





A Visual Analytics Approach for Analyzing Technological Trends in Technology and Innovation Management

Kawa Nazemi^(✉)  and Dirk Burkhardt 

Research Group on Human-Computer Interaction and Visual Analytics,
Darmstadt University of Applied Sciences, Darmstadt, Germany
{kawa.nazemi, dirk.burkhardt}@h-da.de

Abstract. Visual Analytics provides with a combination of automated techniques and interactive visualizations huge analysis possibilities in technology and innovation management. Thereby not only the use of machine learning data mining methods plays an important role. Due to the high interaction capabilities, it provides a more user-centered approach, where users are able to manipulate the entire analysis process and get the most valuable information. Existing Visual Analytics systems for Trend Analytics and technology and innovation management do not really make use of this unique feature and almost neglect the human in the analysis process. Outcomes from research in information search, information visualization and technology management can lead to more sophisticated Visual Analytics systems that involved the human in the entire analysis process. We propose in this paper a new interaction approach for Visual Analytics in technology and innovation management with a special focus on technological trend analytics.

Keywords: Visual Analytics · Information visualization · Trend analytics · Analysis approach · User-centered design

1 Introduction

Visual Analytics and information visualization are widely used to identify, detect and predict early technological trends for strengthening the competitiveness of enterprises. There exist a variety of methods for analyzing emerging or decreasing trends and predict possible future scenarios. From the visualization point of view, the human with his cognitive abilities plays an essential role in the analysis process. This aspect of so called human-in-the-loop approach, is crucial for the interaction design of every interactive visual application and in particular for complex analytical visualizations. Interactive visualizations for trend analytics and technology and innovation management commonly focus on processing data and integrating algorithms for detecting trends. The interaction capabilities are

limited and a real exploration and discovery is not supported by involving the human in the analysis process. On one the hand there exist a number of search and exploration approaches of interactive visualizations or in technology and innovation management system that include humans in the analytics loop. These systems commonly are either designed to solve general visualization tasks or enable humans to solve trend analytics tasks without visualizations. So there exist a gap between the given approaches coming from economics and the human- and task-centered approaches in information visualizations and Visual Analytics.

We propose in this paper a new approach for Visual Trend Analytics that in particular includes the economic tasks in a human-centered way by investigating the different perspectives on trend analytics. We therefore first introduce the state-of-art in trend visualization and interactive Visual Trend Analytics. Thereafter, we propose a general approach for Visual Trend Analytics that aims at bringing together the different viewpoints of economy, analytics, search, exploration, technology and innovation management with interactive visualization and will illustrate the way how this approach was applied to Visual Analytics system for technology and innovation management.

2 Literature Review

Current trend mining methods provide useful indications for discovering trends. Nevertheless, the interpretation and conclusion for serious decision making still requires the human and his knowledge acquisition abilities. Therefore, the representation of trends is one of the most important aspects for analyzing trends. Common approaches often include basic visualization techniques. Depending on the concrete results, line graphs, bar charts, word clouds, frequency tables, sparklines or histograms are utilized to impart different aspects of trends. *TheRiver* and *Tiara* represents thematic variations over time in a stacked graph visualization with a temporal horizontal axis [1]. The variation of the stream width indicates the strength of a specific topic over time [2]. *ParallelTopics* [3] includes a stacked graph for visualizing topic distribution over time. Although the system was not designed for discovering trends but rather for analyzing large text corpora, it allows users to interactively inspect topics and their strengths over time and thus allows the exploration of important trend indicators in the underlying text collection. *Parallel Tag Clouds* (PTC) is based on multiple word clouds that represent the contents of different facets in the document collection [4]. Temporal facets can be used to identify the difference of certain keywords over time and to infer the dynamics of themes in a text collection. Another extension of word clouds is *SparkClouds* that includes a sparkline for each word [5]. These sparklines indicate the temporal distribution of each term and allow conclusions about the topic trends. A user study reveals that participants are more effective with *SparkClouds* compared to three other visualization techniques in tasks related with trend discovery [5]. A similar approach [6] also includes co-occurrence highlighting. In contrast to *SparkClouds*, this technique includes a histogram for representing the temporal relevance of each tag. Additional overlays in the histograms show the co-occurrences over time for a selected word

to enable a more comprehensive analysis of trend indicators. Han et al. introduce with *PatStream* a visual trend analysis system for technology management [7]. Their system measures similarity between pairs of patents using the cosine metrics and extends the work of Heimerl et al [8] in particular in regards of visualization. The evolution and structure of topics that indicates the trends is visualized through a *Streamgraph* [1] using a river metaphor, it shows the evolution of topic-specific corpus structure.

Wei et al. [9] placed word clouds onto the streams to convey the topic content and Heimerl et al. [8] extended stream graphs to visually link topics in a scientific literature corpus with the cited communities to enable a joint analysis. In contrast to this previous works, *Patstream* breaks down the streams into vertical time slices, which represents periods of years. These time slices are based on their introduced concept that uses the term score, the ratio between the radiative frequency of a term in the given patent collection and its relative frequency in a general discourse reference corpus [7]. Although, their concept makes use of term frequencies, *title score* and *claims score* [7], the most useful approach seems to be the term score, thus it relies on a relative score and investigates the entire document or patent corpus. Beside the main visual representation, the stream visualization, they provide four further visual representation, such as a scatterplot with brushing and linking [7]. The most advanced interactive visual representation is *PatStream* [7]. It provides more than one view, makes use of relative scores and co-occurrences and visualize the temporal spread of the topics with the related categories. Furthermore, it provides a kind of process of functionalists to support trend analysis and technology management in particular for patents. They propose a five-step approach derived from the works of Ernst [10] and Joho et al. [11] that starts with (1) obtaining an overview of different technology topics in a given field, (2) identifying relevant trends according to individual information needs, (3) evaluating the importance of technological trends, (4) observing the behavior and productivity of different players relevant to specific trend, and (5) spotting new technologies related to a trend.

As the literature review revealed, there are already a number Visual Analytics systems for analyzing and visualizing trends. We could outline that existing systems do not really make use of the huge interaction capabilities of Visual Analytics and the human in the loop.

3 General Analysis Approach

Different approaches from information visualization, Visual Analytics, search behavior and from the area of technology and innovation management arose to support the analysis process in particular through visual interactive systems. One of the main works in this area is the “Visual Information Seeking Mantra” proposed by Shneiderman [12] with the basic principle of an interactive visualization to give first an overview, then zoom and filter functionalities and then details on demand [12]. In context of large graph-visual exploration to support users with their degree of interest, van Ham and Perer [13] proposed a more

bottom-up approach, starting with search, then showing the context of a certain graph and expanding this on demand. Although, this work is proposed for large graph exploration, the main idea can be adapted to a more user-centered approach and complement the basic principle of Shneiderman in particular by the “search” task. From the area of search and exploration, Marchionini [14] proposed an exploratory search approach based on Bloom’s taxonomy [15]. With three kinds of search activities, *Lookup*, *Learn*, and *Investigate*, he proposed that the activities are overlapping and searchers may be involved in more than one activity in parallel. *Lookup* is the lowest level while *Learn* is the next step of his search process model, which is assigned together with *Investigate* as exploratory activities. This search activity involves multiple iterations of searching and result evaluation to enhance the knowledge about a certain domain or topic. The most complex cognitive activity in the search process is *Investigate* that includes tasks as analysis, synthesis and evaluation. This search activity includes not only finding and acquiring new knowledge and information. It involves analytical tasks such as discovering gaps in the knowledge domain. In this process of knowledge discovery they [14, 16] construct knowledge by investigating various sources and ideas [17].

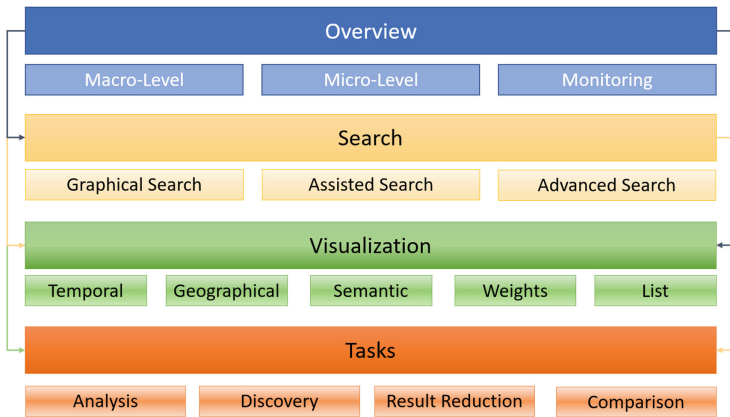


Fig. 1. General visual-interactive analysis approach for trend analytics.

From the viewpoint of technology and innovation management and trend analytics a number of processes arose that should enable strengthening the market position of enterprises through a well-defined analytics process. In this context, commonly patents are used as data corpus and as technology trigger. Bonino et al. [18] proposed that the tasks pursued by patent information users can be subdivided into the three main classes of *search*, *analysis* and *monitoring*. They correlated five *search* tasks to three main questions of *when*, *why* and *what*, and a “focus” that can be either specific or broad. Considering the introduced works the broad-focus can be seen as exploratory search [14]. Such a correlation matrix could be *portfolio survey* as search task, “when” is set to *business*

planning, “why” to identifying the technical portfolio of different players, “what” to patents, scientific and technical publications in a given technology area, and the focus is broad. Analysis is subdivided according to Bonino et al. into *micro*- and *macro-analysis*. Thereby micro-analysis involves a single (patent) document, while macro-analysis investigates a portfolio of documents (patents in their case) [18]. The analysis is usually performed to evaluate and assess the Intellectual Property (IP), to map and chart the IP, to identify trends and competitors and also to identify new areas of potential to exploit. The analysis task is therefore pretty similar to the definition of Marchionini’s exploratory investigation step [14]. The *monitoring* task is about keeping the users up to date about new upcoming (patent) information in the specified domain of interest [18]. Joho et al. [11] used the classification of Bonino et al. and investigated in an empirical study beside demographic aspects, question on how search tasks are performed, which search functionalities are perceived as important, and what the ideal (patent) search system is. They first derived from these questions “important system features” that were then roughly grouped into (1) *query formulation*, (2) *result assessment and result navigation*, and (3) *search management, organization and history*. *Query formulation* included aspects like Boolean search, query expansion, field operators etc., while *result assessment and result navigation* focused on highlighting, navigation and relevance score, and *search management, organization and history* included aspects of combining queries, search or navigation history and timeliness. In their evaluation they could find out that users of such analytics systems are willing to adopt and leverage functionality in contrast to web searchers [11]. Nazemi et al. [19], proposed the following main question in Visual Analytics systems for technology and innovation management: (1) *when* have technologies or topics emerged and when established? (2) “*where* are the key-players and key-locations, (3) *who* are the key-players, (4) *what* are the core-topics (4) *how* will the technologies or topics evolve, and (5) *which* technologies or topics are relevant for an enterprise?” [19, p. 4].

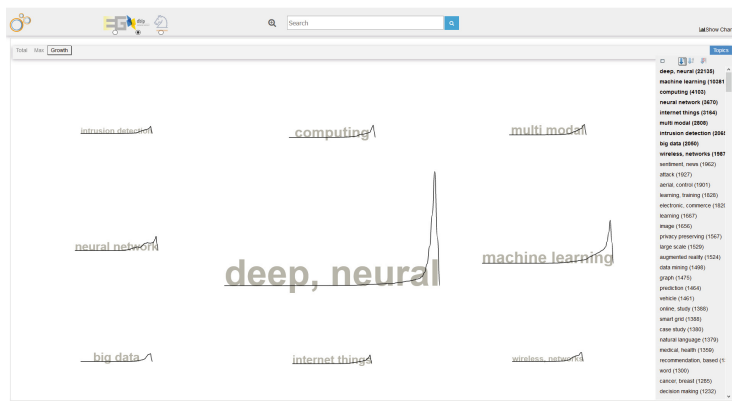


Fig. 2. Overview on macro-level: emerging trends of a certain data-set (in this example the entire DBLP database) at a glance.

Based on the above approaches and models, we derived a more general approach for investigating the entire analysis process for visual trend analytics. Figure 1 illustrates the four main steps of “Overview”, “Search”, “Visualization” and “Tasks”. The first two steps of *overview* and *search* can be assigned as the initial steps, a user is performing during the analysis process. So the first steps are either searching for a certain term or getting an overview of the data or the sub-set of data. The results are then visualized that enables solving different tasks. Our approach combines more abstract tasks like *analysis* or *discovery* with specific tasks that are commonly performed during the analysis process, e.g. *result reduction* or *comparison*. Although, these tasks are not at the same level of *analysis* or *discovery*, dedicated functionalities for the analysis process are required as the introduced models showed.

4 Overview

The *overview* step is based on the initial work of Shneiderman [12] and aims at giving an overview to entire data-set or a sub-set of data to gain an overall overview of the underlying data. We integrated three overview levels according to the work of Bonino et al. [18], namely overview on macro-level, on micro-level and for monitoring. In trend analytics the overview on macro-level should give an initial overview of emerging trends [20] out of the entire data base. Figure 2 illustrates such an overview on macro-level. Thereby the emerging trends gathered through topic modeling is illustrated as *SparkCloud* [5] and the users are able to see the most emerging trends at a glance. On the right side, other topics of interest can be selected. The user can further choose an overview of the most appeared topics in the entire data-set or the topics with the highest climax. The overview on macro-level is interactive, so that the user is able to get details-on-demand with one click on a certain SparkCloud.

Overview on micro-level in trend analytics contains the temporal spread of topics (trends) for a certain key-term that is either searched or selected from the macro-level overview. It gives an overview of all related topics in a temporal manner and insights of the main technologies and approaches for a certain key-term. Figure 3 illustrates one example of an overview on micro-level. The user has chosen the term “Information Visualization” from the overview on macro-level and gets the temporal spread of the related technologies and approaches ranked based on the frequency of the topics in related documents.

For monitoring as overview, we integrated a personalized word cloud that illustrates the most search or selected terms for a single person. Thereby a simple user model with a *bag-of-words* approach is implemented and enables the user to select the terms that wants to monitor. Figure 4 illustrates the word-cloud for two different persons and illustrates clearly that the amount of monitored terms can vary significantly.

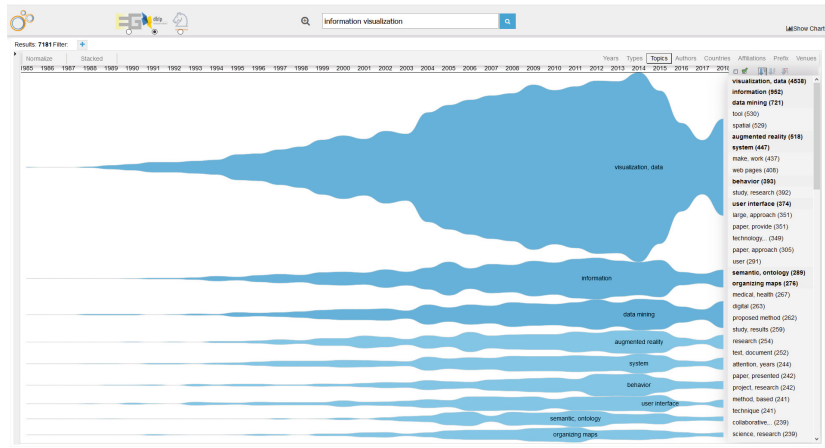


Fig. 3. Overview on micro-level: temporal spread of related topics for a single key-term (in this example for information visualization).



Fig. 4. Overview - monitoring: personalized word-clouds of two different persons.

5 Search

Search functionalities are essential in particular if the amount of the underlying data are large. We have integrated different search approaches to assist users in the search process. A novel approach is the “graphical search” functionality (see Fig. 5). Thereby the initial search term (or selected term) is visualized as a circle in the center of the screen. Users are able to define further search terms as “points-of-interest” (POI), whereas each POI is visualized as a small circle with a certain color. These POIs can be dragged into the initial search circle that illustrates the amount of results including both search terms. The graphical search functionality further enables a nested search, so that more than two circles can be dragged. With this action, the amount of the results is always the amount of all nested POIs.

The assisted search was applied according to Nazemi et al. [19] and extends the search functionalities beside traditional linguistic methods with a topic-based approach. The approach incorporates the information of the generated topics to

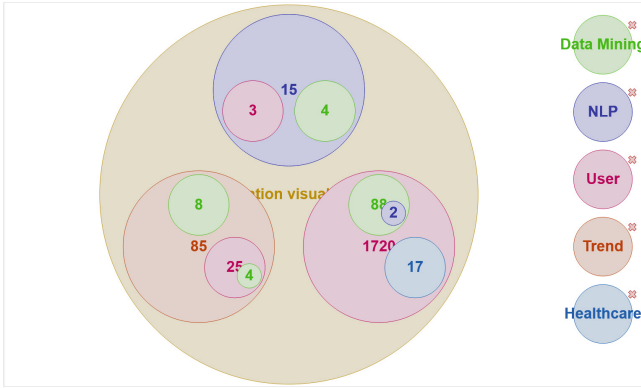


Fig. 5. Search - graphical search: the initial search-term is represented as a circle, users can define further search terms and drag them into the initial search-term to get a nested number of combined search.

enhance search-terms from the query. The topics provide N-Grams, which are the most used phrases within each topic. These phrases often represent different ways to articulate the idea of the topic and consequently can be used as a key phrase to represent the topic. We use data from the phrases and choose the top five most used phrases as additional search-terms to extend the result set. Based on the initial users' search the most dominant topic in the result set is identified. In the next step, semantically similar phrases are extracted based on the identified dominant topic. Additional searches are performed by the system using the semantically similar phrases as search-terms. The functionality can be activated with one click on top next to the search field. The advanced search is detecting automatically all data rows in the database and provides dedicated search in those tables.

6 Visualization

We integrated a set of different visualizations in particular to respond to the questions according to Nazemi et al. [19]: (1) *when* have technologies or topics emerged and when established? (2) *where* are the key-players and key-locations, (3) *who* are the key-players, (4) *what* are the core-topics (4) *how* will the technologies or topics evolve, and (5) *which* technologies or topics are relevant for an enterprise?" [19, p. 4]. To address these questions, we integrated different data models that are the foundation of the underlying visual structures. The visualizations are automatically detecting the data model and visualize the data in appropriate manner. A semantic data model serves as the primary model for all other data models. The visualization step includes a set of temporal, geographical, semantic and topic (weights) visual layouts that are all interactive and enable solving different tasks. A simple list-view is added to select single entities. Figure 6 illustrates examples of the integrated visualizations that lead to solve

different analytical tasks. The visualizations can be used as single interactive visual interface or be combined in dashboards.

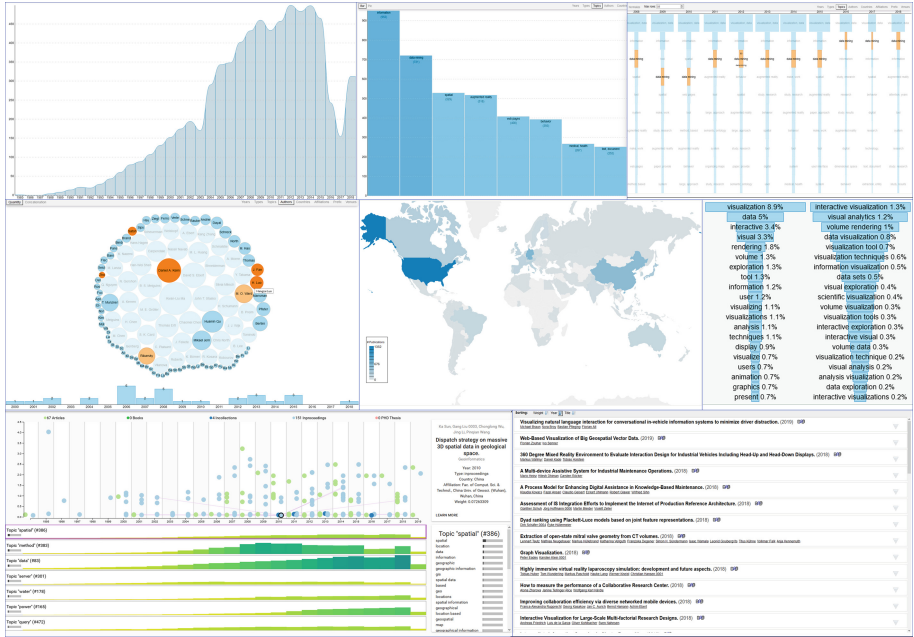


Fig. 6. Visualization: examples of integrated temporal, geographical, semantic, weight and list visualizations.

7 Tasks

The visualizations and the overall user interface enables user to solve different tasks. The tasks vary in every domain and is strongly dependent to the users of a visual system. The analysis tasks in our approach responds mainly to the questions according to Nazemi et al. [19] (see Sect. 6). Whereas aspects of predicting technologies, detecting emerging trends [20] and strengthening the market position of enterprises plays an essential role. The analysis tasks incorporates solving a complex analytical task with a specific goal in order to strengthen the potentials of an enterprise or other institution. The discovery tasks aims at detecting unexpected patterns, topics, technologies or correlations in data. This tasks can be best supported by aspect-oriented visualizations, where the user gets insights from different perspectives and detects a new correlation. Result reduction as first main task is essential for every research task, e.g. researching for related technologies or competitors. To support this tasks, we integrated a dynamic faceting that is further supported by each of the visual layouts. So the user can explore starting a huge amount of data related to a certain field and reduce the amounts based on his requirements.

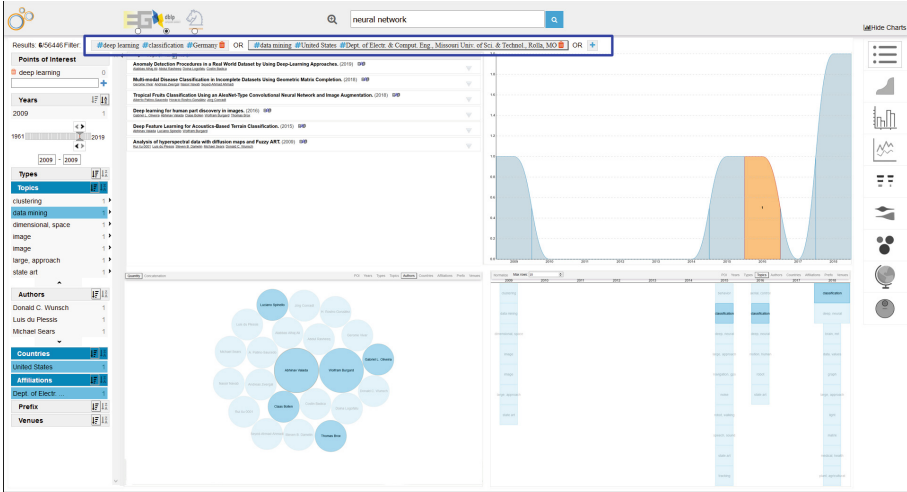


Fig. 7. Tasks: result reduction.

Figure 7 illustrates the user interface of our application including the faceting and visual reduction. The user started with a search-term with about 60,000 results (top left of the screen). He used the graphical search, faceting and visual interactions to reduce the amount to only six relevant papers and chose four visualization to see the results. The faceting is visualized on left and the result reduction is on top (highlighted with a blue rectangle).

The comparison as second main task can be performed on two different levels: Comparison on data-subset with two same visual layouts or comparison of different databases. This differentiation is similar to the proposed micro-level and macro-level overview. Based on the requirements, a user is able to set up a campaign and compare the entire data stored in different database. This way of comparison enable to compare for example enterprise data with scientific publications and see the results at a glance. Figure 8 illustrates a comparison on macro-level through different databases and the visual layout is a macro-level overview too, so that he is able to see the emerging trends in two different databases.

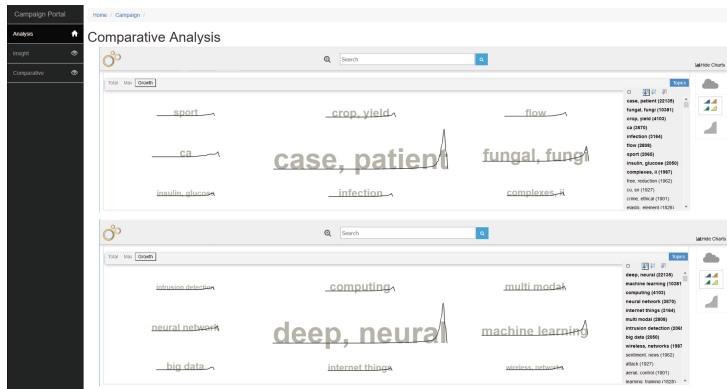


Fig. 8. Tasks: comparison on macro-level by using macro-level overview visual layouts.

8 Conclusion

We proposed in this paper an approach for Visual Trend Analytics that included the economic tasks in a human-centered way. We therefore first introduced the state-of-art in trend visualization and interactive visual trend analytics. Thereafter, we proposed a general approach for visual trend analytics that aimed at bringing together the different viewpoints of economy, analytics, search, exploration, technology and innovation management with interactive visualization. In this context the different approaches was be described let us derive a general approach for Visual Analytics in technology and innovation management. We applied and illustrated each step of our approach with a Visual Analytics system for technology and innovation management with two main contributions: a general approach that brings the different viewpoints together to enhance the state-of-the-art in visual trend analytics and the application of the approach to a complex Visual Analytics system for detecting early trends in technology an innovation management.

Acknowledgments. This work was partially funded by the Hessen State Ministry for Higher Education, Research and the Arts within the program “Forschung für die Praxis” and was conducted within the research group on Human-Computer Interaction and Visual Analytics (<https://vis.h-da.de>). The presentation of this work was supported by the Research Center for Digital Communication & Media Innovation of the Darmstadt University of Applied Sciences.

References

1. Havre, S., Hetzler, E., Whitney, P., Nowell, L.: ThemeRiver: visualizing thematic changes in large document collections. *IEEE TVCG* **8**(1), 9–20 (2002). <http://dx.doi.org/10.1109/2945.981848>

2. Liu, S., et al.: TIARA: interactive, topic-based visual text summarization and analysis. *ACM Trans. Intell. Syst. Technol.* **3**(2), 25:1–25:28 (2012)
3. Dou, W., Wang, X., Chang, R., Ribarsky, W.: ParallelTopics: a probabilistic approach to exploring document collections. In: *VAST 2011* (2011)
4. Collins, C., Viegas, F., Wattenberg, M.: Parallel tag clouds to explore and analyze faceted text corpora. In: *VAST 2009* (2009)
5. Lee, B., Riche, N.H., Karlson, A.K., Carpendale, S.: SparkClouds: visualizing trends in tag clouds. *IEEE TVCG* **16**, 1182–1189 (2010)
6. Lohmann, S., Burch, M., Schmauder, H., Weiskopf, D.: Visual analysis of microblog content using time-varying co-occurrence highlighting in tag clouds. In: *Proceedings of the International Working Conference on Advanced Visual Interfaces, AVI 2012*, pp. 753–756. ACM, New York (2012)
7. Han, Q., Heimerl, F., Codina-Filba, J., Lohmann, S., Wanner, L., Ertl, T.: Visual patent trend analysis for informed decision making in technology management. *World Patent Inf.* **49**, 34–42 (2017)
8. Heimerl, F., Han, Q., Koch, S., Ertl, T.: CiteRivers: visual analytics of citation patterns. *IEEE Trans. Vis. Comput. Graph.* **22**(1), 190–199 (2016)
9. Wei, F., et al.: TIARA: a visual exploratory text analytic system. In: *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2010*, pp. 153–162. ACM, New York (2010)
10. Ernst, H.: Patent information for strategic technology management. *World Patent Inf.* **25**(3), 233–242 (2003)
11. Joho, H., Azzopardi, L.A., Vanderbauwhede, W.: A survey of patent users: an analysis of tasks, behavior, search functionality and system requirements. In: *Proceedings of the Third Symposium on Information Interaction in Context, IiX 2010*, pp. 13–24. ACM, New York (2010)
12. Shneiderman, B.: The eyes have it: a task by data type taxonomy for information visualizations. In: *VL*, pp. 336–343 (1996)
13. van Ham, F., Perer, A.: Search, show context, expand on demand: supporting large graph exploration with degree-of-interest. *IEEE Trans. Vis. Comput. Graph.* **15**, 953–960 (2009)
14. Marchionini, G.: Exploratory search: from finding to understanding. *Commun. ACM* **49**(4), 41–46 (2006)
15. Bloom, B.S.: *Taxonomy of Educational Objectives*. David McKay Co., Inc., New York (1956)
16. White, R.W., Roth, R.A.: *Exploratory Search: Beyond the Query-Response Paradigm*. Synthesis Lectures on Information Concepts, Retrieval, and Services. Marchionini, G. (ed.), vol. 1. Morgan & Claypool Publishers (2009)
17. Bruner, J.S.: The act of discovery. *Harvard Educ. Rev.* **31**, 21–32 (1961)
18. Bonino, D., Ciaramella, A., Corno, F.: Review of the state-of-the-art in patent information and forthcoming evolutions in intelligent patent informatics. *World Patent Inf.* **32**(1), 30–38 (2010)
19. Nazemi, K., Retz, R., Burkhardt, D., Kuijper, A., Kohlhammer, J., Fellner, D.W.: Visual trend analysis with digital libraries. In: *Proceedings of the 15th International Conference on Knowledge Technologies and Data-driven Business - i-KNOW 2015*. ACM Press (2015). <https://doi.org/10.1145/2809563.2809569>
20. Nazemi, K., Burkhardt, D.: Visual analytics for analyzing technological trends from text. In: *2019 23rd International Conference Information Visualisation (IV)*, pp. 191–200. IEEE (2019)