Visualization as Seen Through its Research Paper Keywords

Petra Isenberg, Tobias Isenberg, Michael Sedlmair, Jian Chen, Torsten Möller

Visualization as Seen Through its Research Paper Keywords

Petra Isenberg, *Member, IEEE*, Tobias Isenberg, *Senior Member, IEEE*, Michael Sedlmair, *Member, IEEE*, Jian Chen, *Member, IEEE*, and Torsten Möller, *Senior Member, IEEE*

Abstract—We present the results of a comprehensive multi-pass analysis of visualization paper keywords supplied by authors for their papers published in the IEEE Visualization conference series (now called IEEE VIS) between 1990–2015. From this analysis we derived a set of visualization topics that we discuss in the context of the current taxonomy that is used to categorize papers and assign reviewers in the IEEE VIS reviewing process. We point out missing and overemphasized topics in the current taxonomy and start a discussion on the importance of establishing common visualization terminology. Our analysis of research topics in visualization can, thus, serve as a starting point to (a) help create a common vocabulary to improve communication among different visualization sub-groups, (b) facilitate the process of understanding differences and commonalities of the various research sub-fields in visualization, (c) provide an understanding of emerging new research trends, (d) facilitate the crucial step of finding the right reviewers for research submissions, and (e) it can eventually lead to a comprehensive taxonomy of visualization research. One additional tangible outcome of our work is an online query tool (http://keyvis.org/) that allows visualization researchers to easily browse the 3 952 keywords used for IEEE VIS papers since 1990 to find related work or make informed keyword choices.

Index Terms—Keywords, data analysis, research themes, research topics, taxonomy, visualization history, theory.

1 Motivation

One of the main reasons why visualization is such a fascinating field of research is its diversity. There is not only a diversity of applications but also a diversity of research methods being employed, a diversity of research contributions being made, as well as the diversity of its roots.

Diversity of roots: The term ***visualization*** can be understood very broadly, expressing a long history of its use in common language. Therefore, it is not surprising that concepts of visual thinking have penetrated many areas of science, engineering, and philosophy. The field of modern (computer-based) visualization has been greatly influenced by research methods from the fields of numerics and computer graphics, which have given it its birth in 1990. The impact of human-computer interaction affected the birth of the InfoVis community in 1995 and the influence of applied statistics (such as data mining) and cognition has led to the establishment of VAST in 2006.

Diversity of research methods: Given its diverse roots, visualizations remains a highly inter-disciplinary field that borrows and extends research methods from other fields. Methods come from fields as diverse as the broader computer science, mathematics, statistics, machine learning, psychology, cognitive science, semiotics, design, or art.

Diversity of contributions and applications: Based on these diverse influences, the results of visualization research can be manifold: from engineering solutions to dealing with large data sources (such as real-time rendering solutions, distributed and parallel computing technologies, novel display devices, and visualization toolkits) to understanding design processes (as in perceptual guidelines for proper visual encodings and interaction or facilitating collaboration between different users through visual tools) to scientific inquiries (such as improved understanding of perceptual and cognitive processes).

While all these diverse influences make the field of visualization research an exciting field to be part of, they also create enormous challenges. There are different levels of appreciation for all aspects of visualization research, communication challenges between visualization researchers, and the challenge of communicating visualization as an independent field of research to the outside. These issues lead, in particular, to the frequently asked question "what is visualization?"—among funding agencies or even between colleagues. Given our field's broad nature, we need to ask how we can comprehensively describe and summarize all on-going visualization research. These are not just theoretical and philosophical questions, but the answer to these questions has many real-world (e. g., career-deciding) impacts—from finding the right reviewers during peer-review to administrative strategic decisions on conference and journal structures and foci.

So while "what is visualization?" is a fundamental question, it has not been discussed to a large extent within our community. In fact, thus far the approaches have mostly focused on understanding some sub-field of visualization (e. g., [17, 38, 42]) but the question for the broader community has rarely been tackled beyond general textbook definitions (e. g., [6, 34, 46]). Those who have approached the problem, did so in a top-down approach based on the opinion and experiences of the authors. For example, several taxonomies were suggested by experts based on tasks, techniques, or data models (e. g., [7, 38, 43]). Another way of splitting visualization into more focused areas has been through specific application foci (e. g., VisSec, BioVis, SoftVis, etc.).

What is missing in this picture is a bottom-up analysis: What types of visualization research are actually happening as expressed by single research contributions in the visualization conferences and journals. Our paper is one of the first steps in this direction. We analyze one type of data that can shed light on the diversity of visualization research: author-assigned keywords as well as author-selected taxonomy entries in the submission system for the three IEEE VisWeek/VIS conferences. Based on this analysis, we make the following contributions:

Mapping visualization research: In Sect. 4, through the vehicle of keyword analysis, we build a conceptual map of all visualization work as indexed by individual authors. Our main assumption here is that, while each single keyword might be understood in a slightly different way by different researchers, their co-occurrence with other keywords clarifies their meaning, especially when aggregated over many different usages (i. e., many research papers in a major publication venue). This co-occurrence analysis is the basis for deriving clusters and, therefore, research sub-fields. The use of keywords seen over the past years also allows us to understand historical trends and we report on the most prominent declining and rising keywords within all of visualization.

Taxonomy and Terminology Discussion: Visualization research is influenced by a diverse set of application domains. The vocabulary of

- *Petra Isenberg is with Inria, France. E-mail: petra.isenberg@inria.fr .*
- *Tobias Isenberg is with Inria, France. E-mail: tobias.isenberg@inria.fr .*
- *Michael Sedlmair is with the University of Vienna, Austria. E-mail: michael.sedlmair@univie.ac.at .*
- *Jian Chen is with the University of Maryland, Baltimore County, USA. E-mail: jichen@umbc.edu .*
- *Torsten Möller is with the University of Vienna, Austria. E-mail: torsten.moeller@univie.ac.at .*

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org.
Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx/

visualization is influenced by all these application areas. However, the resulting diversity of terms are only understood by a small set of visualization researchers and, therefore, hinders the dissemination of research results and insights across all visualization sub-fields. This is very well articulated in the collection of keywords throughout all of visualization. We are the first that collected and "cleaned" this data, making it available to the benefit of our community, allowing its systematic analysis. In Sect. 5 we argue for a unification of different vocabulary and further discussion on the establishment of a visualization taxonomy.

keyvis.org: In Sect. 6 we describe a Web-based search tool that (a) makes the keyword meta-data available to a broad set of people, that (b) helps researchers/users of visualizations to quickly find the right papers that relate to a given topic, and that (c) helps visualization researchers find descriptive keywords for their publications.

Table 1. Top ten keywords ***only*** of IEEE VIS/VisWeek; 2000–2015.

#	author keywords		PCS taxonomy (lower level; only 2008–15)	
1	information visualization	136	visual knowledge discovery	416
2	visualization	136	graph/network data	373
3	visual analytics	121	coordinated and multiple views	334
4	volume rendering	117	biomedical and medical visualization	284
5	flow visualization	67	volume rendering	278
6	interaction	56	time series data	276
7	isosurface	41	user interfaces	270
8	volume visualization	40	geographic/geospatial visualization	269
9	scientific visualization	38	data transformation and representation	267
10	evaluation	35	quantitative evaluation	266

2 RELATED WORK

We collected, analyzed, and coded author-assigned keywords of 26 years of IEEE VIS (VisWeek) papers to start a discussion on a refined visualization taxonomy. In this section we discuss related research efforts and explain why we focus on co-word analysis as a research method and author-assigned keywords as a data source to analyze.

2.1 Topic Detection in the Scientific Literature

Topic detection techniques have been used to analyze the scientific literature to understand global research trends or see links and patterns amongst scientific documents. Example techniques include co-citation analysis, co-word analysis, co-author analysis, word frequency analysis, and recently probabilistic methods such as those based on latent Dirichlet allocation [2], or the use of self-organizing maps (e. g., [39]).

In this paper we focus on co-word analysis as a method. Co-word analysis has tackled the problem by analyzing the scientific literature according to the co-occurrence of keywords, words in titles, abstracts, or even in the full texts of scientific articles [4, 8, 13, 16, 30, 47]. Callon et al. [5], in particular, wrote a seminal book on the topic that provides several methods that others have used and extended upon.

The closest co-word analysis studies to our work are Coulter et al.'s [9] work on the software engineering community, Hoonlor et al.'s [18] general investigation of the computer science literature, and, most recently, Liu et al.'s [33] analysis of the human-computer interaction literature. Liu et al. examined papers of the ACM CHI conference from 1994–2013, identified research themes and their evolution, and classified individual keywords as popular, core, or backbone topics. We employ similar approaches as in this previous work. However, we also engaged in a manual coding process of keywords and conducted a comparative evaluation of the emerged topics.

Similar to this previous work we focused on co-word analysis as a simple method that has been shown to work well on keyword data. We focused on the analysis of keywords as primary descriptors provided by paper authors themselves. Alternative data sources could have been paper titles, abstracts, or full texts. However, the extraction of domain-specific key terms from this data is difficult and often requires human intervention and manual processing even with more recent topic modeling techniques [8]. We, thus, opted for a relatively simple, tested, and well established analysis method to apply to a manageable set of domain-scientist keywords. Co-word analysis in our case rests on the assumptions that authors picked their paper keywords out of a somewhat finite and codified repository of domain terms; that they use terms together to propose non-trivial relationships; and finally that the proposal of the same relationship by different authors is informative of the structure of visualization research [11].

2.2 Understanding the Field of Visualization

We are not the first to have made an effort to summarize a large set of visualization papers in order to understand topics or trends. One of the earliest such efforts was a summary and clustering of visualization research papers by Voegele [44] in 1995 in the form of a two-dimensional clustering of all visualization papers up to this point. A more recent related study [23] discussed topic relationships between the fields of visualization, graphics, and data mining by comparing articles from IEEE TVCG, ACM TOG, and IEEE TKDE using a hierarchical topic model. In the process, the authors computed prevalent topics from IEEE TVCG papers from 2005–2014 which we later compare to our results. A second recent study [26] compared information visualization and data visualization using co-citation networks and co-occurring keywords. The authors found many shared research themes between the domains and also highlighted a few differences.

Other efforts of describing the domain of visualization have focused on specific aspects of visualization research. Sedlmair et al. [37], for example, did a thorough analysis of design study papers to summarize practices and pitfalls of design study approaches. Further, Lam et al. [28] studied the practice of evaluations in Information Visualization papers which was then extended to include all visualization papers by Isenberg et al. [22]. Others have surveyed, for instance, the literature on interactive visualization [27, 48], on tree visualizations [36], on quality metrics in high-dimensional data visualization [1], on human-computer collaborative problem-solving [10], or on visualization on interactive surfaces [21]. In addition, several textbooks (e. g., [6, 15, 46]) give overviews of visualization methodologies, tools, and techniques.

2.3 Topic Analysis in Visualization and Visual Analytics

In the visualization and data analysis literature, the closest work to ours is Chuang et al.'s [8] machine learning tool for topic model diagnostics, visual text analysis using Jigsaw [14], and the CiteVis tool [40] based on a dataset of visualization publications [20]. While these lines of work are not per se co-word analyses, their data sources also include visualization papers. In contrast, we primarily focus on the results of our analysis of themes and trends in the visualization literature rather than on the description of any specific tool or algorithm.

3 CO-WORD ANALYSIS OF THE VISUALIZATION LITERATURE

For our analysis of the visualization research literature we use a mixed quantitative and qualitative approach as outlined next.

3.1 Datasets

We collected the following datasets:

Author-assigned keywords: are freely assigned by the authors to their research paper. We collected this data manually from the PDFs of all papers from 1990 to 2015.

PCS taxonomy keywords: are chosen from a pre-defined visualization keyword taxonomy and are assigned by the authors during the submission of their research paper to the precision conference system (PCS) [41]. We used the complete data from 2008 to 2015.

Table 1 provides an overview of the top ten author keywords for IEEE VisWeek/VIS papers for 2000–2015.

3.2 Keyword Analysis

Our general approach was first to extensively clean the data and then to analyze similarities and differences of both keyword sets.

Table 2. Percentage of papers with keywords per year (rounded).

Year	'90	'91	'92	'93	'94	'95	'96	'97	'98	'99
%	0	2	6	6	4	12	27	48	49	50
Year	'00	'01	'02	'03	'04	'05	'06	'07	'08	'09
%	70	73	82	97	84	96	99	98	85	91
Year	'10	'11	'12	'13	'14	'15				
%	84	86	87	100	99	95				

3.2.1 Keyword Topic Coding

We first collected the full set of 2 431 previously published IEEE VisWeek/VIS full papers (incl. case studies) from 1990–2015. From this set, we extracted 4 319 unique keywords. Next, we consolidated keywords based on singulars/plurals, spelling mistakes, or acronyms This yielded a cleaned dataset that contained 3 952 unique keywords.

Next, we engaged in a manual, multi-pass coding of these cleaned keywords in order to find higher-level clusters of keyword topics in this dataset. All authors of this article participated in the coding.

Pass 1—initial coding: We conducted the first coding pass on a subset of the data covering 10 years of the conference from 2004–2013. Each coder assigned one or more higher-level topics from a freely-evolving, personal code set to 2 823 keywords.

Pass 2—clustering: We refined our topics through a combination of automatic clustering and manual re-coding and re-fining of the topic clusters. The resulting set of keywords contained 156 unique topics.

Pass 3—consolidation: Next, all co-authors met for a two-day workshop in order to compare and consolidate topics together with the established keyword set from the PCS taxonomy. During the workshop we consolidated keywords, discussed ambiguities, and re-clustered the keywords into higher-level categories such as *Applications, Evaluation, Theory, etc.* This step resulted in a list of 210 topics, organized into 15 higher-level categories, our first version of a new taxonomy.

Pass 4—refining: We then manually assigned all 3 952 keywords (from 1990–2015) to the 210 unique topics from this first version taxonomy. Four coders assigned one topic per keyword. With this coding pass we validated and refined the first version of the new taxonomy. Each coder collected problems that occurred in the assignment processes such as ambiguities or when a keyword appeared to cover too broad of a concept. We then carefully scrutinized this list of occurred problems and used it to refine the taxonomy into a second version.

Pass 5—re-coding: We then re-coded the 3 952 keywords again with this refined taxonomy, with 2 coders per keyword, and manually resolved conflicts through discussion for 34% of all keywords.

Pass 6—fine-tuning: In a final remote meeting, we discussed further ambiguities and refined the taxonomy based on term frequencies. This step lead to some additional consolidations of topics, and a few splittings which we manually re-coded. After these multiple passes, we then felt confident to have reached saturation with the third and final version of the new taxonomy that included 180 topics in 14 categories (see supplementary material).

Next, we determined which part of the data to include in our final analysis by calculating the percentage of papers with keywords per year (Table 2). To achieve representative coverage for a historical assessment of keyword use, we chose to include years with a $> 70\%$ coverage. We, thus, limited our analysis to the years 2000–2015. This filtered dataset contained 1 760 published papers. Out of these, 185 contained no author-assigned keywords, yielding a set of 1 575 papers we analyzed. These papers contained a total of 3 634 unique keywords.

3.2.2 PCS Taxonomy Keyword Dataset

The PCS taxonomy keyword data consists of the authors' classification of their papers according to a visualization taxonomy (called "PCS taxonomy" from now on). This PCS taxonomy was created in its present form in 2009[1] and was used for the different visualization venues starting that year. For the year 2008 we mapped keywords from a previous

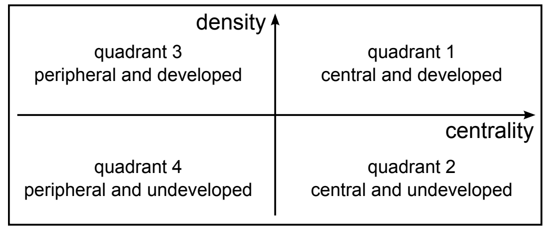

Fig. 1. Strategic diagram to characterize the topic clusters (after He [16]).

Table 3. The analyzed keyword datasets with their occurrence thresholds applied to keywords, number of remaining keywords analyzed, and target number of clusters set for the cluster analysis.

dataset	occur. thresh.	excluded terms	remaining keywords	target clusters
author	8	visualization, information visualization, scientific visualization, visual analytics	112	16
topic	15	visualization, information visualization, scientific visualization, visual analytics	116	16
PCS	0	none	127	16

to the current taxonomy. We analyzed anonymous data on 3 430 paper submissions that included the submission IDs and keywords, but the titles only of accepted papers. We did not have access to author names or any other identifying information for rejected papers. Considering data from both accepted and rejected papers allowed us to analyze topics the community was working on in a given year. In total, the PCS taxonomy includes 14 higher-level categories and 127 lower-level keywords (4–17 lower-level keywords per higher-level category).

3.2.3 Analysis Process

We analyzed three datasets: (1) cleaned ***author***-assigned keywords, (2) manually coded ***topics***, and (3) ***PCS*** keywords. We first filtered each keyword dataset by removing keywords that occurred less than a minimum threshold and also excluded higher-level terms as outlined in Table 3. Next, we generated document-keyword matrices with the keywords as variables (rows) and papers as observations (columns). Each cell contained a 0 if a keyword was not present in a paper and a 1 if it was. On each matrix, we performed a correlation computation using Python's NumPy `corrcoef` function that yielded a correlation matrix holding keyword-keyword correlation coefficients. On each correlation matrix we performed a hierarchical clustering using Ward's method [45] and a squared Euclidean distance metric. We also generated a keyword network in which two keywords were linked if their correlation was > 0 and each link was assigned its respective correlation value. From this network we computed the density of each cluster with the median of all inter-cluster links and the centrality by computing the sum of square of all links from within a cluster to another cluster. We plotted centrality and density in ***strategic diagrams*** [4, 9, 16, 19, 32, 33]. These diagrams distinguish four quadrants (Fig. 1) that characterize the different clusters based on their centrality within the research domain and on how developed they are. The diagram axes are centered on the median of the observed density and centrality values.

4 Results

Ultimately, we were interested in providing information and guidance on developing visualization keyword taxonomies. Thus, we focus our analysis on topic-coded keywords and PCS keywords but provide more information on the analysis of the cleaned author keywords in the supplementary material. Two main research questions drove our analysis of the data. In particular, we were interested in understanding major research themes and their relationship to other themes (Sect. 4.1) and the importance and evolution of individual keywords (Sect. 4.2).

[1]The first version of the taxonomy was created by the 2007 EuroVis paper chairs, Ken Museth, Torsten Möller, and Anders Ynnerman. This version was refined by several people over the next two years. After a broad consultation within in the visualization community, the final version was assembled by Hans-Christian Hege, Torsten Möller, and Tamara Munzner. This effort was supported by the VGTC—the IEEE Visualization and Graphics Technical Committee.

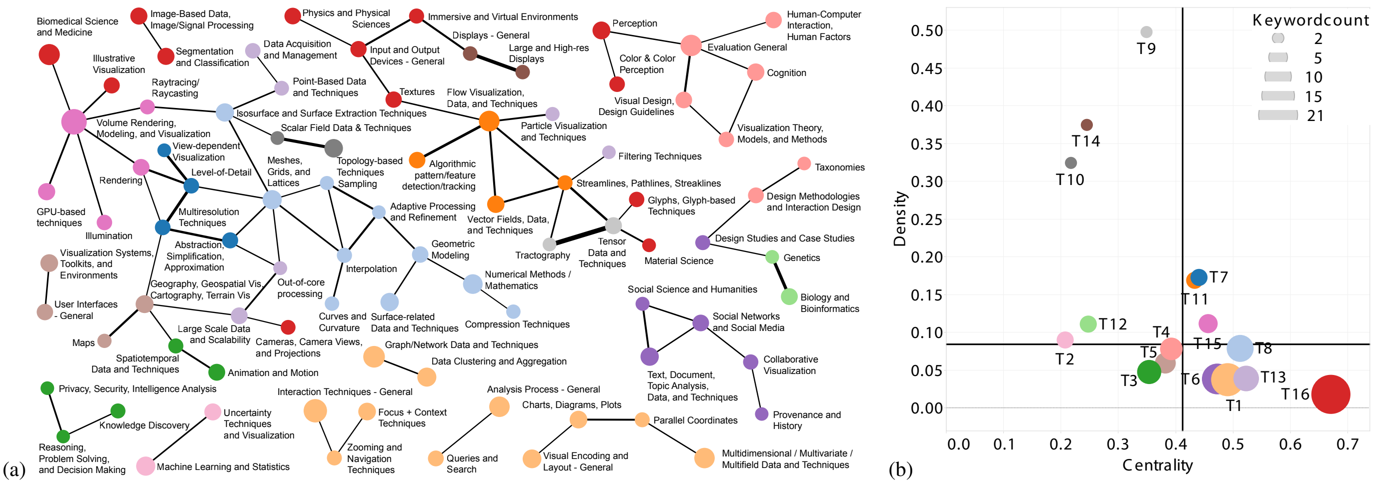

Fig. 2. Hierarchical clustering of the topic-coded keywords: (a) keyword map of nodes with correlation strength ≥ 0.12; circle size encodes keyword frequency. (b) Strategic diagram; circle size encodes number of keywords per cluster; colors match those used in the network graph.

Table 4. Resulting clusters of topic-coded keywords; represented by their two most frequent keywords. N = number of keywords per cluster.

ID	keywords (InfoVis, Vis/SciVis, VAST; 2004–2013)	N
T1	interaction techniques—general; graph/network data & techniques	15
T2	machine learning & statistics; uncertainty techn. & vis.	4
T3	timeseries, time-varying data & techn.; animation & motion	8
T4	evaluation general; cognition	7
T5	programming, algorithms, & data structures; geography, geospatial vis, cartography, terrain vis	6
T6	applications—general & other; text, document, topic analysis, data, & techn.	13
T7	abstraction, simplification, approximation; multiresolution techn.	4
T8	numerical methods/mathematics; meshes, grids, & lattices	10
T9	tensor data & techn.; tractography	2
T10	topology-based techn.; scalar field data & techn.	2
T11	flow vis., data, & techn.; vector fields, data, & techn.	4
T12	biology & bioinformatics; molecular science & chemistry	4
T13	large scale data & scalability; data acquisition & management	9
T14	displays—general; large & high-res displays	2
T15	volume rendering, modeling, & vis.; gpu-based techn.	5
T16	biomedical science & medicine; vis. techn. & tools—general	21

4.1 Analysis of Major Topic Areas

To understand major themes we analyzed the clustering results and the generated network graphs.

4.1.1 Topic-coded Keywords

Cluster Analysis: We created 16 clusters from 116 topic keywords (after removing 64 that each occurred $< 15\times$). We chose the cluster number from a manual inspection of the clustering result.

Table 4 gives an overview of the created clusters (T1–T16). In the table we report the two most frequent keywords in each cluster and the size of the cluster (N). The supplementary material provides a larger table with additional descriptive statistics such as density and centrality measures. The keyword map that results from the hierarchical clustering is shown in Fig. 2(a). Please note that keywords are clustered together when they frequently co-occurred on a paper—not necessarily because they are semantically similar. For example T16 contains the keywords *input and output devices* and *immersive and virtual environments*. These two keywords are in the same cluster because they frequently appeared together on papers, showing joint contributions in these areas. The strong correlation is indicated by a thicker line between the two keywords in Fig. 2(a) (two red circles top, middle).

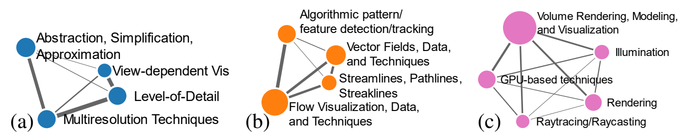

Fig. 3. Topic clusters that emerged as 'motor themes': (a) T7, (b) T11, and (c) T15. Line width encodes correlation strength.

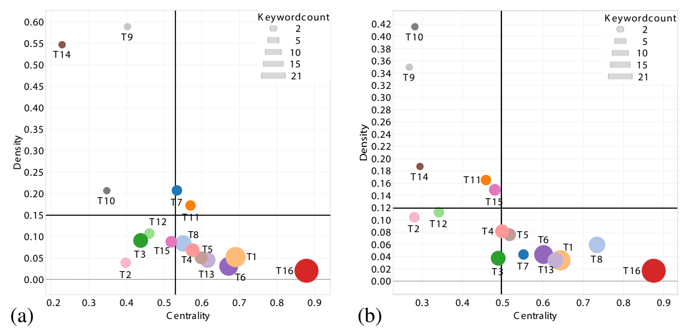

Fig. 4. Strategic diagrams of the topic keywords clustering of Fig. 2(b) (a) for 2000–2007 and (b) for 2008–2015. Black lines indicate the medians.

Network Analysis: The strategic diagram for topic keywords is shown in Fig. 2(b). The data contained 2 189 out of 6 670 possible links between topic keywords. The clusters in Quadrant I of the strategic diagram (top right) are considered motor themes or mainstream topics as they are both internally coherent and central to the research network. Compared to other research fields such as consumer behavior research [35], psychology [31], or software engineering [9], Visualization seems to lack clear motor themes. Yet, three clusters were situated within the lower left corner of the motor themes quadrant and could potentially emerge as more central motor themes in the future: T7, T11, T15—each represented in Fig. 3 with all internal links. All three motor theme clusters include topic keywords from research central to the historically largest subcommunity in visualization (Vis/SciVis) and this may be the reason that they emerge with a slightly more prominent role.

While visualization seems to lack strong motor themes, it has three clearly developed but isolated themes in Quadrant III (top left): T9 (*tensor data & techniques; tractography*), T10 (*topology-based techniques; scalar field data & techniques*), and T14 (*displays-general; large & high-res displays*). These topic clusters represent themes of research that have high internal density but are not central to the larger network—they are not strongly connected to topics in other clusters. They, can thus be seen as specializations.

Visualization also has several clusters (T6, T8, T11, T13, T16) in Quadrant II (bottom right) which are considered 'basic and transversal' themes as they are weakly linked together (low density) but well connected to the remainder of the network (high centrality). Hence, work here can be of significance to the entire network. Finally, themes in the lower left quadrant are weakly developed and marginal and are considered either declining or emerging. Thus, only a temporal analysis can reveal their significance to the field.

In order to gain a better understanding on how these clusters changed over time, we split the analyzed time period in two ranges: 2000–2007 and 2008–2015, covering 8 years each. To observe movement patterns, we recreated the strategic diagrams for these time periods while keeping the clusters from the overall analysis fixed. The two resulting strategic diagrams are shown in Fig. 4(a) and (b). Between the two time periods, several interesting movements can be observed. While T11 was still marginally a motor theme in 2000–2007, it recently moved towards the top left quadrant, indicating that flow visualization and its related topics are becoming a specialization. In contrast, T7 made the move from motor theme to a more central, yet undeveloped theme (bottom right). The quadrant with specialized (developed and isolated) themes in the top left also saw movement with T10 strongly rising and T14 strongly falling, indicating increased development for *topology-based techn.; scalar field data & techn.* and decreased development for *displays-general; large & high-res displays* between the time periods. The bottom left quadrant includes themes of either rising or declining importance to the field. According to this analysis, T2 (*machine learning & statistics; uncertainty techn. & vis.*) and T12 (*biology & bioinformatics; molecular science & chemistry*) became more developed but less connected to other themes; T15 (*volume rendering, modeling, & vis.; gpu-based techn.*) also became more developed while its connections to other themes remained relatively stable. Finally, T3 (*timeseries, time-varying data & techn.; animation & motion*) became more central to the overall field but less developed internally.

4.1.2 PCS Taxonomy Keywords

PCS keywords come from an existing taxonomy of visualization keywords. There is overlap in term usage with the topic keywords, as we used the PCS taxonomy to inform and enrich our hand-coded set of visualization keywords. Yet, before discussing the analysis of PCS keywords and their comparison to the other keyword categories, it is important to note that their usage in practice is different than for author keywords. Author keywords are generally used to label a paper and to explain in a few words what main research topics it covers—for example, with the intention to make the paper discoverable in searches. While the PCS keywords can be used in the same way, the choices are limited and it is not always possible to find the one keyword that exactly expresses a contribution. For the PCS data we analyzed, authors were instructed to choose only a limited subset of PCS keywords (e. g., a primary and up to four secondary keywords). Authors, thus, had to balance their choices among all possible keywords that describe a paper contribution. A fundamental difference to author keywords, however, is that the PCS taxonomy is used not only by the paper authors but also by the reviewers to describe their expertise levels for each topic. The selection of PCS keywords for papers and by reviewers then informs how reviewers are assigned to papers. Since this process is known to many paper authors, PCS keywords are often carefully selected to rank more highly for reviewers with a certain expertise.

Given these differences in use, PCS keywords may not as accurately reflect the topic of a paper as the author keywords and the respectively coded topic keywords do. Nevertheless, by conducting the same analysis on this keyword set as we did before for topic keywords we can speculate on differences in use as well as on topics that may be missing in each type of keyword set.

Cluster Analysis: The result of the analysis of the PCS taxonomy is reported in the same way as for the other keyword set. Table 5 provides the generated clusters P1–P16, their top two keywords, and total number of keywords. The supplementary material holds a table with all keywords and additional cluster metrics. Fig. 5(a) shows a filtered network graph generated from the keyword correlations, while Fig. 5(b) depicts the strategic diagram.

The cluster analysis includes several interesting observations. When the PCS taxonomy was created, keywords were grouped into 14 higher-level categories such as *applications*, *evaluation*, or *perception & cognition*. Most of our clusters did not align with these higher-level categories, showing that most papers make contributions across those categories. In particular, the keywords from the *application* category were separated across a variety of our clusters. This finding suggests that applications spread out because they are being more closely correlated with specific visualization techniques. P1, for instance, combined *volume rendering* and *biomedical/medical visualization* and P5 included *high-dimensional data*-related keywords along with *bioinformatics*. Both suggest a close relation between this data type/techniques, and certain application domains. P14 contained another interesting mix, including *evaluation* and *perception & cognition* keywords. This combination shows that research on *perception & cognition* is strongly based on user studies. On the other hand, there were also more uniform clusters. P15, for instance, contained only *interaction techniques*-related keywords, which suggests that contributions made in this research area could often be well described with this set of keywords alone.

Table 5. Resulting clusters of PCS keywords; represented by their two most frequent keywords. N = number of keywords per cluster.

ID	keywords (InfoVis, Vis/SciVis, VAST; 2008–2015)	N
P1	volume rendering; biomedical and medical visualization	8
P2	time-varying data; vis. in physical sciences and engineering	8
P3	feature detection and tracking; vector field data	8
P4	visual knowledge discovery; graph/network data	13
P5	high-dimensional data; multidimensional data	5
P6	coordinated and multiple views; time series data	11
P7	collaborative and distributed vis.; large and high-res displays	4
P8	intelligence analysis; situational awareness	8
P9	scalability issues; streaming data	10
P10	multiresolution techniques; compression techniques	3
P11	scene perception; stereo displays	3
P12	data transformation and representation; data aggregation	9
P13	illustrative vis.; multimedia (image/video/music) vis.	10
P14	quantitative evaluation; laboratory studies	10
P15	user interfaces; interaction design	5
P16	visualization models; qualitative evaluation	12

Network Analysis: Next, we analyzed the strategic diagram in Fig. 5(b). The data contained 3 607 out of 8 001 possible links between PCS keywords. Interestingly, in contrast to the topic-keyword sets, PCS keywords showed a stronger set of motor themes (see Fig. 6). P12 (*data transformation and representation; data aggregation*), in particular stood out. This motor theme indicates that visualization is indeed not only about visual encoding, but that data pre-processing has always been one of the most integral parts of visualization research. While this notion is usually clear to community insiders, it is often overlooked by outsiders. P8 and P14 are less strong motor themes but have the potential to become more driving themes. P8 includes many keywords that are typically attributed to the VAST conference such as *intelligence analysis* or *emergency and disaster management*. P14 outlines the importance of perception and evaluation to visualization research with keywords centered around these topics.

Topics that are only weakly-linked together but well-connected to the remaining network (basic and transversal topics, bottom right) include P4, 9, 13, and 16. Many of the keywords in these clusters are cross-cutting the entire field of visualization, but are not necessarily topics that would individually characterize a subfield on their own. P16, for example, includes a variety of research methods used in visualization. While clearly relevant to study many different aspects of visualization, these methods are not frequently studied in and of itself, i. e., papers solely focusing on methods usually are published at workshops such as BELIV (www.beliv.org) [29]. In terms of clearly developed but isolated specializations (top left), our analysis included *multiresolution and compression techniques* (P10) as well as *scene perception and*

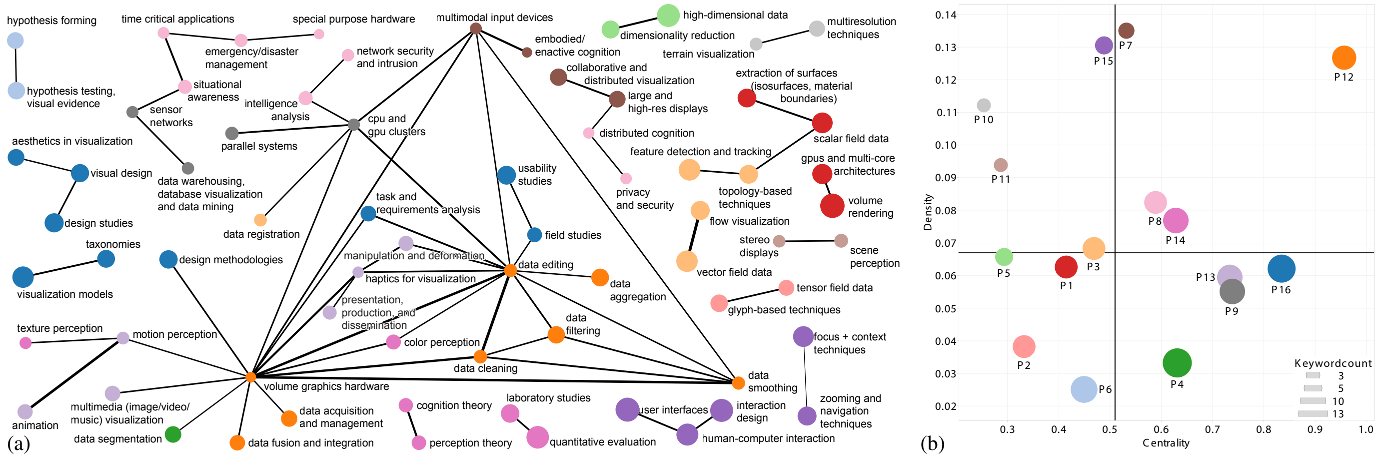

Fig. 5. Hierarchical clustering for PCS taxonomy keywords: (a) keyword map of nodes with correlation strength ≥ 0.15; (b) strategic diagram.

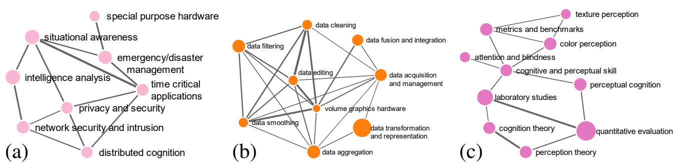

Fig. 6. Motor-theme clusters for PCS keywords: (a) P8, (b) P12, (c) P14. Line width encodes correlation strength (correlations <0.1 not shown).

stereo displays (P11). Finally, P2, which includes a combination of specific methods/data types along with application areas, indicated a clear cluster with either declining or emerging keywords (bottom left).

4.2 Analysis of Individual Keywords

In this section, we analyze individual keywords in more detail across all three of our data sources.

4.2.1 Relationship between Topic-coded and PCS Keywords

As both topic-coded and PCS keywords present a higher-level assessment of the field of visualization, we were interested in analyzing their relationship. Given that we had previously integrated our topics with the PCS taxonomy, the two keyword sets are relatively comparable. Table 6 shows the top 15 keywords per category. In both sets, keywords related to the topics: *graphs/networks*, *biomedial/medical*, *MD/high-dimensional data*, *evaluation*, and *temporal data* were frequent.

Some of the difference between the most frequent keywords can be explained by the fact that PCS keywords often lack a "general" category. For example, the hand-coded topic *interaction-techniques–general* caught all keywords that were not already covered by more specifically listed techniques such as *Manipulation and Deformation* or *Zooming and Navigation Techniques* and subsequently ranked highly amongst the topics. The PCS taxonomy, on the other hand, lists 9 keywords in the *Interaction Techniques* category, none of which is a clear choice for authors trying to submit work on specific interaction techniques such as selections or brushing & linking. Similarly, the hand-coded topic keyword *evaluation general* caught evaluation-related keywords that were not already covered by other related keywords, and ended up ranking highly amongst the other topics.

We also saw several notable differences in the data: the most common PCS keyword *visual knowledge discovery*, for instance, ranked only 77$^{\text{th}}$ amongst the hand-coded topics (used 28×). Similarly, the 10$^{\text{th}}$ most frequent PCS keyword *quantitative evaluation* (266×) only occurred 10× in the hand-coded topics, despite including the summed counts of 5 author keywords (*null hypotheses*, *quantitative analysis*, *quantitative empirical study*, *quantitative evaluation*, and *quantitative study*). This difference could be indicative of the different usage of the PCS and author keywords: the PCS keyword may have been added to receive reviewers with expertise in quantitative studies.

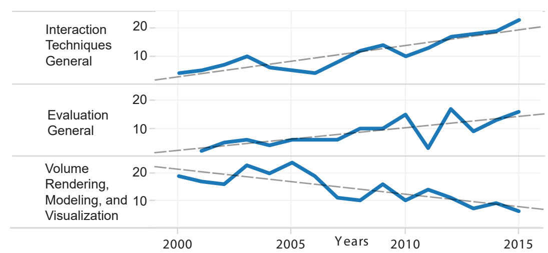

Fig. 7. The two strongest rising and the strongest falling topic keyword.

Table 7. Top 10 topic keywords for two time-periods of eight years each. Keywords that appear in both periods are bolded.

2000–2007	N	2008–2015	N
volume rendering, modeling & vis.	149	**interaction techniques – general**	126
meshes, grids, & lattices	67	evaluation – general	93
flow visualization, data, & tech.	64	**volume rendering, modeling & vis.**	83
biomedical science & medicine	61	**graph/network data & tech.**	80
numerical methods / mathematics	60	multidim./-var./-field data & techn.	75
graph/network data & tech.	54	analysis Process - general	74
interaction techniques – general	49	timeseries, time-varying data & tech.	72
isosurface & surface extraction techn.	49	machine learning & statistics	64
surface-related data & techniques	44	**biomedical science & medicine**	61
programming, algor. & data struct.	41	text, doc., topic analysis, data, & tech.	54

4.2.2 Rising and Declining Keywords

Next, we were interested in historic trends for individual keywords. We calculated linear trend lines and their significance for the 15 most frequent keywords in topic and PCS datasets. Table 6 gives an overview of this data, ordered by slope from increasing to decreasing. Significance was established via individual one sample 2-sided t-tests using Tableau 9.3's "describe trend model" feature. A p-value of $< .05$ was considered significant. Significant results are highlighted in bold.

Comparing the topic and PCS keywords, several common rising themes emerge: *interaction*, *evaluation*, and *time-series*. In addition, 10 different keywords (6 topics, 4 PCS keywords) show a significant increase in usage. Fig. 7 highlights the data for *interaction* and *evaluation* topics. It shows that both have become increasingly prominent research topics and that the need for increased interactive data analysis and validation has been recognized by the community. The figure also visualizes the strong decline of volume visualization research, which has been a historically important and frequent topic as part of motor theme T15. This could perhaps be due to the fact that many fundamental research questions have been tackled and that researchers are now moving elsewhere or use more specific or other keywords.

We also looked at the most common topic keywords for the 2000–2007 and 2008–2015 time periods (see Table 7). A large overlap in top

Table 6. Historical trends for 15 most frequently used keywords for each of the topic and PCS datasets. Significant trends highlighted.

topic (2000–2015)								PCS (2008–2015)							
keyword	chart	#	slope	SE	df	*p*-val.	*t*-val.	keyword	chart	#	slope	SE	df	*p*-val.	*t*-val
interaction techniques—general		175	1.12	0.15	14	**<.001**	7.503	**quantitative evaluation**		266	4.40	0.73	6	**.001**	7.457
evaluation general		128	0.81	0.20	13	**.001**	4.068	**visual knowledge discovery**		416	4.33	1.07	6	**.007**	6.296
machine learning and statistics		85	0.73	0.11	14	**<.001**	6.881	**time series data**		276	4.31	0.76	6	**.001**	6.190
timeseries, time-varying data and techniques		109	0.60	0.18	14	**.005**	3.294	**geographic/geospatial vis.**		269	4.20	1.05	6	**.007**	4.000
multidim./multivar./multifield data and techn.		109	0.57	0.12	14	**<.001**	4.555	**coordinated & multiple views**		334	4.17	1.47	6	**.030**	2.962
analysis process—general		110	0.47	0.18	12	**.024**	2.580	**data transf. and repres.**		267	3.75	1.20	6	**.020**	3.354
graph/network data and techniques		134	0.46	0.15	14	**.007**	3.169	interaction design		246	1.45	1.21	6	.275	1.951
visual encoding and layout—general		78	0.32	0.09	13	**.004**	3.455	graph/network data		373	0.89	1.03	6	.419	1.467
data clustering and aggregation		83	0.21	0.08	12	**.030**	2.465	multidimensional data		255	0.75	0.32	6	.056	2.633
visualization techniques and tools—general		82	0.18	0.10	13	.085	1.853	user interfaces		270	0.60	1.93	6	.768	1.983
biomedical science and medicine		122	-0.02	0.13	14	.855	-0.187	high-dimensional data		236	0.45	1.61	6	.788	1.572
flow visualization, data, and techniques		114	-0.11	0.16	14	.488	-0.713	biomedical and medical vis.		284	0.07	1.07	6	.949	0.204
numerical methods / mathematics		94	-0.42	0.15	14	**.016**	-2.752	vis. system and toolkit design		262	-0.31	1.78	6	.868	-0.174
meshes, grids, and lattices		86	-0.68	0.10	14	**<.001**	-6.869	time-varying data		263	-0.77	0.71	6	.315	0.486
volume rendering, modeling, and vis.		232	-0.92	0.19	14	**<.001**	-4.776	volume rendering		278	-2.52	1.33	6	.107	-1.870

keywords could indicate that the visualization community has become cohesive. Yet, within the top 10 keywords only four can be found in both time periods (highlighted in bold), showing that the landscape of visualization research is still rapidly changing and evolving.

4.2.3 Infrequent Keywords

Author keywords exhibited a power-law distribution, indicating that a very large number of keywords appeared very infrequently. While the frequency of PCS and topic keywords did not follow the same trend, still a large number of keywords in each category appeared very infrequently. The 10 least frequently used PCS keywords, for example, each made up less than 1 ‰ of the PCS keyword usage: *Special Purpose Hardware*, *Sonification*, *Volume Graphics Hardware*, *Haptics for Visualization*, *Embodied/Enactive Cognition*, *Privacy and Security*, *CPU and GPU clusters*, *Distributed Cognition*, and *Texture Perception*. The 10 least frequent topic keywords were each used for 3 or fewer papers in our corpus. Seven of these keywords were shared with other low-frequency PCS keywords (< 40 papers): *CPU and GPU clusters*, *PDE's for Visualization*, *Sensor Networks*, *Data Editing*, *Field Studies*, *Time Critical Applications*, and *Education*. Two of the remaining low-frequency topic keywords were not part of the PCS taxonomy—*Cutting Planes* and *Multi-core processing*. An outlier was the infrequent topic keyword *Usability Studies* which occurred comparatively frequently (121×) within the PCS dataset.

4.3 Comparison To Related Work

The PCS taxonomy is not the only taxonomy that has been applied to visualization papers. Both the IEEE and INSPEC have a fixed set of keywords that are applied by "experts in the field" to each IEEE VIS paper in the IEEEXplore digital library. After having removed general research area keywords such as data visualization, visualization, or computer science—the top 10 keywords for all IEEE VIS papers from 2000–2015 can be seen in Table 8. Not much can likely be learned from the IEEE terms as keywords such as *human*, *shape*, or *chromium* seem nonsensical among the most common keywords. Similarly, while the INSPEC controlled index's top terms seem more related to the field, they do not align well with our findings as there is little to no overlap between the terms in any of the three keyword sets we considered. Finally, we can compare our work to Jiang and Zhang's [23] analysis of TVCG paper titles, keywords, and abstracts. Their 16 topical hierarchies of 4 sub-topics are also shown in Table 8. The authors used a hierarchical topic model, and thus, the keywords found are of a different nature than ours. Yet, some similarly important topic areas seemed to emerge such as those related to graphs, medical visualization, or surfaces.

Table 8. Related Keyword Taxonomies and Analyses

IEEE terms	INSPEC Controlled	Jiang and Zhang [23] cluster topics
rendering (CG)	rendering (CG)	scatterplot, space, coordinate, fiber
computer graphics	data analysis	data, set, contour, scalar
visual analytics	computational geometry	medical, vessel, diagnosis, blood
comp. modeling	interactive systems	memory, gpu, compression, performance
data mining	medical image proc.	virtual, environment, walk, reality
displays	user interfaces	graph, layout, node, map
data analysis	graphical user interfaces	display, projector, registration, camera
chromium	feature extraction	mesh, isosurface, reconstruction, interpolation
humans	solid modeling	uncertainty, event, ensemble, temporal
shape	pattern clustering	video, color, motion, texture

5 Discussion and How to Move Forward

Our analysis of the keyword data has revealed several major themes as well as declining and rising keywords. From the strategic diagrams we saw that motor themes only clearly emerged for the limited PCS keyword taxonomy, while the topic keywords indicated no strong motor themes. This as an interesting finding: it could be indicative of visualization being a highly diverse and context-dependent field.

While we have collected a large amount of author keyword data, invested heavily in clustering related author keywords, and compared this data to a standardized taxonomy, our analysis is only a first step in the direction of two larger research goals as discussed next.

5.1 Creating Motor Themes and a Common Vocabulary

Our analysis of raw author keywords and our subsequent coding of topic keywords has revealed that authors choose many variants of similar keywords based on: singulars and plurals (e. g., *glyphs* vs. *glyph*), abbreviated versions (e. g., *DTI* vs. *diffusion tensor imaging*; *InfoVis* vs. *information visualization*), spelling (*multidimensional* vs. *multi-dimensional*), specificity (e. g., *multi-dimensional* vs. *multi-dimensional data* vs. *multi-dimensional visualization*), or not yet universally established definitions (*focus+context* vs. *focus-in-context* vs. *overview+detail*). Such a diversity of terms may be a reflection of the diversity of influences on the visualization field—but it is not helpful, in particular when one wants to search for keywords or—like us—gain an overview of the research themes of a community. We hope that keyvis.org (Sect. 6) will help paper authors find common terms and reflect on their keyword usage before submitting a camera-ready paper.

One can also think about the problem of creating a common vocabulary for visualization more broadly. By identifying key terms and providing clear definitions, sub-communities in visualization may be able to communicate more clearly about similar approaches. This can also help to collaborate more effectively with people outside the community. Finally, a common vocabulary can also allow us to more easily understand emerging and declining research trends within the field.

A discussion on common vocabulary is, in a wider-sense, also related to the questions on whether visualization has any overarching theories, mainstream research methodologies, or clearly shared and accumulated knowledge. The lack of clear motor themes and large number of basic and transversal themes we observed for topic keywords is similar to observations made by Liu et al. [33] for the field of Human-Computer Interaction (HCI). The authors partially attribute this to the fact that HCI research tends to be highly contextual, instead of universal as in other disciplines. Therefore, in HCI it is difficult to accumulate knowledge that is applicable to the field in general. A similar explanation could hold for the field of visualization with its three main streams InfoVis, SciVis, and VAST—as well as the field's dependence on changing data and work contexts. The large number of transversal themes we found is also indicative of the fact that the field is growing—again, similar to observations made for HCI [33]. Finally, the lack of a clear common vocabulary and motor themes is also an indication of the field's focus on novel techniques and tools instead of being more incremental by improving upon existing solutions.

5.2 Establish a Comprehensive Taxonomy

One of the goals that we had initially set out to accomplish is not yet achieved. Perhaps the holy grail of a keyword analysis is to amount into a taxonomy of the analyzed field, in our case visualization research. This could serve two purposes. One the one hand, a taxonomy will help to better communicate "what is visualization" to other disciplines, i. e., to researchers and practitioners not part of the VIS community. On the other hand, we are hoping to be able to facilitate the crucial step of matching reviewers with papers and grants such that the peer review process improves and new contributions are seen in the right context.

Yet, how to exactly create and in particular **maintain** a comprehensive taxonomy of visualization keywords is still an open question. As can easily be seen from an inspection of both the top IEEE and INSPEC terms in Table 8, broader taxonomies are not very successful in capturing the diversity of the visualization field. The PCS taxonomy, in turn, has been developed by a team of dedicated experts with the goals to improve the reviewing system for visualization papers. Yet, what is the right process for maintaining and changing the taxonomy? The visualization field is evolving and, thus, a visualization taxonomy should be regularly updated. Should an analysis such as ours be used to find trends (and keywords representing these trends) that have not been captured? Would it be possible to automate our process without requiring experts to clean and code the data? Should certain keywords be split or merged as sub-areas increase or decrease in popularity? Does it make sense to keep keywords in the PCS taxonomy that are rarely used—or should the taxonomy provide a broad view of the field in order to capture its topic diversity? Even when one has answers to these questions, how would one choose the right level of granularity for keywords? Finally, should there be separate taxonomies: one for visualization as a whole (e. g., to conduct analyzes on emerging and declining topics, topic coverage across subcommunities, etc.) and one for the submission process for academic research? The former taxonomy could be large and evolving while the latter would have to be reduced in size and remain stable across conferences for given time periods to remain manageable for papers chairs, editors, reviewers, and authors.

5.3 Lessons Learned from Topic-Coding

In the work for this paper, we have deeply dived into the question of how to create, refine, and use taxonomies based on keywords. This process has shed light on several challenges:

Objectivity: An ideal taxonomy would objectively reflect the underlying research area, offering uniformly and equally distributed taxons. When creating or refining taxonomies, however, human's subjective views [25] can lead to an overemphasis of areas that oneself works in, or those that are felt to be trending at the moment and lead to unbalanced taxonomies (see Sect. 4.2.3). For instance, we felt that we had frequently encountered keywords related to sports visualization. Yet, after adding the topic, we found out that it was only included in 5 papers. Similar problems may have occurred when creating the PCS taxonomy. For example, it includes 11 *perception & cognition*-related keywords, half of which ranked in the bottom 25% of keywords according to frequency. At the same time, subjectivity might also lead to the opposite phenomena, that is, underspecification of areas that are less well known. This might lead to overly broad terms. For example, the most frequent topics *volume rendering*, *interaction techniques*, or *graphs/networks* may be worth splitting up to be more discriminative.

Clarity: Specifically with respect to author keywords, a large practical challenge that we faced was dealing with ambiguities. The most challenging cases stemmed from ambiguities based on missing context. *Coding* might, for instance, refer to coding as in programming or as in open coding for grounded theory; or *statistics*, which might refer to statistics used for analyzing user study data, or statistics used to aggregate data before feeding into a visualization. Similar challenges are encountered by authors and reviewers when choosing PCS keywords. For example, what is the difference between *multidimensional data* in the *large data vis* category, and *high-dimensional data* in the *non-spatial data and techniques* category?

Higher-level Categorization: Finding meaningful higher-level categories for our topic keywords was particularly challenging. While many strategies for this task such as affinity diagramming exist, there might still be keywords left that would fit equally well into different categories: should *design methodologies and interaction design*, e. g., go into a category of *interaction techniques and general HCI* or into *theory*? Finding good higher-level categories for keywords is important: they give context and additional meaning—or can be confusing as in the *md data* vs. *high-d data* data example mentioned above.

Concrete suggestions for the PCS taxonomy: Our analysis highlights several lightweight changes to the PCS taxonomy that can be applied before a major restructuring is considered: (a) The taxonomy lacks the possibility to select higher-level keywords (such as *interaction*)—or, alternatively, an *other* keyword in each of the higher-level categories (such as *interaction techniques—other*). (b) In Sect. 4.2.3 and in the supplementary material we highlight several keywords which have only been chosen extremely infrequently and, thus, their inclusion should be re-thought. Similarly, broad terms should potentially be split. (c) Finally, in the creation of our topics for the coding of the author keywords we made several changes to the naming of PCS keywords in order to avoid misunderstandings and ambiguities. While we do not want to claim that our changes are perfect, they could be taken as inspiration for the discussion of new keyword names for the PCS taxonomy. For example, we chose to remove most mention of *visualization in X* to just point to the field *X* more generally. We made this change to more clearly capture contributions in the respective field that may not be primarily visual—such as a field study. We also renamed other keywords to be more general, such as *Collaborative and Distributed Visualization* to *Collaborative Visualization* as it is not necessary to highlight a distributed mode of collaboration (vs. a co-located mode) in the name of the keyword.

5.4 Limitations

While our analysis has revealed a wealth of information, the study results have to be read in light of several analysis limitations. One obvious limitation is, of course, that we only analyzed a subset of publications from the visualization domain. To determine this subset, however, we followed advice from Bradford's law [3] for selecting our data sources. This law states that a small core of publications will account for as much as 90% of the literature in terms of citations received—trying to achieve 100% will add publications at an exponential rate. Thus, it is unreasonable to attempt to cover the whole literature for a field. Given the size and importance of IEEE VisWeek/VIS, we focused on a 16-year dataset from this conference (plus the additional data on the use of the PCS taxonomy). This analysis enabled us to get a rich overview of the field. Yet, compared to several past keyword analysis studies (e. g., [5, 32, 33]), the visualization field is still young and the overall number of keywords was comparably low, in particular for the author-assigned keywords. The low number of overall keywords and the vast difference

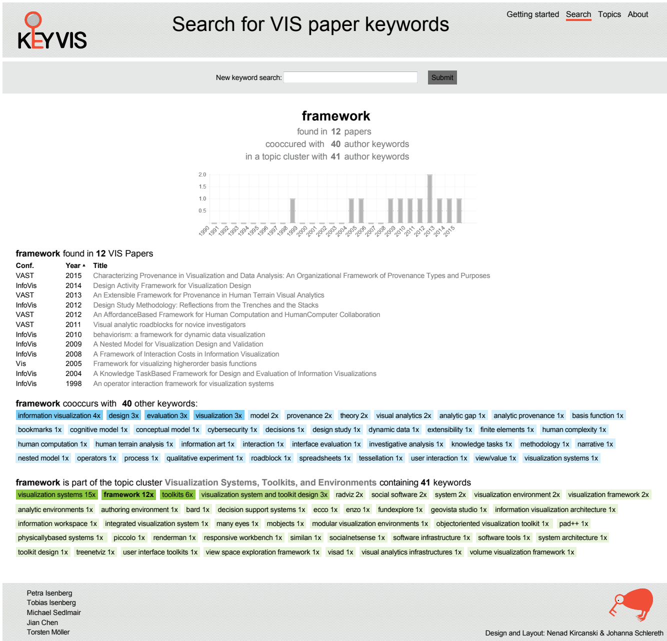

Fig. 8. Screenshot of keyvis.org's search page. Searching for the author-assigned keyword "framework" returns the full list of papers that used that keyword and their temporal distribution (bar chart). It also provides a list of all co-occurred keyword (blue), as well as the topics we manually assigned it to (green). All elements are interactive and can be used to start further investigations from there.

in number of papers accepted to the IEEE VisWeek/VIS conferences is also one of the main reasons we did not study the difference between IEEE VAST, InfoVis, and Vis/SciVis in depth but looked at the whole field together. By expanding the dataset back to 1995 (the first year of InfoVis) a comparative analysis may be more meaningful. Yet, another peculiarity of the field may further impact such a comparison: for a long period of time it was possible to send information visualization papers to both the InfoVis and Vis/SciVis conference. The Vis/SciVis conference even included information visualization sessions. Thus, the borders between the conferences were not always as clear as their current names may suggest.

For the PCS data, we discussed a major limitation earlier that pertains to the different use of these keywords compared to author-assigned keywords. In addition, we also found that several older papers included ACM classifiers and it is possible that authors back then only selected author keywords in addition to these classifiers and did not provide duplicates. As we did not want to speculate on what may have happened, we collected the author keywords as present on each paper.

The expert coding process that led to the topic keywords also, of course, includes an inherent subjective coding bias. While our team consisted of 5 experts with varying backgrounds in VAST, InfoVis, Vis/SciVis, and HCI, it is entirely possible that with other coders other clusters and cluster names may have emerged. Finally, this keyword analysis was conducted by experts in the field of visualization and not professional social scientists. Our own experience and views of our field have certainly tainted our interpretations of the data—as is common with any kind of qualitative analysis [12].

6 KEYVIS.ORG: A KEYWORD SEARCH TOOL

To make the result of our work easily accessible we created a webpage that makes author keywords and expert topics as well as the related papers search- and browsable: http://www.keyvis.org/. Visitors can search all 3 952 unique author-assigned keywords, find out which keywords co-occurred how frequently, which manually coded topics they belong to, and the actual research papers on which they appear (see Fig. 8). It also supports partial keyword searches to broaden the scope to keywords that have not been explicitly used yet. Note that keyvis.org is only meant to make our data publicly available and explorable. In particular, keyvis.org is not designed to directly support the analyses we presented above (which we conducted using a combination of R, Python, and Tableau). Instead, we wanted to provide an easy-to-use, lightweight interface to our keyword data, with two use cases in mind:

1. Supporting Keyword Selection. Picking keywords for publications is often done without much consideration. However, this practice does not necessarily lead to "good" keywords. Good keywords are those that make it easy for others to find papers, e. g. in large databases such as IEEE Xplore (http://ieeexplore.ieee.org/) [24]. Keyvis.org is meant to support visualization researchers in their choices of good keywords. Consider, for instance, a scenario of researchers working on a visualization framework for trajectory analysis. Seeking guidance for picking keywords for the respective publication, they can use keyvis.org to search for keyword candidates. In doing so, they can learn about other potentially relevant keywords that co-occurred in the past with their candidates (blue in Fig. 8), as well as those that we have manually grouped together into larger, semantically meaningful topics (green in Fig. 8). In our scenario, searching for the keyword *framework*, for instance, reveals the related keyword *visualization system* in the topic cluster (green). The researchers might consider *visualization system* as less generic and more discriminative than *framework*, and they also see that it occurred even more often ($\times 15$) than *framework* ($\times 12$). Thus, they might decide to use it instead. Similarly, searching for *trajectory analysis* reveals the co-occurring (blue) keyword *movement data*. While the researchers have not considered depicting their work in this way before, they find it very helpful, and hope that including it might further increase the chance that interested readers will find their work. As illustrated in this scenario, keyvis.org is meant to support broadening the consideration space and, in doing so, making the keyword selection process more conscious and systematic.

2. Identifying Related Work. While better supporting keyword selection was our initial goal, we found that our approach also fosters a second use case, namely, identifying related work. Consider the scenario from above: becoming aware of the close connection of *trajectory analysis* and *movement data*, might not only help to pick good keywords, but also to spot related papers that were previously missed, specifically if they were published in different sub-communities. While *trajectory analysis*, for instance, only reveals papers from VAST, *movement data* also includes an InfoVis paper. Looking further into the keyword *movement* also reveals related work in (Sci)Vis. In the ideal world, we would simply expect researchers to know all such facets. However, psychology research has shown that humans (including researchers) are strongly biased by their context, and hence might easily miss things that are not immediately available/visible to them: *"what you see is all there is"* [25]. We therefore also included a full list of all papers. The papers are queried via the keyword search, and hence can be interactively explored and related to each other through co-occurring keywords and the topic clusters we created.

Since the first version of keyvis.org in 2014, we have used the site ourselves for finding keywords and related work that we were not aware of before. We hope that others will find it similarly useful. The website has undergone multiple rounds of usability tests to ensure its ease of use and understandability. In the long run, our goal is to maintain the website as a platform for visualization keyword access and analysis.

ACKNOWLEDGMENTS

We would like to thank Meghan Haley for providing us with the anonymized PCS data, Vassilis Kostakos for the details about the keyword analysis process for CHI keywords, and Jiawan Zhang for more detailed analysis results of their TCVG paper analysis. We also thank Nenad Kircanski and Johanna Schlereth for their help with keyvis.org. This work was partially funded by FFG project 845898 (VALID) and by US National Science Foundation IIS-1018769, DBI-1260795, and MRI-1531491 and US Department of Defense USAMRAA-13318046.

REFERENCES

[1] E. Bertini, A. Tatu, and D. Keim. Quality metrics in high-dimensional data visualization: An overview and systematization. *IEEE Transactions*

on Visualization and Computer Graphics*, 17(12):2203–2212, 2011. doi: 10.1109/TVCG.2011.229

[2] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

[3] S. C. Bradford. Sources of information on specific subjects. *Journal of Information Science*, 10(4):173–180, 1985. doi: 10.1177/016555158501000406

[4] M. Callon, J. Courtial, and F. Laville. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, 22(1):155–205, 1991. doi: 10.1007/BF02019280

[5] M. Callon, J. Law, and A. Rip, eds. *Mapping the Dynamics of Science and Technology*. Palgrave Macmillan, Hampshire, UK, 1986. doi: 10.1007/978-1-349-07408-2

[6] S. K. Card, J. D. Mackinlay, and B. Shneiderman. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, San Francisco, 1999.

[7] E. H. Chi. A taxonomy of visualization techniques using the data state reference model. In *Proc. InfoVis*, pp. 69–75. IEEE Computer Society, Los Alamitos, 2000. doi: 10.1109/INFVIS.2000.885092

[8] J. Chuang, S. Gupta, C. D. Manning, and J. Heer. Topic model diagnostics: Assessing domain relevance via topical alignment. *Journal of Machine Learning Research: Workshop and Conference Proceedings*, 28(3):612–620, 2013.

[9] N. Coulter, I. Monarch, and S. Konda. Software engineering as seen through its research literature: A study in co-word analysis. *Journal of the American Society for Information Science*, 49(13):1206–1223, 1998. doi: 10.1002/(SICI)1097-4571(1998)49:13<1206::AID-ASI7>3.0.CO;2-F

[10] R. J. Crouser and R. Chang. An affordance-based framework for human computation and human-computer collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2859–2868, 2012. doi: 10.1109/TVCG.2012.195

[11] N. De Bellis. *Bibliometrics and Citation Analysis: From the Science Citation Index to Cybermetrics*. Scarecrow Press, Lanham, USA, 2009.

[12] N. K. Denzin and Y. S. Lincoln, eds. *Handbook of Qualitative Research*. Sage Publications, Inc., Thousand Oaks, London, New Delhi, 3rd ed., 2005.

[13] E. Geisler. *The Metrics of Science and Technology*. Quorum Books, Westport, CT, USA, 2000.

[14] C. Görg, Z. Liu, J. Kihm, J. Choo, H. Park, and J. Stasko. Combining computational analyses and interactive visualization for document exploration and sensemaking in Jigsaw. *IEEE Transactions on Visualization and Computer Graphics*, 19(10):1646–1663, 2013. doi: 10.1109/TVCG.2012.324

[15] C. D. Hansen and C. R. Johnson, eds. *The Visualization Handbook*. Elsevier, Amsterdam, 2005. doi: 10.1016/B978-012387582-2/50001-0

[16] Q. He. Knowledge discovery through co-word analysis. *Library Trends*, 48(1):133–159, 1999. doi: 2142/8267

[17] J. Heer and B. Shneiderman. Interactive dynamics for visual analysis. *Communications of the ACM*, 55(4):45–54, 2012. doi: 10.1145/2133806.2133821

[18] A. Hoonlor, B. K. Szymanski, and M. J. Zaki. Trends in computer science research. *Communications of the ACM*, 56(10):74–83, 2013. doi: 10.1145/2500892

[19] C.-P. Hu, J.-M. Hu, S.-L. Deng, and Y. Liu. A co-word analysis of library and information science in China. *Scientometrics*, 97(2):369–382, 2013. doi: 10.1007/s11192-013-1076-7

[20] P. Isenberg, F. Heimerl, S. Koch, T. Isenberg, P. Xu, C. Stolper, M. Sedlmair, J. Chen, T. Möller, and J. Stasko. Visualization publication dataset. Dataset: http://vispubdata.org/ , 2015. Published Jun. 2015.

[21] P. Isenberg and T. Isenberg. Visualization on interactive surfaces: A research overview. *i-com*, 12(3):10–17, 2013. doi: 10.1524/icom.2013.0020

[22] T. Isenberg, P. Isenberg, J. Chen, M. Sedlmair, and T. Möller. A systematic review on the practice of evaluating visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, 2013. doi: 10.1109/TVCG.2013.126

[23] X. Jiang and J. Zhang. A text visualization method for cross-domain research topic mining. *Journal of Visualization*, pp. 1–16, 2016. doi: 10.1007/s12650-015-0323-9

[24] Y. Joshi. Why do journals ask for keywords? Web site: http://www.editage.com/insights/why-do-journals-ask-for-keywords, 2014. Accessed June 2016.

[25] D. Kahneman. *Thinking, fast and slow*. Macmillan, 2011.

[26] M. C. Kim, Y. Zhu, and C. Chen. How are they different? A quantitative domain comparison of information visualization and data visualization (2000–2014). *Scientometrics*, 107(1):123–165, 2016. doi: 10.1007/s11192-015-1830-0

[27] H. Lam. A framework of interaction costs in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1149–1156, 2008. doi: 10.1109/TVCG.2008.109

[28] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, 2012. doi: 10.1109/TVCG.2011.279

[29] H. Lam, P. Isenberg, T. Isenberg, and M. Sedlmair, eds. *Proc. BELIV*. ACM, New York, 2014. doi: 10.1145/2669557

[30] J. Law, S. Bauin, J.-P. Courtial, and J. Whittaker. Policy and the mapping of scientific change: A co-word analysis of research into environmental acidification. *Scientometrics*, 14(3–4):251–264, 1988. doi: 10.1007/BF02020078

[31] M. Le Marc, J. P. Courtial, E. Drozda Senkovska, J. P. Petard, and Y. Py. The dynamics of research in the psychology of work from 1973 to 1987: From the study of companies to the study of professions. *Scientometrics*, 21(1):69–86, 1991. doi: 10.1007/BF02019183

[32] G.-Y. Liu, J.-M. Hu, and H.-L. Wang. A co-word analysis of digital library field in China. *Scientometrics*, 91(1):203–217, 2012. doi: 10.1007/s11192-011-0586-4

[33] Y. Liu, J. Goncalves, D. Ferreira, B. Xiao, S. Hosio, and V. Kostakos. CHI 1994–2013: Mapping two decades of intellectual progress through co-word analysis. In *Proc. CHI*, pp. 3553–3562. ACM, New York, 2014. doi: 10.1145/2556288.2556969

[34] B. H. McCormick, T. A. DeFanti, and M. D. Brown. Visualization in scientific computing (ViSC): Definition, domain and recommendations. *ACM SIGGRAPH Computer Graphics*, 21(6):i–x, 1–14, A1–A17, B1–B16, C1–C20, D1–D8, E1–E8, 1987. doi: 10.1145/41997

[35] F. Muñoz-Leiva, M. I. Viedma-del Jesús, J. Sánchez-Fernández, and A. G. López-Herrera. An application of co-word analysis and bibliometric maps for detecting the most highlighting themes in the consumer behaviour research from a longitudinal perspective. *Quality & Quantity*, 46(4):1077–1095, 2011. doi: 10.1007/s11135-011-9565-3

[36] H.-J. Schulz. Treevis.net: A tree visualization reference. *IEEE Computer Graphics and Applications*, 31(6):11–15, 2011. doi: 10.1109/MCG.2011.103

[37] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012. doi: 10.1109/TVCG.2012.213

[38] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proc. Visual Languages*, pp. 336–343. IEEE Computer Society, Los Alamitos, 1996. doi: 10.1109/VL.1996.545307

[39] A. Skupin, J. R. Briberstine, and K. Börner. Visualizing the topical structure of the medical sciences: A self-organizing map approach. *PLoS ONE*, 8(3):e58779, 2013. doi: 10.1371/journal.pone.0058779

[40] J. Stasko, J. Choo, Y. Han, M. Hu, H. Pileggi, R. Sadanaand, and C. D. Stolper. CiteVis: Exploring conference paper citation data visually. In *Posters of IEEE InfoVis*, 2013.

[41] J. Stewart. Precision conference solution. Web site: https://precisionconference.com/, 1999. Accessed June 2016.

[42] J. J. Thomas and K. A. Cook, eds. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE Computer Society, Los Alamitos, 2005.

[43] M. Tory and T. Möller. Human factors in visualization research. *IEEE Transactions on Visualization and Computer Graphics*, 10(1):72–84, 2004. doi: 10.1109/TVCG.2004.1260759

[44] K. Voegele. Annotated bibliography of the visualization conference proceedings. In *Proc. IEEE Visualization*, p. xxii. IEEE Computer Society, Los Alamitos, 1995. doi: 10.1109/VIS.1995.10003

[45] J. H. Ward. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301):236–244, 1963. doi: 10.1080/01621459.1963.10500845

[46] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, San Francisco, 2nd ed., 2004.

[47] J. Whittaker. Creativity and conformity in science: Titles, keywords and co-word analysis. *Social Studies of Science*, 19(3):473–496, 1989. doi: 10.1177/030631289019003004

[48] J. S. Yi, Y. a. Kang, J. Stasko, and J. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, 2007. doi: 10.1109/TVCG.2007.70515