues, citations, and keywords, indicating the target audience for both systems is similar.

Between participants, the order of the presented systems was alternated to compensate for learning effects. Test users were given tutorials of both systems before they started working on the respective tasks. A pre-evaluation with two different subjects showed the need for another training step between tutorial and evaluation with FacetSearch as the time to learn [16] with this system was



(a) Time per question answered correctly in seconds with mean of difficulty given by for FacetSearch and Semantic Scholar.

	FaSe			SeSc		
	✓	×	diff <sub>mean</sub>	✓	×	diff <sub>mean</sub>
Q1	8	2	3.5	9	1	4.1
Q2	9	1	2.4	2	8	4.3
Q3	9	1	2.4	9	1	3.2
Q4	8	2	4.9	6	4	4.9
Q5	9	1	3.2	8	2	3.4
Q6	3	7	3.9	7	3	4.4

(b) Number of right and wrong answers per system and question  
answers per system and question  
test users.

Fig. 8: Comparison of evaluation results from FacetSearch (FaSe) and Semantic Scholar (SeSc).

primarily underestimated. Two basic training tasks were included before users worked on the evaluated questions. Participants solved these in training with help of an instructor. The whole session lasted about one hour for each member of the evaluation group.

Figure 7 gives an overview of the six types of questions posed for each of the systems, variables were changed between them. The first two questions were directed at correct handling of authors, the next two aimed at exposure with publications and the last tasks required usage of the trend visualisation. Not only the correctness of responses was measured, answering time was also taken. Additionally, users were asked to rate the difficulty of solving the posed questions with the respective systems on a scale from 1 to 7 (1 being very easy, 7 being very difficult) and judge functionalities of FacetSearch in the end.

## 4.2 Results and Discussion

**FacetSearch** Some works observe a change in search habits when a grouping mechanism such as faceted search is available [4]. This corresponds with our observation of users requiring additional time and training before being able to fully grasp the functionality of FacetSearch in *Q1*. *Q2* and *Q3* were answered quickly and correctly by the test users. Because *Q4* uses references, this task was difficult for most participants as the required facet *Papers Referencing Paper* behaves unintuitively. Choosing elements from this facet leads to the display of publications in the item view, of which the chosen elements are a reference. The next question *Q5* was designed for users to look at the trend visualization but most of them chose another path to determine the right answer. With the last task *Q6*, a lot of test users made the mistake of limiting the observed years in the wrong place or in the decision of the selected trend type. Our evaluation showed the adequacy of the prototype for the different tasks except for *Q6* where a better distinction of restriction of years needs to be found.



Question	Avg. Rating
How intuitive was working with FacetSearch? ( <i>1 - not intuitive to 7 - very intuitive</i> )	3.2
How did you like the implementation of search on facets? ( <i>1 - very bad to 7 - very good</i> )	5.2
How did you like the implementation and usage of the trend visualization? ( <i>1 - very bad to 7 - very good</i> )	5.1
How did you like searching for information using FacetSearch compared to Semantic Scholar? ( <i>1 - clearly worse to 7 - clearly better</i> )	5.0

Table 1: User feedback given in the evaluation.

**Semantic Scholar** With Semantic Scholar’s intuitive operational concept, most users had no difficulties in answering *Q1* and *Q3* correctly. The second task *Q2* was solved wrong by eight users because in this search engine, keywords are ordered without apparent criteria. Users were confronted with the challenge of being unable to sort citations by attribute in *Q4*. The next question *Q5* hardly posed a problem except in the search for a collection as it cannot be searched for directly in Semantic Scholar. The last task *Q6* caused some confusion, several users had problems in finding and comparing the citation diagrams.

**Comparison** In Figure 8, evaluation results from FacetSearch and Semantic Scholar are compared. Figure 8a gives an overview of time needed for correctly answering each of the six questions split by system. The plot indicates over all participants were faster when using Semantic Scholar. Only when solving *Q2* and *Q6*, FacetSearch was the faster option. Figure 8b quantifies the correct and incorrect answers as well as the subjective mean of difficulty for the questions. With FacetSearch, 76.67% of all tasks were finished correctly, with Semantic Scholar, only 68.33% of results were correct. The rating of subjective difficulty of solving the problems by test users lead to a median of 3 for FacetSearch while resulting in 4 for Semantic Scholar. Participants did consider using the faceted search prototype as easier when confronted with our exemplary tasks even though their completion took longer. This phenomenon of initial problems but successful and easy usage when confronting users with faceted search has also been observed in user studies as the users’ ability to use and understand a system evolves due to learning processes [17,11].

There are classes of tasks which cannot be solved by systems such as Semantic Scholar but are manageable with FacetSearch. Such tasks were not put in our evaluation as it would have resulted in failure for the text-based search engine. They include the following example problems: The best publication (e.g. with most citations) from a set of publications is hardly found as sorting options are missing. Exploration of publications from collections is difficult as no direct search is implemented for venues. Trends for referenced publications from an author or publication cannot be displayed as well as citation trends for more complex queries than single authors.

**User Feedback** The last part of the evaluation was a user feedback rating intuitiveness and overall implementation of certain aspects of FacetSearch. Table 1 gives an overview of the assessments.

Some participants of the evaluation were confused because facet elements are simultaneously used for input and output. Filter rules are generated from them

but when a filter rule is added, elements dynamically change to match the query result. This phenomenon is complicated and a reason to classify FacetSearch as an expert system. One user fittingly noted that the more complex requirements for a system become, the more complex its operation becomes.

Other users remarked they trusted the results of FacetSearch far more than the results of Semantic Scholar because of its inconsistent data when for example searching for co-authors in author profiles. The stable structures in FacetSearch were complimented by many, overall consent was that with a sufficient period of professional training, users would be much faster with our prototype.

## 5 Related Work

FacetSearch is an extension of PaperLens [8], an information retrieval system specialized on visualizing trends and relationships in bibliographical data, and the faceted classification system FacetLens [9], where facets are ordered in tiles and elements in them can be selected gradually to refine queries. Just as most visualization and exploration tools for bibliographical data, they do not scale well. Some tools even come with a limited number of elements per facet [26,7] whereas our prototype is scalable and has no restricted value range.

Tiles are used in other research to display distribution and overlap of searched terms in documents [5,14] but in our case with only a small amount of textual data and separated search on each area, such a representation would not be meaningful for the most part.

Daffodil [3] is an information system sitting on top of several digital libraries which supports its users visually by providing so-called stratagems, high-level services which offer the possibility of depth search in different, possibly inconsistent views. Stratagems are the equivalent to facets in our approach. Working with facets has the advantage of being able to avoid empty result sets while also assuring consistency in the different views of the displayed data, as only facet elements are displayed which are fitting to the current rule set.

Some works like FacetMap [18] and FacetedDBLP [1] support search and exploration on faceted data but only hardly provide the possibility to visualize trends and relations in them like our trend visualization does.

PaperQuest [12] focuses on the aspect of visualizing literature review research but does not cover exploration on authors or the intertwined network of venues as important factors in this process which we can access by pivoting. The CiteVis system [19] displays citations of papers in a grid-based layout where it highlights searched for publications, authors and keywords. In contrast to our approach, it is only suitable for small corpora and exploration on facets is unavailable. TIARA [22] can be used for displaying the temporal development of keyword clouds extracted from texts and offers a faceted exploration option. Unfortunately, the exploration is not as powerful and exhaustive as ours. Other research displays ego-centred networks to visualize and explore relations between authors, authors and venues, or keywords and venues in a temporal context but do not offer the possibility to define complex queries on papers [15].

Further systems such as PivotPaths [2], which visualizes the links between different facets as tripartite graphs and focuses on the exploration of connections, seem to scale better but confuse users when too many elements are displayed simultaneously. With Bibliography Explorer++ [25], where the exploration of publications and citations were the focal point, similar problems occurred in user studies. Although users were confused by the complexity of facets dealing with citations and references in our evaluation, FacetSearch’s clear structure prevents an overstraining of users with too much information.

## 6 Conclusion and Future Work

We presented FacetSearch, an information retrieval prototype designed for search and exploration of bibliographic data via facets as well as display and comparison of trends in result sets by usage of a trend visualization.

Our tool is able to keep up with text-based search engines such as Semantic Scholar according to our evaluation, although users thought of it as less intuitive (3.2). This observation is no surprise as there are a lot of features which need understanding and experience. Most persons coped well with usage of facets (5.2) and the trend visualization (5.1). Even compared to Semantic Scholar, our prototype was able to satisfy users (5.0). Nevertheless, FacetSearch is an expert system which needs a certain period of vocational adjustment.

Further research such as PaperQuest [12] or Bow Tie Academic Search [6] use colouring and shapes to present more information on citations while preventing an overload of the interface with numbers and thus confusing a user. These techniques could also be applicable for publications in the item or detail view. In PivotPaths [2] the visualization of connectedness of facets was valued by users; in our case, we could incorporate a graph representing the citations and references between collections and their associated keywords.

Another area for improvement would be the opportunity of incorporating Doc2Vec [13] to extend and diffuse the keyword search and present a broader selection of papers fitting searched terms as this search is rather unspecific and the possible information gain exceeds the cost of displaying unfitting results [21].

Potential for future enhancements lies in the field of collections: Visualizing the citation/reference linkage graph of collections could provide useful insights. Venue ratings from Eigenfactor [23] or the CORE Conference/Journals DB[28] could be included so citations can be rated with them.

## References

1. J. Diederich, W.-T. Balke: *FacetedDBLP - Navigational Access for Digital Libraries*. TCDL 4(2) (2008)
2. M. Dörk, N.H. Riche, G. Ramos, S. Dumais: *PivotPaths: Strolling through Faceted Information Spaces*. TVCG 18(2): 2709–2718 (2012)
3. N. Fuhr, C.-P. Klas, A. Schaefer, P. Mutschke: *Daffodil: An Integrated Desktop for Supporting High-Level Search Activities in Federated Digital Libraries*. ECDL 2002: 597–612

4. M. Hearst: *Clustering versus faceted categories for information exploration*. Commun. ACM 49(4): 59–61 (2006)
5. M. Hearst: *TileBars: Visualization of Term Distribution Information in Full Text Information Access*. CHI 1995: 59–66
6. T. Khazaei, O. Hoerber: *Supporting academic search tasks through citation visualization and exploration*. Int J Digit Libr 18(1): 59–72 (2017)
7. C. Lambeck, J. Wojdziak, R. Groh: *Facet Lens - Local Exploration and Discovery in Globally Faceted Data Sets*. DESIRE 2011: 85–88
8. B. Lee, M. Czerwinski, G. Robertson, B. Bederson: *Understanding research trends in conferences using paperLens*. CHI Extended Abstracts 2005: 1969–1972
9. B. Lee, G. Smith, D. Tan, et al: *FacetLens: Exposing Trends and Relationships to Support Sensemaking within Faceted Datasets*. CHI 2009: 1293–1302
10. M. Ley: *DBLP - some lessons learned*. PVLDB, 2(2): 1493–1500 (2009)
11. Y. Liu, S. Barlowe, Y. Feng, et al.: *Evaluating exploratory visualization systems: A user study on how clustering-based visualization systems support information seeking from large document collections*. Information Visualization 12(1): 25–43 (2012)
12. A. Ponsard, F. Escalona, T. Munzner: *PaperQuest: A Visualization Tool to Support Literature Review*. CHI Extended Abstracts 2016: 2264–2271
13. L. Quoc, T. Mikolov: *Distributed Representations of Sentences and Documents*. ICML 2014: 1188–1196
14. H. Reiterer, G. Tullius, T. Mann: *INSYDER: a content-based visual-information-seeking system for the Web*. Int J Digit Libr 5(1): 25–41 (2005)
15. F. Reitz: *A Framework for an Ego-centered and Time-aware Visualization of Relations in Arbitrary Data Repositories*. arXiv preprint arXiv:1009.5183v1 (2010)
16. B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist: *Designing the User Interface - Strategies for Effective Human-Computer Interaction, 6th Edition*. Pearson 2016
17. V. Sinha, D.R. Karger: *Magnet: supporting navigation in semistructured data environments*. SIGMOD '05: 97–106
18. G. Smith, M. Czerwinski, B. Meyers, D. Robbins, G. Robertson, D.S. Tan: *FacetMap: A Scalable Search and Browse Visualization*. TVCG, 12(5): 797–804 (2006)
19. J. Stasko, J. Choo, Y. Han, M. Hu, H. Pileggi, R. Sadana, C.D. Stolper: *CiteVis: Exploring Conference Paper Citation Data Visually*. IEEE VIS 2013 (poster)
20. J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, Z. Su: *Arnetminer: Extraction and mining of academic social networks*. KDD 2008: 990–998
21. J. Teevan, S.T. Dumais, Z. Gutt: *Challenges for Supporting Faceted Search in Large, Heterogeneous Corpora like the Web*. HCIR 2008: 6–8
22. F. Wei et al.: *TIARA: a visual exploratory text analytic system*. KDD'10: 153–162
23. J.D. West, T.C. Bergstrom, C.T. Bergstrom: *The Eigenfactor Metrics™: A Network Approach to Assessing Scholarly Journals*. C&RL 71(3): 236–244 (2010)
24. K.-P. Yee, K. Swearingen, K. Li, M. Hearst: *Faceted metadata for image search and browsing*. CHI '03: 401–408
25. J. Zhang, G. Marchionini: *Evaluation and evolution of a browse and search interface: Relation Browser++*. dg.o '05: 179–188
26. J. Zhao, S. Drucker, D. Fisher, D. Brinkman: *TimeSlice: interactive faceted browsing of timeline data*. AVI 2012: 433–436
27. <https://www.semanticscholar.org>
28. <http://www.core.edu.au/conference-portal>
29. <https://github.com/allenai/science-parse>