The State of the Art in Sentiment Visualization

Kostiantyn Kucher¹, Carita Paradis² and Andreas Kerren¹

¹Department of Computer Science, Linnaeus University, Växjö, Sweden
{kostiantyn.kucher, andreas.kerren}@lnu.se
²Centre for Languages and Literature, Lund University, Sweden
carita.paradis@englund.lu.se

Abstract
Visualization of sentiments and opinions extracted from or annotated in texts has become a prominent topic of research over the last decade. From basic pie and bar charts used to illustrate customer reviews to extensive visual analytics systems involving novel representations, sentiment visualization techniques have evolved to deal with complex multidimensional data sets, including temporal, relational and geospatial aspects. This contribution presents a survey of sentiment visualization techniques based on a detailed categorization. We describe the background of sentiment analysis, introduce a categorization for sentiment visualization techniques that includes 7 groups with 35 categories in total, and discuss 132 techniques from peer-reviewed publications together with an interactive web-based survey browser. Finally, we discuss insights and opportunities for further research in sentiment visualization. We expect this survey to be useful for visualization researchers whose interests include sentiment or other aspects of text data as well as researchers and practitioners from other disciplines in search of efficient visualization techniques applicable to their tasks and data.

Keywords: sentiment visualization, text visualization, sentiment analysis, opinion mining

ACM CCS: • Information systems → Sentiment analysis; • Human-centred computing → Visualization techniques; • Human-centred computing → Visual analytics; • Human-centred computing → Information visualization; • Human-centred computing → Visualization systems and tools

1. Introduction
The development of digital technologies and the Internet has led to an unprecedented increase of text data, creating new opportunities and challenges. Researchers in linguistics and natural language processing (NLP) have access to the data that can be drastically different from traditional corpora or literature with regard to content, scale, and corresponding analyses, for instance, customer reviews or social media messages. One of the research problems investigated by these disciplines is sentiment analysis (the term being often interchangeable with opinion mining and affect analysis), which is generally concerned with detecting attitudinal content in text at various levels of granularity [PL08]. Usually, the text is classified into positive, negative or neutral at the level of words, utterances or complete documents.

Sentiment visualization—understood as a research challenge in information visualization (InfoVis) and visual analytics (VA) to analyze sentiment discovered in text data—is part of the more general research area of text visualization [KK15]. The applications and tasks of sentiment visualization include, for instance, monitoring of public opinion in social media, literature analysis for digital humanities, or support for research of sentiment and stance in linguistics and NLP. Some of the earliest papers mentioning visualization of sentiment actually originate in data mining (DM) or NLP and use basic visual representations in most cases. On the contrary, state-of-the-art techniques often reflect the advances in InfoVis and VA and incorporate sentiment in complex settings involving heterogeneous data. However, existing work in sentiment visualization has not been covered by any comprehensive survey yet, as discussed in Section 3. Therefore, our survey can be beneficial for visualization researchers working on this problem since there is already a significant body of work involving sentiment visualization which can be difficult to explore—the publications are scattered across a large number of outlets and disciplines. Our survey can also be useful for researchers from other fields as well as practitioners interested in visualization/visual analysis methods for sentiment data.

© 2017 The Authors
Computer Graphics Forum published by John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
In this paper, we propose a survey based on the collection and analysis of a substantial number of sentiment visualization techniques described in peer-reviewed papers in InfoVis, VA and other disciplines (NLP, DM, etc.). The main scientific contributions of this paper are the following:

- an overview of a large set of sentiment visualization techniques introduced in multiple research disciplines during the past 15 years (as illustrated by Figure 1);
- a fine-grained categorization that provides evidence about data, tasks and visualization aspects supported by existing techniques;
- an interactive survey browser of sentiment visualization techniques (see below);
- an investigation of existing trends and correlations between categories based on temporal and correlation analyses; and
- suggestions and recommendations for future work in sentiment visualization.

To refine our categorization, discover interesting patterns and facilitate data exploration by the readers of this paper, we have developed an interactive survey browser available at

http://sentimentvis.lnu.se

In this survey, we have limited ourselves only to visualization techniques based on the analysis of text data, and have not included visualization techniques related to emotion measurement with the help of brain–computer interfaces (as opposed to emotions discovered in text) or similar approaches. We refer the interested readers to the recently published survey by Cernea and Kerren [CK15] that covers the corresponding research area. Another related discipline that concerns itself with analysis and, in some cases, visualization of opinions is social network analysis. For example, Du et al. [DYLL15] discuss OpinionRings, a visualization technique for networks with explicit user opinion values. Since such approaches do not involve text data, we have considered them to be beyond the scope of this survey.

The rest of this paper is organized as follows. In the next section, we describe the background of sentiment analysis and related research problems from linguistics and computational linguistics/NLP. Section 3 provides a discussion of existing visualization surveys relevant to our work. Afterwards, we discuss our methodology and initial statistical results for the collected data in Section 4. Our categorization and the corresponding sentiment visualization techniques are discussed in Section 5. We discuss our interactive survey browser, findings and perspectives for the research field in Section 6. Finally, we conclude this paper in Section 7.

2. Background

With the advent of machine-readable corpora, research in linguistics has shown that language is not primarily a means of providing information about facts, but rather to evaluate what we are talking about, to take a stance and to express opinions and emotions. Language use in different contexts is highly viewpointed, interactive and interpersonal. Human communication has a purpose. It is in a constant flux and so is the use of language itself [Eng07, Par15].

Evaluative meanings are not easy to specify in advance because they are not confined to traditional areas of grammar or specific words, but may be expressed by parts of words, words or longer
The term sentiment analysis as it is used in NLP is usually defined as the task of classifying (short) pieces of text (ranging from single words over phrases and sentences to complete documents) into a small number of classes representing different kinds of sentiments (the term sentiment is used here and below synonymously with terms like emotion, affect, attitude, and so on, unless a more specific term is required). In the simplest formulation, sentiment analysis is considered as a binary problem, where we are either interested in detecting the presence of emotionally loaded content, or in distinguishing positively from negatively loaded content. The former of these tasks is closely related to what has been referred to as subjectivity detection [PL08], while the latter is sometimes referred to as polarity detection [Tur02]. Classification itself is usually based on lexical matching of keywords from a previously constructed dictionary/lexicon (such as WordNet-Affect [SV04], MPQA Subjectivity Lexicon [WHH05], SentiWordNet [BES10], LIWC [TP10], or SenticNet [COR14]), knowledge about word/concept similarity (e.g. using a distributional semantics model or an ontology such as WordNet [Mil95]), or a variety of machine learning (ML) classification models.

Arguably, the most common approach to sentiment analysis is to formulate the problem as a three-way categorization task over the categories negative, neutral and positive. More complex formulations of the sentiment analysis task involve a broader range of possible sentiment classes, either in terms of a graded scale (e.g. weakly to strongly negative and positive) [SPW*13], or in terms of a broader palette of sentiment types [SPH*11, He12, ADAC*13].

Sentiment analysis has also been combined and integrated with other NLP and ML techniques such as topic detection and tracking (TDT), in which case we are interested not only in which sentiments are expressed in the data, but also what is the target of the sentiment. As an example, a sentiment analysis system might detect that customers are predominantly negative to the release of a novel product. However, it would be valuable for the product manufacturer to know if there are any specific aspects of the product that are more negatively perceived—it might be the case that the negativity only concerns one specific aspect of the product, in which case it might be reasonably easy for the company to make the necessary adjustments. Such analysis is commonly referred to as aspect-based sentiment analysis [BE10, JO11].

Another closely related task is emotion detection, which typically employs a categorical, dimensional, or hybrid model of emotions. One of the most well-known categorical models is Ekman’s ‘Big Six’ basic emotions: anger, fear, happiness, surprise, disgust and sadness [Ekm92]. Dimensional models describe emotions in terms of continuous spaces along axes such as valence/pleasure, arousal and dominance [RM77]. Finally, hybrid models such as the Plutchik’s wheel of emotions [Plu80] define a set of basic emotions, their relative similarity, and possible combinations resulting in numerous derivative emotions. The relation between emotions and sentiment is analyzed by Munezero et al. [MMSP14], who discuss the definitions of sentiment, opinion, emotion, and affect in detail from the standpoint of linguistics and psychology. They point to subtle differences between these concepts which are often overlooked in work originating from more technical fields such as InfoVis, VA, or DM.

In general, the variation in terminology and computational models of sentiment analysis are covered by several survey papers—the arguably most comprehensive one is that of Pang and Lee [PL08], who discuss work done in linguistics, NLP, DM and ML. More recent survey papers include the works by Tsytserau and Palpanas [TP12], Cambria et al. [CSXH13], Ravi and Ravi [RR15], and a comprehensive survey by Mohammad [Moh16].

Based on this discussion, we can summarize the general properties and challenges of sentiment visualization as follows:

- sentiment visualization covers a variety of sentiment analysis tasks ranging from subjectivity detection to emotion analysis and stance analysis;
- sentiment visualization techniques may have to use data specific to the sentiment analysis model (e.g. lexicon-based or ML-based) and scope (word-level, utterance-level, etc.); and
- sentiment visualization reflects the variety of data domains and user tasks existing in research and applications of sentiment analysis, which range from theoretical research in linguistics and NLP to social media and news monitoring, thus implying the usage of various visual channels and representations.

3. Related Work

Sentiment visualization has not enjoyed the same level of interest in systematic/comprehensive reviews compared to other visualization areas that are also related to data extracted from text, such as the visualization of topic models or events. Only a few text visualization surveys include analysis and visualization of sentiment/opinion as one of their categorization aspects. An example is the survey on the visual analysis of events in text data streams written by Wanner et al. [WSJ*14]. They select polarity extraction as one of the text processing methods used in visual analytics systems: it was utilized by 14 out of 51 papers included in their report. Several surveys mention sentiment and affect analysis as a potential feature extraction method for text visualization; for instance, the paper by Risch et al. [RKFW08] in the context of visual analytics (without any examples) or the paper by Šilić and Bašić [ŠB10] in connection with text stream visualization (with a single example). Given the total amount of work on sentiment visualization, there is clearly a gap in this area in the visualization survey literature.

To the best of our knowledge, there are only two existing works that contain surveys of sentiment visualization techniques. Boumaiza [Bou15] aims to provide a survey of sentiment and opinion visualization techniques with the focus on social media texts. The author does not provide clear criteria and a clear categorization/taxonomy, though. In consequence, numerous techniques are included that are not directly related to sentiment visualization (or even visualization per se); this makes it difficult to navigate the survey and use it for reference. By our estimate, that work
mentions around 35 peer-reviewed sentiment visualization techniques. Shamim et al. [SBT15] provide an overview of 11 techniques and classify them according to visual metaphor. However, the focus of that work is not on categorization, but rather on evaluation: the authors conduct a study to compare the techniques with regard to metrics such as user-friendliness or usefulness. In contrast, our survey focuses on the categorization of a much larger number of techniques with regard to multiple aspects related to computational model, data, user tasks, and visual representation.

4. Methodology

The steps that we took while working on this survey are summarized in Figure 2. The overall methodology can be compared to the model described by Pirolli and Card [PC05], which was adapted to scientific literature analysis by Beck et al. [BKW16].

Based on our previous work on TextVis Browser [KK15], we started with an initial set of text visualization techniques related to sentiment as well as an initial categorization applicable to such techniques. We should state that we use a technique as a unit for this survey as opposed to a publication—therefore, we describe several cases below where multiple techniques originate from the same publication. Since our survey includes work not only from InfoVis, but also from VA and even non-visualization disciplines, a single technique does not necessarily mean a novel metaphor/representation, but also an approach or a system relevant to sentiment visualization.

In addition to the initial set of techniques, we have conducted a search in several visualization outlets: IEEE TVCG, Information Visualization, Computer Graphics Forum, IEEE CG&A, and Journal of Visualization as well as proceedings of IEEE InfoVis, IEEE VAST, EuroVis, IEEE PacificVis, TextVis workshop, ACM CHI, ACM IU, IV, IVAPP, and VINCI. We have also conducted a search in IEEE Xplore, ACM DL, and Google Scholar using such key phrases as ‘sentiment visualization’, ‘emotion visualization’, and ‘opinion visualization’ (considering only literature in English). Finally, we have investigated references from related surveys as well as already detected research publications.

Selection criteria We have used the following criteria for including/excluding techniques in our survey:

- a technique must be related to visualization of sentiment associated with text data (either extracted automatically or annotated manually);
- a technique must be illustrated by at least a single figure in the corresponding publication;
- a technique must be described in a peer-reviewed publication (including poster papers and extended abstracts);
- since we are focusing on techniques as opposed to publications, incremental work in several papers by the same authors is not considered as separate techniques; and
- a technique must actually involve an implementation used for visual representation or analysis (possibly even without interactive features) as opposed to figures generated with third-party tools solely for illustrative purposes in the respective publication.

Some of the candidate techniques had to be excluded with regard to the criteria above. Brath and Banissi [BB15] discuss the usage of text layout and font attributes for the purposes of text-related data visualization and mention sentiment analysis as one of the possible applications, however, the sentiment values are not directly visualized by their technique. OpinionRings by Du et al. [DYLL15] use network data explicitly labelled with user opinions as opposed to extracting the data from text—we have also not included Opinion Space by Faridani et al. [FBRG10] for the same reason. Li et al. [LDS10] mention a visualization module used in their opinion mining system, but provide no further details or figures, so it is

Figure 2: Flowchart describing our methodology for this survey. Colour coding is used to denote the corresponding concept/artefact/activity (yellow for visualization techniques, green for categorization, orange for the article, and white for miscellaneous), and bold font is used to denote activities. Interactive analysis of the techniques data, subsequent discussions, and corresponding categorization changes together form a sensemaking loop.
The categorization of aspects (highlighted in yellow) vary from general to specific (left to right). These 7 aspects, or groups, include the total of 35 categories listed in Table 2. The categorization facilitates the search for visualization techniques and corresponding publications for interested readers based on their data, required analytic and visualization tasks, and even specific encodings used for sentiment. In this section, we discuss the individual categories and corresponding prominent examples; for full details on each technique’s categorization, see the summary tables or our interactive browser (both discussed in Section 6).

### 5.1. Data aspects

Every sentiment visualization technique relies on certain data, and both higher-level (for instance, the original domain) and lower-level (e.g., source representation) data aspects affect the later stages of the pipeline.

#### 5.1.1. Data domain

Most sentiment visualization techniques are designed with a specific data domain in mind, some of which have been historically or currently associated with sentiment analysis and opinion mining tasks in visualization, NLP, and DM communities.

A prominent example of such a domain that provides a lot of text data suitable for sentiment analysis is online social media including forums, blogs, microblogs, and social networks. It is no surprise that the majority of techniques in our data (around 62%) support this category. Some of the early examples include MoodViews by Mishne and de Rijke [MDR06], an NLP system for affect analysis (‘mood’) of blogs which uses explicit mood tags provided by users as well as predicted mood levels for at least 132 mood types proposed by the blog platform. The system applies basic line plots for temporal visualization and allows the users to interactively investigate salient terms and phrases for selected time intervals. Ink Blots by Abbasi and Chen [AC07] is a technique for exploration of documents and corpora that uses a simple bubble metaphor to mark regions of interest in communication and social media (i.e., forums) texts. With the development of microblogs, more and more techniques started to focus on this data.

Diakopoulos et al. [DNKS10] describe Vox Civitas (see Figure 4(c)), a visual analysis tool for investigation of sentiment, relevance, and salient keywords in social media posts related to public events, and social media (i.e., forums) texts. With the development of microblogs, more and more techniques started to focus on this data. Diakopoulos et al. [DNKS10] describe Vox Civitas (see Figure 4(c)), a visual analysis tool for investigation of sentiment, relevance, and salient keywords in social media posts related to public events, which uses a pixel-based stacked bar timeline to represent sentiment values. TwitInfo by Marcus et al. [MBB*11] is a visual event analysis system for Twitter streams which supports sentiment polarity detection. The aggregated polarity value is visualized with a pie chart, and the polarity for individual documents is encoded as colour of the corresponding markers in a geographical map view, as depicted in Figure 3.
The complete categorization used in our survey. Each row contains the number of corresponding visualization techniques, which is also illustrated by heatmap-style icons.

| Data domain          | Analytic tasks                                      | Visual variable | Visual representation |
|----------------------|-----------------------------------------------------|----------------|-----------------------|
| Online Social Media  | 127 - Polarity Analysis / Subjectivity Detection  | 132 - Color    | 132 - Line Plot / River |
| Communication        | 14 - Opinion Mining / Aspect-based Sentiment Analysis |               | 132 - Pixel/Area/Matrix |
| Reviews / (Medical)  | 82 - Emotion/Affect Analysis                        |               | 132 - Clouds/Galaxies |
| Literature/Poems     | 33 - Stance Analysis                                | 107 - Position/Orientation | 57 - Node-Link |
| Scientific Articles/Papers | 4 - Region of Interest          | 84 - Size/Area | 23 - Maps              |
| Editorial Media      | 31 - Clustering/Classification/Categorization       | 32 - Shape     | 35 - Text              |
|                      |                                                    | 77 - Overview  | 36 - Glyph/Icon         |
|                      |                                                    | 120 - Comparison|                      |
|                      |                                                    | 21 - Monitoring |                      |
|                      |                                                    | 87 - Navigation/Exploration |          |
|                      |                                                    | 9 - Uncertainty Tackling |          |
|                      |                                                    |                |                      |
| Data source          | Region of Interest                                  |                |                      |
| Document             |                                                    |                |                      |
| Corpora              |                                                    |                |                      |
| Streams              |                                                    |                |                      |
|                      | 132 - Node-Link                                     |                |                      |
|                      | 20 - Overview                                       |                |                      |
|                      | 114 - Comparison                                    |                |                      |
|                      | 25 - Monitoring                                     |                |                      |
|                      | 29 - Overview                                       |                |                      |
|                      | 82 - Navigation/Exploration                         |                |                      |
|                      | 23 - Uncertainty Tackling                           |                |                      |
|                      | 132 - Node-Link                                     |                |                      |
|                      | 132 - Line Plot / River                             |                |                      |
|                      | 57 - Node-Link                                       |                |                      |
|                      | 120 - Comparison                                    |                |                      |
|                      | 21 - Monitoring                                     |                |                      |
|                      | 87 - Navigation/Exploration                         |                |                      |
|                      | 9 - Uncertainty Tackling                            |                |                      |
|                      | 99 - Geospatial                                     |                |                      |
|                      | 29 - Text                                           |                |                      |
|                      | 82 - Time Series                                    |                |                      |
|                      | 23 - Networks                                       |                |                      |
|                      | 82 - Location                                        |                |                      |
|                      | 23 - Networks                                       |                |                      |

Table 2: The complete categorization used in our survey. Each row contains the number of corresponding visualization techniques, which is also illustrated by heatmap-style icons.

in Figure 5(a). Other examples for this category include BLEWS by Gamon et al. [GBB*08], ForAVis by Wanner et al. [WRK11], webLYzard by Scharl et al. [SHHW*12], MoodLens by Zhao et al. [ZDWX12], Whisper by Cao et al. [CLS*12], Opinion-Flow by Wu et al. [WLY*14], PEARL by Zhao et al. [ZGZW14], SentiCompass by Wang et al. [WSK*15], uVSAT by Kucher et al. [KSBK*16], and many more.

Communication such as emails and chats can also be subject to sentiment analysis and visualization (used by 11% of techniques). The earliest example in our data is CrystalChat by Tat and Carpendale [TC06], a visualization technique for personal chat history that represents individual messages as circles and organizes them in a 3D layout with regard to temporal order and chat contact. The technique encodes the emotional content of conversations based on detected emoticons as the background plane colour. Another technique for communication texts is described by Gobron et al. [GAP*10], and it is a rather unusual emotion visualization technique bordering on computer graphics rather than InfoVis. The authors use the Facial Action Coding System (FACS) to create animated 3D avatars whose faces convey the emotions related to the corresponding text data. Chen et al. [CFKA14] conduct visual analysis of chat logs using emoticons classified by valence/polarity. Mail data is supported by Ink Blots by Abbasi and Chen [AC07], the work of Mohammad [Moh12], 5W Summarization by Das et al. [DBG12], the technique by Guzman [Guz13], and TargetVue by Cao et al. [CSL*16].

Before the rise of social media, sentiment analysis was almost solely used to analyze product reviews and customer feedback. The category Reviews / (Medical) Reports is supported by 25% of sentiment visualization techniques, including the ones such as Affect Inspector by Subasic and Huettner [SH01] which uses star plots to visualize affect profiles of text documents, including movie reviews. Pulse by Gamon et al. [GACOR05] uses a tree map to represent a clustering of sentence-level sentiment classification of car reviews (see Figure 4(a)). Opinion Observer by Liu et al. [LHC05] provides a visualization of customer opinions using modified bar charts. AMAZING by Miao et al. [MLD09] depicted in Figure 4(b) is an opinion mining system for product reviews that visualizes NLP processing results with a line chart (using the review timestamps) and a pie chart (a simple summary for the proportion of positive/negative reviews). In general, a lot of techniques in this category originate from NLP and DM rather than the visualization community, focus mostly on the analytical part, and use rather simple visual representations and interactions. In contrast, Chen et al. [CISSW06] use multiple analytical and visualization techniques for investigation of conflicting opinions in customer reviews. Hao et al. [HRJ*13] present four novel visualization techniques for customer feedback analysis: pixel sentiment geo map, key term geo map, pixel-based sentiment calendar, and self-organizing term association map. Some other examples for this category include map of opinion clusters introduced in the work by Oelke et al. [OHR*09], OpinionSeer by Wu et al. [WWL*10], and comparative relation maps by Xu et al. [XLLS11].

Literature/Poems is a surprisingly underrepresented domain with regard to sentiment visualization techniques, which includes only four entries or 3% of our current data set. Subasic and Huettner [SH01] demonstrate how their Affect Inspector can be applied for affect analysis of a poem by T. S. Eliot. Affect Colour Bar by Liu et al. [LSL03] uses a pixel-based metaphor to represent the emotions associated with text document segments. The resulting visualization provides an overview of affective structure as well as navigation over the text. The work of Mohammad [Moh12] focuses on analysis of emotions in mail and books that is based...
on lexical matching of terms associated with eight emotions as well as positive/negative categories. The author uses multiple basic representations such as line plots, bar charts, and word clouds to analyze the distribution and temporal trends of emotion-bearing word usage in individual documents and corpora. Finally, Weiler et al. [WGS15] apply their text stream analysis and visualization system called Stor-e-Motion for a combined text of the whole ‘Harry Potter’ series. The resulting visualization represents sentiment as a river-like stream graph with an overlay of salient topic terms list for each time interval (in this particular case, position in text is treated as timestamp).

We did not expect to find a lot of work focusing on sentiment visualization in Editorial Articles/Papers due to the style standard in this genre. The few existing techniques (2%) detect polarity or stance in text that surrounds citations and use this information for analysis and visualization of citation networks. Schäfer and Spurk [SS10] classify polarity and reuse of citations in scientific articles and use the classification results for colour coding of the citation graph edges. Small [Sma11] discusses sentiment analysis and visualization of citations in scientific literature with Maps of Science. His approach results in classification of multiple categories beyond the standard polarity-related ones. For instance, uncertainty and differentiation/contrast can be considered as categories of stance. The results of analysis are used to calculate the layout of a node-link diagram which resembles a map. The recent work of Wang et al. [WLQ*16] involves polarity classification and visualization of citations to represent a citation graph for paper review purposes.

The last category of data domains is Editorial Media such as news or pre-moderated websites (e.g. Wikipedia), and this category is supported by 14% of techniques. Some of the early examples here include SATISFI, ‘Sentiment and Time Series: Financial Analysis System’ by Taskaya and Ahmad [TA03], which uses lexical matching with specific markers to analyze the polarity of financial news documents. The polarity values are aggregated and treated as time series which are visualized with simple line plots. Fukuhara et al. [FNN07] use line charts to visualize topic data associated with eight affect categories in news and blogs, and vice versa, affect data associated with a specific topic. Gamon et al. [GBB*08] describe BLEWS, a system dedicated to the analysis of relations between news articles and political blog posts that refer to such articles. The number of detected subjective posts is encoded visually as the amount of glow around the corresponding bars representing the number of liberal and conservative blog posts. Zhang et al. [ZKKT09] introduce Sentiment Map, a lightweight visualization based on a geographical map. The tool detects eight emotion categories in news articles over time and visualizes the resulting time series with line plots for corresponding geographical regions based on the zoom level. In contrast to some of these techniques with simple and standard representations, TextWheel by Cui et al. [CQZ*12] introduces a combination of several complex and novel representations for monitoring of sentiment associated with specific entities in the news streams. Figure 5(b) displays some of the representations used in the system: a radial node-link diagram representing relations between entities (a keyword wheel), a U-shaped transportation belt that acts as a substrate for moving document glyphs, and a significance trend chart that is represented by a line plot.

We have also found several techniques (around 4%) that are not designed for any particular data domain. Duan et al. [DQP*12] describe VISA, a system for temporal aspect-based visual sentiment analysis which extends the more general system TIARA with sentiment analysis capabilities. VISA uses a river-like stream graph with embedded tag clouds as well as several auxiliary views (pie charts, bar charts, text views) with multiple interactions to support a number of user tasks. The authors demonstrate VISA with a use case involving hotel reviews data, but their approach is not specific to this data domain. Semantize by Wecker et al. [WLM*14] is a lightweight web-based visualization technique that uses font style and background colour to encode the word-level polarity, sentence-level subjectivity, and paragraph-level polarity directly in the HTML document. Neviarouskaya et al. [NAPI14] visualize the results of document analysis with their AM model, including the judgement and appreciation aspects which go beyond typical polarity and emotion categories. Gold et al. [GREA15] list sentiment polarity as one of the possible annotation types for their Lexical Episode Plots technique. In this case, particular words as well as bar segments representing text document regions can be highlighted to facilitate the understanding of affective structure of the text. Typographic Set Graph by Brath and Banissi [BB16] provides a map representation of a word-emotion association lexicon. This technique manages to simultaneously represent membership of individual words in ten sets (eight emotion and two polarity categories) by using font style and colour.

5.1.2. Data source

The type of text data source has implications for the design of sentiment visualization techniques in most cases. However sometimes the categorization can become somewhat fuzzy, since a single large document (e.g. a novel or a transcript) can be treated as a collection or sequence of its sections; and vice versa, multiple separate texts can be concatenated for analysis and visualization.

First of all, we have identified techniques that support data from an individual Document (15% of the total set). In most of such cases, visualization of a single document is used either to support details on demand, or to represent a subset of data from another type of data source. For instance, Affect Inspector by Subasic and Huettner [SH01] can visualize affect profiles for a single document or several documents at a time. uVSAT by Kucher et al. [KSBK*16] provides a separate representation of individual documents based on selected subsets of aggregated emotion/stance value series initially presented to the user (see Figure 6(b)). Ruppert et al. [RBB*16] provide visual summaries of subjectivity and polarity analysis for selected individual documents in their system PolicyLine. Techniques by Gobron et al. [GAP*10] and Krcaadinac et al. [KJD*16] support either a real-time conversation text stream, or its transcript as a single document.

In some cases, though, the focus of the visualization technique is set on a single document. Besides Affect Colour Bar [LSL03], Semantize [WLM*14], and Lexical Episode Plots [GREA15] discussed above, Li and Ren [LR09] visualize 2D/3D heatmap visualizations of emotion matrices for individual text documents. Attitude Radial Plot by Almutairi [Alm13] displays a document as a 2D/3D map using the dimensions of affect, judgement, and appreciation.
as well as the position of text fragments in the document. The technique by Bembenik and Andruszkiwicz [BA16] takes an unstructured text document as input, extracts proposals and arguments alongside their polarity, and visualizes this data using a node-link diagram.

The absolute majority of techniques in our data set (86%) support text data from a collection of documents, a corpus, or corpora. In contrast to the visualization of individual documents, techniques that support such data sources typically have to address challenges related to larger data set sizes, varying text lengths, relationships between documents, and additional data properties (see Section 5.1.3). A typical example here would be a technique oriented at a collection of customer reviews [XLLS11] or previously fetched tweets [WLY*14]. The corresponding category Corpora includes techniques spanning the complete time range of our survey, from Affect Inspector [SH01], Opinion Observer [LHC05], and Pulse [GACOR05] to OpinionSeer [WWL*10] and Vox Civitas [DNKS10] to ConVis [HC14], SocialHelix [CLL*15], and SemEval-2016 dataset visualization [MKS*16].

During the last five years, research in streaming data visualization has also produced a number of sentiment visualization techniques which support Streams as data sources (19% of our collected set). Most of these techniques consume data from microblogs such as Twitter or Weibo, for instance, TwitInfo by Marcus et al. [MBB*11] or MoodLens by Zhao et al. [ZDWX12]. In some cases, the major focus of a technique is set on event detection, and sentiment analysis/visualization plays an auxiliary role. For example, Krstajić et al. [KRHW12] describe a visual analysis system for event detection in Twitter data which uses the aspect-based sentiment analysis method introduced by another technique [RHD*12] to calculate polarity scores for individual documents among other features.

Other streaming data visualization techniques treat sentiment as the first class data. Liu et al. [LZX*13] use a river representation for sentiment polarity values extracted from a Twitter stream with several co-trained classifiers. Kranjc et al. [KSP*15] describe the usage of a general-purpose cloud-based platform ClowdFlows for sentiment analysis and visualization. It detects the polarity of streaming Twitter data by using an active learning classifier and visualizes it with standard line plots, stream graphs, and word clouds. Matisse by Steed et al. [SBBI*15] is a VA system for text stream data that focuses specifically on temporal sentiment analysis of Twitter data. Polarity of individual tweets is visualized with stream graphs and geospatial heat maps, and a fine-grained analysis of emotions in the text data is available with a scatter plot representation. PaloPro by Tsirakis et al. [TPTV16] is a brand monitoring platform which conducts opinion mining of data streams from multiple social media and news sources. Its dashboard visualization includes line plots and bar charts representing polarity for specific topics or named entities.

5.1.3. Data properties

Besides the data source type, we have also analyzed special properties of the data used by sentiment visualization techniques. The results of our analyses discussed in Section 6 confirm that these properties are correlated with the later stages of the sentiment visualization pipeline.

Some of the recent techniques (22%) make use of Geospatial information, starting with Sentiment Map by Zhang et al. [ZKT09]. This tool identifies affective content (namely, eight emotion categories) in news articles. The resulting temporal emotion values are visualized with line plots for corresponding geographical regions based on the zoom level. Scharl et al. [SHHW*12]
present weblyzard, a platform for monitoring and visual analysis of social media, news, and other text documents from the Web. Among other analyses, it supports polarity detection for specific topics and uses these polarity values with line charts, a map and a tag cloud. One of the techniques described by Zhang et al. [ZLW13] provides a geospatial visualization of sentiment in social media. The technique combines a regular geographical map with the representations of kernel density estimation (KDE) analysis for sentiment polarity in specific cities and edges between cities representing the retweeting network. Caragea et al. [CSS*14] combine geospatial information present in tweets with polarity analysis results to study social media patterns emerging during disaster scenarios.

The majority of techniques in our survey (62%) support temporal data, denominated by the category Time Series. Most of such techniques support overview of sentiment value series over time with representations such as standard line plot [MDR06, DBG12, HHHM*16] or river/stream graph [DQP*12, LZX*13, WLY*14, ZGWZ14]. Other popular approaches to visualize temporal sentiment data involve glyphs aligned along a timeline [WRM*09, WWS12, LKT*14] and animation-based methods [GOC12, ZLW13, EAGA*16].

Finally, we have identified the techniques that use Networks present in input or processed data (17% of the total set). The examples of such networks include citation graphs [SS10, Sma11, WLO*16] and graphs of relationships between terms [BHHS12], comments [CCYT15], users [WXL*13], communities [CLL*15], and named entities [SHHJ*16]. Hoque and Carennini [HC14] discuss ConVis, a visualization system for discourse analysis of social media discussions that uses a network of users, topics/concepts, and opinions. The resulting visualization displayed in Figure 5(c) combines a radial node-link diagram with an indented sequence of stacked bars that represents the conversation tree. The subsequent work of the same authors on MultiConVis [HC16] extends the analysis to multiple discussions and temporal data. A similar network of users, entities/concepts, and opinions is also visualized with ORCAESTRA by Prasoso et al. [PDK15]. The authors propose a planetary metaphor for their main node-link representation which represents nodes as stars, planets, and asteroids, and uses a heliocentric radial layout. Asteroid nodes represent individual comments and use colour to encode the corresponding sentiment polarity.

5.2. Tasks

The process of designing a sentiment visualization technique includes the analysis of intended data, intended audience, and intended tasks. We focus on the latter aspect in this subsection and analyze the tasks related to both the computational and visual/exploratory methods.

5.2.1. Analytic tasks

High-level analytic tasks supported by sentiment visualization techniques are based on the respective sentiment analysis models introduced in Section 2.

Polarity Analysis/Subjectivity Detection is the most common analytic task in our survey associated with 81% of techniques. The techniques which support only this analytic task provide a summary about the overall polarity of text data, for example, SATISFI by Taskaya and Ahmad [TA03] or Vox Civitas by Diakopoulos et al. [DNKS10]. Annett and Kondrak [AK08] extend a blog visualization tool eNulog with sentiment analysis by classifying movie blog posts into positive, negative, and neutral/uncertain, and using these labels for colour coding of the blog map nodes. Pupi et al. [PDPA14] highlight the polarity of individual sentiment-bearing terms in their system Ent-it-up. Agave by Brooks et al. [BRT*14] is one of the few sentiment visualization systems that focus on collaborative visual analysis of social media data. The system visualizes the aggregated polarity of temporal text data using representations such as line plots and stream graphs. The final example in this category is BLEWS by Gamon et al. [GBB*08], a system dedicated to the analysis of relations between news articles and political blog posts that refer to such articles. The characteristic detail of BLEWS is that it focuses only on detecting subjectivity in the blog posts without more detailed polarity analysis. The number of detected subjective posts is visually encoded as the amount of glow around the corresponding bars representing the number of liberal and conservative blog posts.

Another category with high support in our data set (64%) is Opinion Mining/Aspect-based Sentiment Analysis. We have used this item for techniques supporting sentiment analysis at the level of particular aspects/features, topics, named entities, or clusters detected in text. This task has often been supported historically with customer reviews data. For instance, Review Spotlight by Yatani et al. [YNTT11] is a simple visualization tool for summarizing customer reviews with a tag cloud of salient adjective/noun word pairs. The sentiment polarity for each word is calculated with lexical matching over its counterparts and used for colour coding. OpinionBlocks by Alper et al. [AYHK11] provides an aspect-based sentiment overview for customer reviews. The visualization combines multiple coordinated bar charts and text tags that can be explored interactively to investigate the polarity and salient keywords associated with specific product features. SentiVis by Di Caro and Grella [DCG13] visualizes results of aspect-based sentiment analysis of customer reviews. After selecting a specific aspect, the users are provided with a visual representation of polarity scores for review objects (in this case, restaurants) that combines a scatter plot and a line plot. Görg et al. [GLK*13] discuss the text analysis and visualization features added to Jigsaw, a general-purpose VA system. The support for sentiment analysis includes document-level polarity detection, which can also be used for analysis of specific aspects when combined with topic analysis, as demonstrated with car reviews data. The visual representations involving the polarity data include word clouds and pixel-based Document Grid Views that can encode document-level polarity.

Another prominent application of techniques in this category is the analysis of social media texts. Wensel and Sood [WS08] describe several basic visualizations of sentiment with regard to the specific topics discovered in personal blog posts with their system VIBES. The authors estimate valence of the texts (hence their claims about analyzing the emotional content of the text data—nevertheless, valence detection on its own is similar, if not
et al. [ZGWZ14] support visual analysis of social media conversations.

Examples of more recent sentiment visualization techniques from 2011–2014. (a) TwitInfo [MBB*11] uses colour-coded markers to represent polarity of individual tweets detected as part of visual event analysis. Adam Marcus et al. TwitInfo: Aggregating and Visualizing Microblogs for Event Exploration. CHI 2011, p. 230, https://doi.acm.org/10.1145/1978942.1978975. ©2011 ACM, Inc. Reprinted by permission. (b) TextWheel [CQZ*12] supports monitoring of sentiment associated with specific entities in the news streams. Weiwei Cui et al. Watch the Story Unfold with TextWheel: Visualization of Large-Scale News Streams. ACM Trans. Intell. Syst. Technol. 3, 2, Article 20, p. 2, https://doi.acm.org/10.1145/2089094.2089096. ©2012 ACM, Inc. Reprinted by permission. (c) ConVis [HC14] visualizes a network of users, topics/concepts, and opinions detected in social media discussions. Image courtesy of Enamul Hoque.

Emotion/Affect Analysis is related to analysis of affective content in text beyond the positive/negative categories, usually involving a categorical or dimensional emotion model. Approximately 24% of techniques in our survey support this task. Affect Inspector by Subasic and Huettner [SH01] uses star plots to visualize affect profiles of text documents. The authors combine lexical matching with fuzzy tagging to label documents with eight affect categories (including basic emotions), and support visual exploration and annotation with their tool. Gregory et al. [GCW*06] conduct analyses of affect bearing words in customer reviews using lexical methods, an annotation tool and a general-purpose visualization system INSPIRE. They propose a novel metaphor based on a rose plot to visualize statistics for eight affect categories. The users can investigate the rose plots for the corpus in general as well as particular clusters of reviews. Kang and Ren [KR11] conduct a joint emotion/topic analysis in blog posts. They analyse the output of their method by examining node-link diagrams generated for topic networks that use colour coding based on one of the eight emotions dominating for the respective topic. Zhao et al. [ZGWZ14] support visual analysis of emotions detected in tweets over time for an individual user with their system PEARL. They use two emotion models, the categorical Plutchik’s model with eight basic emotions and the dimensional VAD model. PEARL uses river-like stream graphs as the main visual representation for temporal emotion data alongside multiple auxiliary representations (line and area charts, glyphs, scatter plots, and tag clouds) to support overview and detailed exploratory analysis. Wang et al. [WSK*15] discuss a technique for emotion analysis of Twitter data called SentiCompass. As displayed in Figure 6(a), the visual representation uses star plots that naturally correspond to the polar valence-arousal space and organizes them in a nested fashion resembling a spiral to support the simultaneous visual analysis of multiple temporal intervals. Additionally, the same data is represented by linearly ordered line plots as an auxiliary representation.

Finally, we have introduced the category Stance Analysis to represent analyses that encompass not only sentiment/affect, but also other categories of subjectivity/evaluation expressed in text, for instance, certainty or judgement. Currently, very few techniques support this task (6%), including the works of Small [Sma11], Almutairi [Alm13], Neviarouskaya et al. [NAPI14], and Bembenik and Andruszkiewicz [BA16] discussed above. Torkildson et al. [TSA14] complement Ekman’s six emotions with two additional categories representing support and accusation for their visualization of social media posts on the Gulf Oil Spill crisis. uVSAT by Kucher et al. [KSBK*16] is a VA system dedicated to sentiment and stance analysis which detects the expressions of Ekman’s six emotions as well as certainty and uncertainty in social media data over time. These time series are visualized with line plots, and the contents of corresponding documents are represented with highlighted text, scatter plot thumbnails, and bubble charts, as displayed in Figure 6(b). Mohammad et al. [MKS*16] introduce a dashboard visualization of the SemEval-2016 stance-annotated data set which provides an overview of the data with regard to class distributions by using several bar charts, a tree map, and confusion matrices. In their work, stance is treated as agreement/disagreement with a
certain topic besides the expressed sentiment. Finally, El-Assady et al. [EAGA*16] provide an animation-based visualization of conversation transcripts, e.g., political debates transcripts, with their system ConToVi. The system supports several stance categories related to argumentation as well as sentiment, certainty, and politeness. It allows the users to monitor the stance of individual speakers with regard to specific topics.

### 5.2.2. Visualization tasks

Besides the higher-level analytic tasks related to sentiment analysis model, we have included a number of more concrete representation and interaction tasks that are directly supported by sentiment visualization techniques.

For example, the category **Region of Interest** includes techniques which support automatic detection and highlighting of interesting or anomalous items [CQZ*12, ZCW*14, ZGZW14, CSL*16, LSB*16], regions [KSBK*16], or peaks [MBB*11, SHHW*12] (overall, 8% of the total set of techniques). Time Density Plots by Rohrdantz et al. [RHD*12] allow the users to focus on sequences of subjective text documents related to the specific features (aspects) in customer reviews and RSS news streams. The technique supports visual analysis of interesting data regions and automatic extraction of such patterns. The actual visual representation combines a series of bar glyphs with an area chart for the corresponding time region.

TargetVue by Cao et al. [CSL*16] uses sentiment polarity values to analyze and visualize anomalous user behaviour analysis in social media and email communications, employing glyphs to represent detected anomalous users.

A lot of techniques in our survey (58%) are related to **Clustering/Classification/Categorization**. Here, we have identified the techniques that involve additional (semi-)automatic tagging or grouping of data elements besides the actual sentiment classification, represented or facilitated by the visualization. The work by Oelke et al. [OHR*09] on opinion analysis for customer reviews uses several techniques for visual analysis of aspects/features, including a matrix-based visual summary report and a circular correlation map (a sort of combination of a node-link diagram and a parallel coordinates plot). The authors also introduce a novel technique for the visual analysis of opinion clusters that combines a Voronoi diagram with thumbnails containing cluster details as tables. Brew et al. [BGAC11] support the temporal sentiment polarity analysis for groups of Twitter users with their SentireCrowds system. They cluster users into groups using tweet contents and calculate the aggregated polarity values for each cluster/time step. SentireCrowds provides an overview of the overall sentiment over time with an area chart as well as multiple treemaps for individual time steps which represent the polarity and salient keywords for user clusters. Kim and Lee [KL14] discuss a dimensionality reduction technique called Semi-Supervised Laplacian Eigenmaps which they apply to the customer reviews data. The method involves extraction of features (terms) related to positive/negative categories and dimensionality reduction of the feature space based on graph and matrix computations. The resulting 2D embedding of reviews, which highlights the clustering in the data, is visualized with a colour-coded scatterplot.

ToPIN by Sung et al. [SHS*16] identifies topics in student comments by using a clustering algorithm and then represents them as nodes. The average polarity of comments belonging to the same cluster is encoded by node brightness.

**Comparison** of several entities is facilitated by most visualization techniques in our survey (91%). For instance, Xu et al. [XLLS11] propose a method for opinion analysis of product reviews that directly takes the comparisons of particular features into account. The resulting visualization uses a node-link diagram to represent a probabilistic graphical model as a bipartite graph, where the polarity and direction of comparison are displayed for each feature. Kuksenok et al. [KBR*12] use a timeline visualization to represent occurrences of several affect categories in their annotated data set and to identify relationships between such categories. One of the techniques used in the SocialBrands system by Liu et al. [LXG*16] is BrandWheel, which visualizes the scores of various brand aspects estimated by the analysis of social media posts and employees’ reviews. The system is capable of simultaneously displaying two BrandWheels for comparison purposes and even visualizing their differences as a derived representation.

A similar comparison mode was previously used in EmotionWatch by Kempfert et al. [KSPM14] to compare emotional reactions to a pair of topics or named entities on Twitter. Lariat by Chen et al. [CBK*17] is designed for visual comparison of Twitter queries using multiple attributes, including polarity analysis results.

Another task supported by the absolute majority of techniques (95%) is **Overview**—“The big picture” achieved by displaying (1) a significant portion of the data set or (2) aggregated representations. The former (1) is the case for techniques such as City Sentiment by Wu et al. [WFL*13], which displays a grid of bubble glyphs representing Chinese cities. Each glyph encodes the number of Weibo posts from the respective city as well as the overall detected polarity. BrandSediments introduced in the SocialBrands system by Liu et al. [LXG*16] provide an overview of brand traits/aspects such as ‘Sincerity’ and ‘Excitement’ by laying out clouds of bubbles (which represent individual brands) along the corresponding axes. (2) Aggregated representations are used to provide overview by techniques such as AMAZING by Miao et al. [MLD09] or Westeros Sentinel by Scharl et al. [SHJJ*16], which include pie charts to provide a summary of polarity values distribution. SentiWheel by Gali et al. [GOCD12] provides daily summaries of the sentiment on several Canadian banks in social media posts with a variation of a sunburst diagram which is also capable to display relations based on shared keywords.

**Monitoring**, which could be described as dynamic observation and alerts on changes in the data. Some of the techniques use a continuously updated timeline metaphor to represent the changes in data [KRHW12, KSP*15, WGS15], others involve dedicated representations heavily relying on animation [GAP*10, GOC12, EAGA*16, KIDP16], or combine multiple representations [MBB*11, CQZ*12, SHHJ*16]. Calderon et al. [CAHF14] discuss how combining several visual representations can facilitate the monitoring task for streaming sentiment data.

We have used the category **Navigation/Exploration** for the 66% of techniques which support interactive exploration of the data. Most of such techniques support details on demand,
typically by providing the user with a detailed (text) view for the
selected timeline [DQP*12, BRT*14] or map [LWW*16, LSB*16] region. Another option is to display specific analysis details for a
selected/hovered item [HRJ*13, TSA14, CCYT15].

Finally, we have added the category Uncertainty Tackling
for sentiment visualization techniques that make use of uncertainty present in data and/or computations (currently supported by
roughly 7% of techniques). In several works, uncertain [AK08],
undefined [SS10], or controversial [DNKS10] sentiment is used be-
side the neutral category. Gamon et al. [GBB*08] use the certainty
level of their subjectivity classifier to filter out blog posts which
should not be represented in the visualization with the threshold of
50% certainty. OpinionSeer by Wu et al. [WWL*10] is a complex
system for the visual analysis of customer feedback that extracts po-
larity values associated with certain features (aspects). The further
computational steps include uncertainty modelling and aggregation
of opinions based on a model called ‘subjective logic’. The authors
provide a visualization that combines multiple radial views as well
as a scatter plot laid out inside of a triangle to represent temporal,
geospatial, aspect, sentiment, uncertainty, and customer profile data
at multiple levels of details. Rohrdantz et al. [RHD*12] estimate
the uncertainty of their sentiment analysis method and encode such
values directly in their Time Density Plots alongside the polarity and
data entry counts. The work of Makki et al. [MBM14] focuses on
sentiment lexicon expansion/annotation. Their aspect-based senti-
mement analysis model predicts the polarity values of individual terms.
The users can then validate and edit the predicted sentiment values
by using several visual representations: a tree cloud and a scatterplot
with embedded word clouds. Uncertainty in this case is related to the
ambiguous polarity predictions, which are highlighted in yellow
in the tree cloud representation. The VA framework for event cueing
by Lu et al. [LSB*16] involves the polarity analysis of RSS news
messages based on ML methods. The authors address the task of
uncertainty visualization with stream graphs: stacked layers encode
the volumes for several sentiment classification certainty levels over
time. The previous work from the same authors [LHW*15] also uses
blurred map glyphs to represent uncertainty of Twitter data polarity
analysis for disaster scenarios.

5.3. Visualization aspects

The final two groups of our categorization are related to specific
aspects of representing sentiment data visually.

5.3.1. Visual variable

One of the interesting research questions that we aimed to an-
swer with this survey was related to how sentiment values are
usually represented by various techniques. Based on the work by
Bertin [Ber83], we have introduced several categories for the visual
variables used to encode sentiment.

The most popular choice by far is Colour, used by 89% of
techniques. Most techniques that focus on polarity values use green
for positive and red for negative polarity [WS08, TPTV16], al-
though several techniques reverse this colour map to use green/blue
as a cold hue corresponding to negative polarity, and orange/red for
positive polarity [WFL*12, WFL*13, CAHF14, CFKA14, HC14,
HC16, YWC*16]. The techniques related to emotions or affect cat-
egories other than positive/neutral/negative use either an ordinal
set of colours representing categories [KR11 ZGWZ14, EAGA*16]
or interpolate the colours based on the corresponding dimensional
model [Alm13, WSK*15].

Position/Orientation holds the second rank with the sup-
port of 49% techniques. The most obvious example here would be a
line plot whose vertical axis corresponds to the aggregated po-
larity value [TA03, CGG*15]. Vista by Hoeber et al. [HHM*16]
uses line plots to represent counts of positive, neutral, and nega-
tive tweets over time. In other cases, horizontal position is used to
encode the polarity [MBM14, CCT15]. Eventscapes by Adams
et al. [APV11] is a system for visual analysis of events, topics,
and emotions (moods) in RSS news feeds. The main visual representation of Eventscapes uses a timeline metaphor with document thumbnails laid out linearly according to their timestamps and valence values (vertical and horizontal axis, respectively). Also, several techniques which use a gauge metaphor rely on orientation rather than position [WS08, ZQLT16].

- **Size/Area** is used to represent sentiment by 44% of techniques in our survey. The basic examples include the techniques which use bar charts [AYHK11, MKS*16], pie charts [MBB*11, DBG12], area/stream charts [LZX*13, KSP*15], or filled star plots [KSMP14, MMMS15]. More unusual cases in this category include BLEWS by Gamon et al. [GBB*08], which encodes the subjectivity level with amount of glow, and Semantize by Wecker et al. [WLM*14], which adjusts the font size of intensifying and diminutive words detected in text.

The remaining choices of visual variable are much less popular.

- **Shape** is used by 8% of techniques. In several cases, the corresponding techniques use a glyph [LSL03] or glyph-like [GAP*10] representation. The work by Lee et al. [LFO*10] on analysis of financial blog posts uses simple glyphs to represent polarity values in text lists alongside line plots and bar charts. Tour-pedia by Cresci et al. [CDG*14] analyses customer reviews of touristic locations and visualizes them by embedding smiley glyphs (based on the average review polarity) in a map. Kim and Lee [KL14] use several marker shapes as well as colour coding to differentiate between labeled/unlabelled positive and negative reviews in their scatterplot representation. Other techniques use the contour of a heat map or similar representation to convey emotion values [LR09, Alm13]. Kempter et al. [KSMP14] conduct an emotion analysis of social media texts with 20 emotion categories. Their visualization system EmotionWatch uses a star plot with filled area to represent the emotions detected for the selected time interval. Muneezer et al. [MMMS15] focus on the temporal emotion analysis and visualization for an individual Twitter user with their tool EmoTwitter. One of the visual representation used for the resulting eight basic emotions is a star plot with filled area.

Finally, we have discovered that only four techniques in our complete data set (3%) use **Texture/Pattern** to represent sentiment. Gali et al. [GOCD12] introduce three techniques involving polarity detection in timestamped social media posts related to five Canadian banks. One of these techniques, Emotional Tapestry, generates monthly summaries as woven patterns different for various sentiments, which can be combined for several banks at a time. Zhang et al. [ZLW13] introduce a visualization technique for temporal visualization of social media sentiment that is based on the Electron Cloud Model. The authors calculate the polarity of posts over time for individual users and then visualize this data using a special layout algorithm, which results in a texture-like rendering of trajectory lines. Kuang et al. [KTL14] describe ImgWordle, a visualization tool that is designed for social media monitoring. One of its visual representations is a choropleth map that provides an overview of frequency and sentiment polarity of posts for each region. The authors use textures to represent sentiment, since the colour channel is used to convey topic data. Finally, Krcadinac et al. [KJDP16] propose an artistic visualization technique for chat conversations called Synemania. They analyze the emotions present in chat messages using the Ekman’s six emotions and then use this data in an animated particle simulation. The visualization can therefore be used for monitoring of emotions during a conversation by observing the overall texture and colours.

### 5.3.2. Visual representation

The last group in our categorization includes visual representations (or metaphors) that make use of sentiment. The statistics in this group are not so heterogeneous as for visual variable, for instance: the categories described below are supported by 17% to 43%. The majority of techniques in our survey (70%) have more than one category assigned—in many cases, novel or complex representations combine several traits, or multiple coordinated views are used to represent data that includes sentiment information.

The first category in this group is **Line Plot/River**, supported by 43% of techniques. Basic line plots/Charts have been used primarily for temporal data in sentiment visualizations since earlier works such as SATISFI [TA03] and MoodViews [MDR06] up to the recent techniques such as uVSAT [KSBK*16] and Westeros Sentinel [SHH*16]. We also include the techniques that use area charts in this group since they usually convey the same data as line plots, for example, the work by Fukuharu et al. [FNN07] or LingoScope by Diakopoulos et al. [DZES14]. Then there are river-like representations [HNN00, BW08] that range from a simpler stacked area chart used, for instance, in VisTravel by Li et al. [LW*16], to stream graphs used in VISA by Duan et al. [DQP*12], PEARL by Zhao et al. [ZGWZ14], and OpinionFlow by Wu et al. [WLY*14]. The flow metaphor is also used by a recent work by Wang et al. [WLC*16]. Their system called IdeaFlow combines the ideas of storyline visualizations and topic modelling to facilitate the lead analysis of ideas discovered in social media; the results of opinion mining are integrated in the visual representation displayed in Figure 6(c) to convey polarity towards certain ideas. Finally, we also include several techniques in this category which do not use a standard line plot per se, but rather representations which consist of line segments or contours drawn in a certain coordinate system. The corresponding examples include a star plot used in Affect Inspector by Subasic and Huettner [SH01], a line plot variation used in SentiVis by Di Caro and Grella [DCG13], and a parallel coordinates plot used by Zhao et al. [ZLT16].

The second category in the visual representations group is **Pixel/Area/Matrix** (used by 39% of techniques). Here, we have tried to collect the techniques which use space-filling approaches and other representations which rely on the size/area variable. One of the basic representations here is a pie chart; for instance, AMAZING by Miao et al. [MLD09], TwitInfo by Marcus et al. [MBB*11], and 5W Summarization by Das et al. [DBG12] use pie charts to provide an overall summary about the polarity distribution. Kumamoto et al. [KWS14] describe an emotion detection tool for personal Twitter data that uses six emotion categories organized in polar pairs. The tool presents a simple visualization consisting of line charts and pie charts for the individual user’s data over time. Many other techniques use various forms of bar charts, including regular histograms used in FAVe by Guzman et al. [GBB14], VisOHC by Kwon et al. [KKL*16] and GeoSentiment by Pino et al. [PKS16], and stacked bar charts used by Dehiya and Mueller [DM16] and
Mohammad et al. [MKS*16]. Wanner et al. [WRK11] introduce a visualization system for internet forum data called ForAVis. The main visual representation used by the authors comprises pixel-based stacked bars that are organized and coloured according to the selected set of features, including sentiment polarity. Another example of this category is a tree map, which is used in Pulse by Gamon et al. [GACOR05] and SentireCrowds by Brew et al. [BGAC11]. Wu et al. [WFL*12] use a cartogram for their system City Flow to represent sentiment of Weibo posts originating from specific cities. The sizes of nodes represented by squares are calculated similarly to a regular tree map, but the layout takes the geographical positions of cities into account.

Various forms of Node-Link representations are used by sentiment visualization techniques (17%) on their own, e.g. in works of Kang and Ren [KR11], Small [Sma11], or Maki et al. [MBM14], or in combination with other representations such as a map (e.g. in Whisper by Cao et al. [CLS*12] or the work by Zhang et al. [ZLW13]) or a (stacked) bar chart (for instance, in the work by Xu et al. [XLLS11] or the works by Hoque and Carenini [HC14, HC16]). Several techniques use arc diagram variations, e.g. the work by Chen et al. [CISSW06] or News Flow by Brayœuvreanu et al. [BHHS12]. Arc diagrams are also used in the recent work by Fu et al. [FZCQ17], whose system iForum uses multiple representations for visual analysis of MOOC forum data. Individual threads are represented by a combination of a river and an arc diagram called Thread River, and this representation can be user-configured to display sentiment analysis results.

We have used the next category for representations that use multiple visual items (such as dots, bubbles, glyphs, or words/tags) to give rise to associations with Clouds/Galaxies (supported by 27% of the techniques set). Some of the techniques in this category use various forms of word/tag clouds, for instance, Review Spotlight by Yatani et al. [YNTT11] and webLyzard by Scharl et al. [SHHW*12]. FishEye Word Cloud by Wang et al. [WDN13] is a technique for temporal sentiment visualization that detects the polarity of individual key terms in a set of Twitter posts. The visual representation is based on a word cloud whose layout takes the temporal order into account. The technique makes heavy use of focus+context for interactive exploration. Other techniques rely on clouds of dots or similar markers, e.g. RadViz-based Attribute Astrolabe used in SentiView by Wang et al. [WXL*13] and the work by Kim and Lee [KL14]. Opinion Zoom by Marrrese-Taylor et al. [MTVBM13] provides a lightweight visualization of customer reviews using an aspect-based sentiment analysis model. The system uses basic visual representations such as bar charts and bubble charts. Lu et al. [LKT*14] describe a VA framework for classification and prediction that uses sentiment analysis to predict movie gross based on social media texts and movie reviews. Polarity values are used in several visualizations used in the framework, namely, word clouds and temporal bubble charts. Finally, the galaxy metaphor is also used in a more literate sense in CosMovis by Ha et al. [HKH*14], a system for emotion analysis in movie reviews which displays a constellation map of movies positioned according to the specific affect-bearing words detected in reviews. The visualization also contains the centroids of corresponding word clusters and artistic representations of emotions as constellations.

The category Maps includes techniques (27%) which use either (1) an actual geographical map or (2) an abstract map which somehow allows the user to identify interesting regions or peaks in the overall landscape. In the former case (1), the maps are often augmented with markers (e.g. in TwitInfo by Marcus et al. [MBB*11]) or overlays (e.g. in the work by Zhang et al. [ZLW13]). The visualization of public opinions on educational institutions in eduMRS–II by Qiu et al. [QRQ15] provides a map with overlaid circles/bubbles which represent the aggregated positive or negative polarity. A similar approach is used by Dai and Prout [DP16] to represent the aggregated positive sentiment on the Super Bowl teams extracted from Twitter. Tweetviz by Sjitsma et al. [SQC16] uses a map view to represent business locations reviewed in tweets and encodes the review polarity with the map marker colour. Some techniques use the choropleth approach, for instance, MoodLens by Zhao et al. [ZDWX12] and ImgWordle by Kuang et al. [KTLS14]. The work by Yu et al. [YWC*16] uses a choropleth map among other representations to display sentiment analysis results for Weibo microblogs in China. Zhao et al. [ZQLT16] introduce Social Sentiment Sensor, a visualization system for monitoring topics and emotions in microblog streams. The resulting values are aggregated for a specific topic or geographical region and visualized with line plots, pie charts, bar charts, choropleth maps, and gauge-based glyphs. (2) Some of the abstract map examples in this category include heatmap-like representations. Sentimap by Hennig et al. [HBM*14] creates a heatmap for temporal sentiment data based on specific Twitter search terms (hashtags). A later technique, Cluster Heat Map [HBB*15], uses an algorithm based on dimensionality reduction to change the vertical layout of the heatmap. This facilitates the discovery of patterns and clusters in the data. Finally, some techniques use a combination of ideas, for instance, Hao et al. [HRJ*13] as well as Steed et al. [SBBI*15] create a pixel-based heatmap on the top of a geographical map. In contrast, GeoSentiment by Pino et al. [PKS16] uses a regular heatmap in the overlay, which is calculated for aspect-based sentiment analysis results for social media and news data.

While most of the techniques in our survey visualize sentiment extracted from text data, some of them (25%) use Text as one of visual representations. Techniques such as Ink Blots by Abbasi and Chen [AC07], Jigsaw by Görg et al. [GLK*13], Semantize by Wecker et al. [WLM*14], and uSAT by Kucher et al. [KSBK*16] visualize the content of a text document with words, phrases, or regions highlighted based on the sentiment analysis results. In addition to such an approach, LDA-based sentiment visualization by Chen et al. [CCT15] presents an overview of sentiment-bearing word pairs for a selected topic in hotel reviews by positioning them in a grid layout. The horizontal position encodes the polarity in this case. Typographic Set Graph by Brath and Banissi [BB16] demonstrates how multiple font attributes such as font family, style, and weight can be used simultaneously with colour coding to represent membership of words in 10 sets corresponding to emotion and polarity categories.

The final category in this group is Glyph/Icon, which is supported by 29% of techniques. Multiple techniques in our survey use glyphs in combination or as part of other visual representations. For instance, Wanner et al. [WRM*09] propose a visualization of RSS feeds comprising series of curved bar glyphs that encode temporal,
topical (whether the RSS news is related to the Democratic or Republican party), and polarity information. A later work by Wanner et al. [WWS12], Topic Tracker, is a system for temporal analysis of Twitter streams that combines topic monitoring and gradual sentiment polarity detection. The authors use basic colour-coded triangle glyphs representing the timestamp and polarity of individual tweets. The glyphs are positioned in a dense fashion, and the final result resembles pixel-based representations. FluxFlow by Zhao et al. [ZCW*14] is a VA system for investigating anomalous patterns of information spreading on social media, namely, retweeting threads on Twitter. The system uses lexical matching of emotion-bearing words as one of the features for estimating the anomaly score and visualizes these values as part of thread glyphs. Several other techniques use gauge-like glyphs, for instance, VIBES by Wensel et al. [WS08] or Social Sentiment Sensor by Cresci et al. [ZQLT16]. Even smiley faces are used in visualizations as glyphs (e.g. in Affect Colour Bar by Liu et al. [LSL03] and Tour-pedia by Cresci et al. [CDG*14]), which in a way is taken to the next level by actual faces of virtual avatars by Gobron et al. [GAP*10].

6. Discussion

After several iterations of adding new techniques and refining the categorization, we have been able to summarize the state of the art in sentiment visualization based on the statistics for our data. In addition, we have investigated the relations between categories in general.

Correlation between categories

We have conducted a correlation analysis for categories assigned to sentiment visualization techniques. Technique entries were treated as observations, and categories were treated as dimensions/variables (see the supplementary material for the complete categorization results). Linear correlation analysis was then used to measure the association between pairs of categories. The resulting matrix in Figure 7 contains Pearson’s r coefficient values which reveal certain patterns and interesting cases of positive (green) and negative (red) correlation between categories. The interpretation of the coefficient values seems to differ in the literature: Cohen [Coh88] defines the range 0.30–0.50 (absolute values) as moderate correlation and 0.51–1.00 as strong correlation; Taylor [Tay90] mentions the corresponding ranges 0.36–0.67 and 0.68–1.00 used in earlier works; and Evans [Eva96] defines the ranges 0.40–0.59 for moderate correlation, 0.60–0.79 for strong correlation, and 0.80–1.00 for very strong correlation. Based on this, we have focused on cases with an absolute value of 0.40 or higher.

The interesting cases with negative correlation mostly include categories from the same groups, implying ‘competition’ or a kind of paradigmatic relation between them. For example, the correlation of -0.49 between document and social media could be explained by a general shift from data sets of well-defined product reviews to the data extracted from social media (including posts associated with some brands)—see the discussion of such temporal trends below. The correlation of -0.51 between the document and corpora categories is interesting, given the discussion in Section 5.1.2. While there are many techniques using both data source types, this value could be explained by techniques focusing exclusively on individual documents or corpora. The document category is also negatively correlated to the temporal data property with a value of -0.45: while there are some techniques which treat position in text as the temporal dimension, the majority of the techniques which support time series apparently use other data sources. Finally, the strongest negative correlation in our survey is -0.63 between the analytic tasks of polarity and emotion/affection analysis, which is explained by a more specialized nature of the latter task.

The cases of positive correlation tend to include categories from different groups in a kind of syntagmatic relation. For instance, the correlation of 0.42 between the literature domain and the document data source type can be explained easily: the corresponding techniques focus on individual novels/poems. The correlation of 0.40 between the document data source and emotion/affection analysis is more surprising in this regard—perhaps, it is affected by a rather low support for both categories. The other cases involving visual tasks and representations are generally easier to interpret. The positive correlation of the region of interest and monitoring with the respective values of 0.48 and 0.64 was expected. The analytic task of opinion mining/aspect-based sentiment analysis involves the support for various kinds of classification/clustering/categorization, hence the positive correlation of 0.42 (it is a little surprising that this value is not higher, though).

Temporal data is often visualized using representations such as pie/donut charts. The cases of temporal data source with such visualization tasks as region of interest and monitoring with the respective values of 0.48 and 0.64 was expected. The analytic task of opinion mining/aspect-based sentiment analysis involves the support for various kinds of classification/clustering/categorization, hence the positive correlation of 0.42 (it is a little surprising that this value is not higher, though).

The strongest positive correlation of 0.74 in our survey exists between network data and node-link representations, which is not surprising at all. We could also expect the correlation between geospatial data and map representations to be higher than 0.58—but we have noted in Section 5.3.2 that the latter category includes also abstract maps. Finally, the correlation of 0.51 between pixel/area/matrix representations and the visual variable of size is explained by such representations as bar charts and pie/donut charts.

Popular approaches

Table 2 presents the statistics for the collected data based on the final categorization. It supports our expectations of the most common aspects of existing sentiment visualization techniques. An average technique is used for visualization of temporal data (stored as corpora) from social media based on polarity analysis/subjectivity detection. The popularity of the more specific aspect-based sentiment analysis/opinion mining is explained by the existing interest for topic analysis and visualization. Based on the statistics given in Table 2, we can also identify a standard set of four visualization tasks relevant to sentiment, which reflect the visual information seeking mantra [Shm96] and visual analytics mantra [KAF*08] to some extent: if possible, cluster the data into groups first, then provide an overview of these results, compare interesting items, and explore them in detail. More specific visualization aspects related to variable and representation are discussed below. Also, we should note that the absolute majority of techniques rely only on 2D representations (even though we have not discussed it explicitly).

Temporal trends

In addition to the overall statistics, it is also interesting to analyze temporal trends with regard to the occurrence...
of individual categories in our collected data set. Figure 8 comprises sparkline-style plots based on the category counts normalized by the total technique counts for each year (for example, 6 out of 28 techniques from 2016 support streaming data). It allows us to detect global trends and compare trends within each group of categories. For instance, it confirms our previous statements about popularity of social media and decreasing role of reviews as data sources. The popular approaches discussed above demonstrate stable support throughout the years, but it is also interesting to trace the usage of underrepresented categories over time, as discussed below.

**Underrepresented categories**

Emotion/affect analysis is supported by a relative minority of techniques, and only a few techniques address the issues of stance analysis. According to the temporal trends (see Figure 8), the interest has varied over the years for the former task and started to emerge only recently for the latter task. These analytical tasks present multiple opportunities for future research with applications in several domains (social media, literature, etc.). As for visualization tasks, uncertainty tackling is currently underrepresented. This also presents future research opportunities, since uncertainty is an inherent aspect of many popular ML models (e.g. SVM or CRF) as well as data sets characterized as ‘Big Data’. The increasing interest for such data will for sure also affect the number of techniques supporting streaming data sources as well as the related visualization task of monitoring. While such data sets are mostly associated with online sources, we have been surprised by the statistics for several other domains. For instance, very few techniques (mostly from the past years) address the literature domain as opposed to the overall growth of interest.

**Figure 7:** The matrix of correlation values for categories (see Table 2) calculated for our techniques data set.
for digital humanities. Also note the low number of techniques that focus on a single document—the interest for such data sources has diminished over the years. We also had to exclude the domain of patents from our categorization since no detected techniques support it. This can be explained by the generally formal and objective style of language used in such texts—on the other hand, several techniques already support the related domain of scientific articles.

**Visual representations** The statistics on the visual variable state that colour is the most common visual channel to convey sentiment/emotion, which was expected by us even before collecting the data. The rather large number of techniques using position/orientation and size/area to encode sentiment can be explained by the usage of line plots and stream graphs for temporal sentiment data, as well as pie charts and bar charts for simple visual summaries which are often used by techniques originating in non-visualization disciplines. This leads us to the issue of categorizing the visual representations used for sentiment into simple and complex, which was initially one of our intentions. The existing work investigating complex representations (for instance, [TKT*00, WBWK00, YN05]) refrains from providing exact definitions of such. Shamim et al. [SBT15] have already raised the issue of evaluating sentiment visualization techniques including the perception aspects—this problem presents interesting opportunities for future research.

**Interactive exploration with a survey browser** We have developed an interactive survey browser similar to our TextVis Browser to accompany this paper and used it extensively ourselves while working on the paper. Figure 9 demonstrates its user interface that is focused on individual technique thumbnails and the interaction panel comprising category filters and a search field. In general, it follows the design decisions used by several existing browsers [Sch11, TA, KK14, KK15, BKW16]. SentimentVis Browser is implemented as a client-side web application using HTML, JavaScript, and D3. After loading the page, the user is presented with a list of visualization technique thumbnails organized in a grid. The entries are ordered by publication year first and then by the prime author’s surname. Clicking on a thumbnail opens a dialog box with details such as a complete bibliographical reference, a URL link to the source publication webpage (if available), a BIBTeX file link, and a complete list of categories assigned to the corresponding technique. The categories are also presented in form of filters in the main interaction panel on the left. Additionally, the panel includes a text search field, a time range slider, and a histogram showing the temporal distribution of techniques before and after filtering. The users can also access category statistics via the ‘About’ dialog (similar to the data in Table 2) and a summary table with an overview of the complete categorization (see below). We encourage the users to submit additional entries to SentimentVis Browser via a form available from the top panel—the process is not entirely automated, though, since we intend to continue careful curation of the survey.

**Survey summary and citation counts** Besides the interactive survey browser, our entire resulting categorization is presented in the supplementary material (see Tables S1–S3). These results take the aforementioned scope of particular categories into account. For instance, most visualization techniques use several visual representations/views to represent a single complex data set. In such cases, we considered only the representations involving sentiment, and categorized the techniques appropriately. The supplementary material also includes citation counts for the publications corresponding to visualization techniques (see Table S4). While this information is external to the actual techniques, it allows us to make assumptions about the impact and adoption of sentiment visualization approaches. When using these citation counts as guidelines, though, it is important to take the aspects related to publication year and discipline into account. For example, the most cited publications from our survey belong to disciplines such as data mining and NLP—these are the works on Affect Inspector [SH01], Pulse [GACOR05] (see Figure 4(a)), Opinion Observer [LHC05], comparative relation maps [XLLS11], and MoodLens [ZDWX12]. The most cited publications from visualization outlets are related to techniques such as TwitInfo [MBB*11] (see Figure 5(a)), Vox Civitas [DNKS10] (see Figure 4(c)), and OpinionSeer [WWL*10]. In general, these techniques fit the discussion of popular categories and temporal trends above well, thus reinforcing its validity.

---

Computer Graphics Forum published by John Wiley & Sons Ltd.
7. Conclusions

In this paper, we have analyzed the state of the art in visualization of sentiment detected in text data. We have discussed 132 visualization techniques originating from peer-reviewed publications using a fine-grained categorization comprising 7 groups with 35 categories in total. We have also introduced an interactive survey browser which supports the categorization, and discussed insights as well as opportunities for future research in sentiment visualization. The collected survey data indicates the growing multidisciplinary interest for visualization of sentiment with regard to multiple data domains and tasks. Our future work on this survey includes updates to the survey data set, eventual refinements of the categorization, additional features for the interactive browser, and more analytical work based on the resulting changes.

Acknowledgements

This work was supported by the framework grant ‘The Digitized Society—Past, Present, and Future’ from the Swedish Research Council (Vetenskapsrådet) [grant number 2012-5659]. We would like to thank Magnus Sahlgren for his expertise related to CL/NLP aspects. We are also thankful to the authors of original techniques for providing us with permissions and images of their tools.

References

[AC07] Abbasi A., Chen H.: Categorization and analysis of text in computer mediated communication archives using visualization. In Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries (2007), JCDL ’07, ACM, pp. 11–18.

[ADAC*13] Amigó E., de Albornoz J. C., Chugur I., Corujo A., Gonzalo J., Martín T., Meid E., de Rijke M., Spina D.: Overview of Replab 2013: Evaluating online reputation monitoring systems. In Information Access Evaluation. Multilinguality, Multimodality, and Visualization. Springer, 2013, pp. 333–352.

[AK08] Annet M., Kondrak G.: A comparison of sentiment analysis techniques: Polarizing movie blogs. In Advances in Artificial Intelligence, Bergler S., (Ed.), vol. 5032 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2008, pp. 25–35.

[Alm13] Almutairi B. A. A.: Visualizing patterns of appraisal in texts and corpora. Text & Talk 33, 4–5 (2013), 691–723.

[APV11] Adams B., Phung D., Venkatesh S.: Eventscapes: Visualizing events over time with emotive facets. In Proceedings of the 19th ACM International Conference on Multimedia (2011), MM ’11, ACM, pp. 1477–1480.

[AYHK11] Alper B., Yang H., Haber E., Kandogan E.: OpinionBlocks: Visualizing consumer reviews. In Proceedings of the IEEE Workshop on Interactive Visual Text Analytics for Decision Making (2011), TextVis ’11.

[BA16] Bembenik R., Andruszkiewicz P.: Towards automatic argument extraction and visualization in a deliberative model of online consultations for local governments. In Proceedings of the 20th East European Conference on Advances in Databases and Information Systems (ADBIS 2016), Pokorný J., Ivanović M., Thalheim B., Šaloun P., (Eds.), vol. 9809 of Lecture Notes in Computer Science. Springer International Publishing, 2016, pp. 74–86.

[BB15] Brath R., Banissi E.: Using text in visualizations for micro/macro readings. In Proceedings of the IUI Workshop on Visual Text Analytics (2015), TextVis ’15.

[BB16] Brath R., Banissi E.: Typographic sets: Labeled set elements with font attributes. In Proceedings of the International Workshop on Set Visualization and Reasoning (2016), SetVR 2016, pp. 29–43.

[BE10] Brody S., Elhadad N.: An unsupervised aspect-sentiment model for online reviews. In Proceedings of Human Language
Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics (2010), HLT ’10, Association for Computational Linguistics, pp. 804–812.

[Ber83] BERTIN J.: Semiology of Graphics: Diagrams, Networks, Maps. University of Wisconsin Press, 1983. Translated by W. J. Berg.

[BES10] BACCINELLA S., ESULI A., SEBASTIANI F.: SentiWordNet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In Proceedings of the Seventh Conference on International Language Resources and Evaluation (2010), vol. 10 of LREC ’10, European Language Resources Association, pp. 2200–2204.

[BGAC11] BREW A., GREENE D., ARCHAMBAULT D., CUNNINGHAM P.: Deriving insights from national happiness indices. In Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops (2011), ICDMW ’11, pp. 53–60.

[BBHS12] BRĂȘOVEANU A. M., HUBMANN-HARDOVGL A., SCHARL A.: Interactive visualization of emerging topics in multiple social media streams. In Proceedings of the International Conference on Advanced Visual Interfaces (2012), AVI ’12, ACM, pp. 530–533.

[BKW16] BECK F., KOCH S., WEISKOPF D.: Visual analysis and dissemination of scientific literature collections with SurVis. IEEE Transactions on Visualization and Computer Graphics 22, 1 (Jan. 2016), 180–189.

[Bou15] BOUMAZA A.: A survey on sentiment analysis and visualization. Journal of Emerging Technologies in Web Intelligence 7, 1 (Nov. 2015), 35–43.

[BRT*14] BROOKS M., ROBINSON J. J., TORKILDSON M. K., HONG S. R., ARAGON C. R.: Collaborative visual analysis of sentiment in Twitter events. In Proceedings of the 11th International Conference on Cooperative Design, Visualization, and Engineering (CDVE ’14), LUO Y., (Ed.), vol. 8683 of Lecture Notes in Computer Science. Springer International Publishing, 2014, pp. 1–8.

[BW08] BYRON L., WATTENBERG M.: Stacked graphs — geometry & aesthetics. IEEE Transactions on Visualization and Computer Graphics 14, 6 (Nov. 2008), 1245–1252.

[CAHF14] CALDERON N. A., ARIAS-HERNANDEZ R., FISHER B.: Studying animation for real-time visual analytics: A design study of social media analytics in emergency management. In Proceedings of the 47th Hawaii International Conference on System Sciences (2014), HICSS ’14, pp. 1364–1373.

[CBK*17] CHEN N.-C., BROOKS M., KOCIELNIK R., HONG S. R., SMITH J., LIN S., QU Z., ARAGON C.: Lariat: A visual analytics tool for social media researchers to explore Twitter datasets. In Proceedings of the 50th Hawaii International Conference on System Sciences (2017), HICSS ’17, pp. 1881–1890.

[CCT15] CHEN Y.-S., CHEN L.-H., TAKAMA Y.: Proposal of LDA-based sentiment visualization of hotel reviews. In Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (2015), ICDMW 2015, pp. 687–693.

[CCYT15] CHEN Y.-S., CHEN L.-H., YAMAGUCHI T., TAKAMA Y.: Visualization system for analyzing user opinion. In Proceedings of the IEEE/SCIE International Symposium on System Integration (2015), SII 2015, pp. 646–649.

[CDG*14] CRESCI S., D’ERRICO A., GAZZÉ D., LO DUCA, A., MARCHETTI A., TESCONI M.: Tour-pedia: A web application for sentiment visualization in tourism domain. In Proceedings of the LREC 2014 OpeNER Workshop (2014), pp. 18–21.

[CFKA14] CHEN N.-C., FELDMAN L. B., KROLL J. F., ARAGON C. R.: Emoticons and linguistic alignment: How visual analytics can elicit storytelling. In Poster Abstracts of IEEE VAST (2014), VAST ’14, pp. 237–238.

[CGG*15] CERVANTES O., GUTIÉRREZ F., GUTIÉRREZ E., CASTILLO E., SÁNCHEZ J. A., WAN W.: Expression: Visualizing affective content from social streams. In Proceedings of the Latin American Conference on Human Computer Interaction (2015), LCIHC’15, ACM, pp. 10:1–10:8.

[CISSW06] CHEN C., IBEKWE-SANJUAN F., SANJUAN E., WEAVER C.: Visual analysis of conflicting opinions. In Proceedings of the IEEE Symposium on Visual Analytics Science and Technology (2006), VAST ’06, IEEE, pp. 59–66.

[CK15] CERNEA D., KERKEN A.: A survey of technologies on the rise for emotion-enhanced interaction. Journal of Visual Languages and Computing 31, Part A (2015), 70–86.

[CLL*15] CAO N., LI L., LIN Y.-R., WANG F., WEI Z.: SocialHelix: visual analysis of sentiment divergence in social media. Journal of Visualisation 18, 2 (2015), 221–235.

[CLS*12] CAO N., LIN Y.-R., SUN X., LAZER D., LIU S., QU H.: Whisper: Tracing the spatiotemporal process of information diffusion in real time. IEEE Transactions on Visualization and Computer Graphics 18, 12 (Dec. 2012), 2649–2658.

[Coh88] COHEN L. H.: Measurement of life events. In Life events and psychological functioning: Theoretical and methodological issues. SAGE Publications, 1988, pp. 11–30.

[COR14] CAMBRIA E., OLIVER D., RAAGopal D.: SenticNet 3: A common and common-sense knowledge base for cognition-driven sentiment analysis. In Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence (2014), AAAI ’14, AAAI Press, pp. 1515–1521.

[CQZ*12] CUI W., QU H., ZHOU H., ZHANG W., SKIENA S.: Watch the story unfold with TextWheel: Visualization of large-scale news streams. ACM Transactions on Intelligent Systems and Technology 3, 2 (Feb. 2012), 20:1–20:17.

© 2017 The Authors
Computer Graphics Forum published by John Wiley & Sons Ltd.
for document exploration and sensemaking in Jigsaw. *IEEE Transactions on Visualization and Computer Graphics* 19, 10 (2013), 1646–1663.

**[GOCI12]** GALI G., OLIVER S., CHEVALIER F., DIAMOND S.: Visualizing sentiments in business-customer relations with metaphors. In *CHI ’12 Extended Abstracts on Human Factors in Computing Systems* (2012), CHI EA ’12, ACM, pp. 1493–1498.

**[GREA15]** GOLD V., ROHRDANTZ C., EL-ASSADY M.: Exploratory text analysis using lexical episode plots. In *Short Papers of the EG/VGTC Conference on Visualization* (2015), Bertini E., Kennedy J., Puppo E., (Eds.), EuroVis ’15, The Eurographics Association.

**[Guz13]** GUZMAN E.: Visualizing emotions in software development projects. In *Proceedings of the 2013 First IEEE Working Conference on Software Visualization* (2013), VISSOFT ’13, IEEE, pp. 1–4.

**[HBB*15]** HENNIG P., BERGER P., BREHM M., GRAS Nick B., HERDT J., MEINEL C.: Hot spot detection — an interactive cluster heat map for sentiment analysis. In *Proceedings of the 2015 IEEE International Conference on Data Science and Advanced Analytics* (2015), DSAA 2015.

**[HBM*14]** HENNIG P., BERGER P., MEINEL C., PIRL L., SCHULZE L.: Exploring emotions over time within the blogosphere. In *Proceedings of the 2014 International Conference on Data Science and Advanced Analytics* (2014), DSAA 2014, pp. 587–592.

**[HC14]** HOQUE E., CARENINI G.: ConVis: A visual text analytic system for exploring blog conversations. *Computer Graphics Forum* 33, 3 (2014), 221–230.

**[HC16]** HOQUE E., CARENINI G.: MultiConVis: A visual text analytics system for exploring a collection of online conversations. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (2016), IUI ’16, ACM, pp. 96–107.

**[He12]** HE Y.: A Bayesian modeling approach to multi-dimensional sentiment distributions prediction. In *Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining* (2012), WISDOM ’12, ACM, pp. 1:1–1:8.

**[HHM*16]** HOEBER O., HOEBER L., MISSEERY M. E., ODORI K., GOP R.: Visual Twitter analytics (Vista): Temporally changing sentiment and the discovery of emergent themes within sport event tweets. *Online Information Review* 40, 1 (2016), 25–41.

**[HHN00]** HAVRE S., HETZLER B., NOWELL L.: ThemeRiver: visualizing theme changes over time. In *Proceedings of the IEEE Symposium on Information Visualization* (2000), InfoVis ’00, IEEE, pp. 115–123.

**[HKH*14]** HA H., KIM G.-n., HWANG W., CHOI H., LEE K.: CosMovis: Analyzing semantic network of sentiment words in movie reviews. In *Poster Abstracts of IEEE LDAV* (2014), LDAV ’14, pp. 113–114.

**[HRJ*13]** HAO M. C., ROHRDANTZ C., JANETZKO H., KEIM D. A., DAVAI U., HAUG L.-E., HSU M., STOFFEL F.: Visual sentiment analysis of customer feedback streams using geo-temporal term associations. *Information Visualization* 12, 3–4 (2013), 273–290.

**[Hun11]** HUNSTON S.: *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. Routledge, 2011.

**[JO11]** JO Y., OH A. H.: Aspect and sentiment unification model for online review analysis. In *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining* (2011), WSDM ’11, ACM, pp. 815–824.

**[KAF*08]** KEIM D., ANDRIENKO G., FEKETE J.-D., GÖRG C., KÖHLHAMMER J., MELANÇON G.: Visual analytics: Definition, process, and challenges. In *Information Visualization: Human-Centered Issues and Perspectives*, Kerren A., Stasko J. T., Fekte J.-D., North C., (Eds.), vol. 4950 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2008, pp. 154–175.

**[KBR*12]** KUSENOK K., BROOKS M., ROBINSON J. J., PERRY D., TORKILDSON M. K., ARAGON C.: Automating large-scale annotation for analysis of social media content. In *Proceedings of the 2nd IEEE Workshop on Interactive Visual Text Analytics “Task-Driven Analysis of Social Media”* (2012), TextVis ’12.

**[KJC16]** KRADECINAC U., JOVANOVIC J., DEVEDZIC V., PASQUER P.: Textual affect communication and evocation using abstract generative visuals. *IEEE Transactions on Human-Machine Systems* 46, 3 (June 2016), 370–379.

**[KK14]** KUCHER K., KERRER A.: Text visualization browser: A visual survey of text visualization techniques. In *Poster Abstracts of IEEE VIS* (2014), InfoVis ’14.

**[KK15]** KUCHER K., KERRER A.: Text visualisation techniques: Taxonomy, visual survey, and community insights. In *Proceedings of the 8th IEEE Pacific Visualization Symposium* (2015), PacificVis ’15, IEEE, pp. 117–121.

**[KKL*16]** Kwon B. C., Kim S.-H., Lee S., Cho J., Huh J., Yi J. S.: VisOHIC: Designing visual analytics for online health communities. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (Jan. 2016), 71–80.

**[KL14]** KIM K., Lee J.: Sentiment visualization and classification via semi-supervised nonlinear dimensionality reduction. *Pattern Recognition* 47, 2 (2014), 758–768.

**[KR11]** KANG X., REN F.: Sampling latent emotions and topics in a hierarchical bayesian network. In *Proceedings of the 2011 7th International Conference on Natural Language Processing and Knowledge Engineering* (2011), NLP-KE ’11, IEEE, pp. 37–42.

**[KRHW12]** KRSTAJIC M., ROHRDANTZ C., HUDN M., WEAVER A.: Getting there first: Real-time detection of real-world incidents on Twitter. In *Proceedings of the 2nd IEEE Workshop on Interactive Visual Text Analytics “Task-Driven Analysis of Social Media”* (2012), TextVis ’12.
[KSBK*16] Kucher K., Schamp-Bierede T., Kerren A., Paradis C., Sahlgren M.: Visual analysis of online social media to open up the investigation of stance phenomena. Information Visualization 15, 2 (2016), 93–116.

[KSMP14] Kemper R., Sintsova V., Musat C., Pu P.: Emotion-Watch: Visualizing fine-grained emotions in event-related tweets. In Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media (2014), ICWSM ’14, pp. 236–245.

[KSP*15] Kranc J., Smailović J., Podpečan V., Grčar M., Žnidarič M., Lavrač N.: Active learning for sentiment analysis on data streams: Methodology and workflow implementation in the CloudFlows platform. Information Processing & Management 51, 2 (2015), 187–203.

[KTLS14] Kuang C., Tang J., Liu Z., Sun M.: ImgWordle: Image and text visualization for events in microblogging services. In Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces (2014), AVI ’14, ACM, pp. 371–372.

[KWS14] Kumamoto T., Wada H., Suzuki T.: Visualizing temporal changes in impressions from tweets. In Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services (2014), ůWAS ’14, ACM, pp. 116–125.

[LDS10] Li X., Dai L., Shi H.: Opinion mining of camera reviews based on semantic role labeling. In Proceedings of the Seventh International Conference on Fuzzy Systems and Knowledge Discovery (2010), vol. 5 of FSKD ’10, pp. 2372–2375.

[LFO*10] Lee H., Ferguson P., O’Hare N., Gurrin C., Sweatman A. F.: Integrating interactivity into visualising sentiment analysis of blogs. In Proceedings of the First International Workshop on Intelligent Visual Interfaces for Text Analysis (2010), IVITA ’10, ACM, pp. 17–20.

[LHC05] Liu B., Hu M., Cheng J.: Opinion Observer: Analyzing and comparing opinions on the Web. In Proceedings of the 14th International Conference on World Wide Web (2005), WWW ’05, ACM, pp. 342–351.

[LHW*15] Lu Y., Hu X., Wang F., Kumar S., Liu H., Maciejewski R.: Visualizing social media sentiment in disaster scenarios. In Proceedings of the 24th International Conference on World Wide Web (2015), WWW ’15 Companion, ACM, pp. 1211–1215.

[LKT*14] Lu Y., Krüger R., Thom D., Wang F., Koch S., Ertl T., Maciejewski R.: Integrating predictive analytics and social media. In Proceedings of the IEEE Conference on Visual Analytics Science and Technology (2014), VAST ’14, IEEE, pp. 193–202.

[LR09] Li H., Ren F.: The study on text emotional orientation based on a three-dimensional emotion space model. In Proceedings of the International Conference on Natural Language Processing and Knowledge Engineering (2009), NLP-KE ’09, IEEE.

[LSB*16] Lu Y., Stepien M., Burke S., Wang H., Tsai J.-Y., Davulcu H., Montgomery D., Corman S. R., Maciejewski R.: Exploring evolving media discourse through event cueing. IEEE Transactions on Visualization and Computer Graphics 22, 1 (Jan. 2016), 220–229.

[LSL03] Liu H., Selker T., Lieder H.: Visualizing the affective structure of a text document. In CHI ’03 Extended Abstracts on Human Factors in Computing Systems (2003), CHI EA ’03, ACM, pp. 740–741.

[LWW*16] Li Q., Wu Y., Wang S., Lin M., Feng X., Wang H.: VisTravel: visualizing tourism network opinion from the user generated content. Journal of Visualization 19, 3 (2016), 489–502.

[LXG*16] Liu X., Xu A., Gou L., Liu H., Akkiraju R., Shen H.-W.: SocialBrands: Visual analysis of public perceptions of brands on social media. In Proceedings of the IEEE Conference on Visual Analytics Science and Technology (2016), VAST ’16, pp. 116–127.

[MBB*11] Marcus A., Bernstein M. S., Badar O., Karger D. R., Madden S., Miller R. C.: TwitInfo: Aggregating and visualizing microblogs for event exploration. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2011), CHI ’11, ACM, pp. 227–236.

[MBM14] Makki R., Brooks S., Milios E. E.: Context-specific sentiment lexicon expansion via minimal user interaction. In Proceedings of the International Conference on Information Visualization Theory and Applications (2014), IVAPP ’14, SciTePress, pp. 178–186.

[MDR06] Misine G., De Rijke M.: MoodViews: Tools for blog mood analysis. In Proceedings of the AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs (2006), pp. 153–154.

[MFX*16] Mahmud J., Fei G., Xu A., Pal A., Zhou M.: Predicting attitude and actions of Twitter users. In Proceedings of the 21st International Conference on Intelligent User Interfaces (2016), IUI ’16, ACM, pp. 2–6.

[Mil95] Miller G. A.: WordNet: A lexical database for English. Commun. ACM 38, 11 (Nov. 1995), 39–41.

[MKS*16] Mohammad S. M., Kiritchevskiy S., Sobhani P., Zhu X., Cherry C.: SemEval-2016 task 6: Detecting stance in tweets. In Proceedings of the International Workshop on Semantic Evaluation (2016), SemEval ’16.
[RR15] RAVI K., RAVI V.: A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. Knowledge-Based Systems 89 (2015), 14–46.

[ŠB10] ŠILIĆ A., BAŠIĆ B. D.: Visualization of text streams: A survey. In Knowledge-Based and Intelligent Information and Engineering Systems, Setchi R., Jordanov I., Howlett R. J., Jain L. C., (Eds.), vol. 6277 of Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2010, pp. 31–43.

[SBBI*15] STEED C. A., BEAVER J., BOGEN II P. L., DROUHARD M., PYLE J.: Text stream trend analysis using multiscale visual analytics with applications to social media systems. In Proceedings of the IUI Workshop on Visual Text Analytics (2015), TextVis ’15.

[SBT15] SHAMIM A., BALAKRISHNAN V., TAHIR M.: Evaluation of opinion visualization techniques. Information Visualization 14, 4 (2015), 339–358.

[Sch11] SCHULZ H.-J.: TreeVis.net: A tree visualization reference. IEEE Computer Graphics and Applications 31, 6 (Nov. 2011), 11–15. URL: http://treevis.net.

[SH01] SUBASIC P., HEUTTNER A.: Affect analysis of text using fuzzy semantic typing. IEEE Transactions on Fuzzy Systems 9, 4 (Aug. 2001), 483–496.

[SHFA16] SAIF H., HE Y., FERNANDEZ M., ALANI H.: Contextual semantics for sentiment analysis of Twitter. Information Processing & Management 52, 1 (2016), 5–19.

[SHHJ*16] SCARL A., HUBMWANN-HAIDVOGEL A., JONES A., FISCHER D., KAMOLOV R., WEICHSELBAUN A., RAEFELBERGER W.: Analyzing the public discourse on works of fiction–detection and visualization of emotion in online coverage about HBO’s Game of Thrones. Information Processing & Management 52, 1 (2016), 129–138.

[SHHW*12] SCARL A., HUBMWANN-HAIDVOGEL A., WIECHSELBAUN A., WOHLGENANT G., LANG H.-P., SABO M.: Extraction and interactive exploration of knowledge from aggregated news and social media content. In Proceedings of the 4th ACM SIGCHI Symposium on Engineering Interactive Computing Systems (2012), EICS ’12, ACM, pp. 163–168.

[Shn96] SHNEIDERMAN B.: The eyes have it: a task by data type taxonomy for information visualizations. In Proceedings of the IEEE Symposium on Visual Languages (1996), VL ’96, pp. 336–343.

[SHS*16] SUNG C.-Y., HUANG X.-Y., SHEN Y., CHERNG F.-Y., LIN W.-C., WANG H.-C.: ToPIN: A visual analysis tool for time-anchored comments in online educational videos. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (2016), CHI EA ’16, ACM, pp. 2185–2191.

[Sma11] SMALL H.: Interpreting maps of science using citation context sentiments: a preliminary investigation. Scientometrics 87, 2 (2011), 373–388.

[SPH*11] SOCHER R., PENNINGTON J., HUANG E., NG A., MANNING C.: Semi-supervised recursive autoencoders for predicting sentiment distributions. In Proceedings of the Conference on Empirical Methods in Natural Language Processing (2011), EMNLP ’11, Association for Computational Linguistics, pp. 151–161.

[SPW*13] SOCHER R., PERELYGIN A., WU J., CHUANG J., MANNING C. D., NG A. Y., PORTS C.: Recursive deep models for semantic compositionality over a sentiment treebank. In Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (2013), EMNLP ’13, Association for Computational Linguistics, pp. 1631–1642.

[SQC16] SITBSMA B., QVARFORDT P., CHEN F.: Tweetviz: Visualizing tweets for business intelligence. In Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval (2016), SIGIR ’16, ACM, pp. 1153–1156.

[SS10] SCHÄFER U., SPURK C.: TAKE Scientist’s Workbench: Semantic search and citation-based visual navigation in scholar papers. In Proceedings of the IEEE Fourth International Conference on Semantic Computing (2010), ISCS ’10, pp. 317–324.

[SV04] STRAPPARAVA C., VALITUTTI A.: WordNet Affect: an affective extension of WordNet. In Proceedings of the 4th International Conference on Language Resources and Evaluation (2004), vol. 4 of LREC ’04, European Language Resources Association, pp. 1083–1086.

[TA] TOMINSKI C., AIGNER W.: The TimeVis Browser. Accessed May 10, 2017. URL: http://survey.timeviz.net

[TA03] TASKAYA T., AHMAD K.: Bimodal visualisation: a financial trading case study. In Proceedings of Seventh International Conference on Information Visualization (2003), IV ’03, IEEE, pp. 320–326.

[Tay90] TAYLOR R.: Interpretation of the correlation coefficient: A basic review. Journal of Diagnostic Medical Sonography 6, 1 (1990), 35–39.

[TC06] TAT A., CARPENDALE S.: CrystalChat: Visualizing personal chat history. In Proceedings of the 39th Annual Hawaii International Conference on System Sciences (2006), vol. 3 of HICSS ’06, IEEE.

[TKT*00] TRAFTON J. G., KIRSCHENBAUM S. S., TSUI T. L., MIYAMOTO R. T., BALLAS J. A., RAYMOND P. D.: Turning pictures into numbers: extracting and generating information from complex visualizations. International Journal of Human-Computer Studies 53, 5 (2000), 827–850.

[TP10] TAUSSCHER Z. H., PENNBECKER J. W.: The psychological meaning of words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology 29, 1 (2010), 24–54.

© 2017 The Authors

Computer Graphics Forum published by John Wiley & Sons Ltd.
Text Analytics “Task-Driven Analysis of Social Media” (2012), TextVis ’12.

[WXL*13] WANG C., XIAO Z., LIU Y., XU Y., ZHOU A., ZHANG K.: SentiView: Sentiment analysis and visualization for Internet popular topics. *IEEE Transactions on Human-Machine Systems* 43, 6 (Nov. 2013), 620–630.

[XLLS11] XU K., LIAO S. S., LI J., SONG Y.: Mining comparative opinions from customer reviews for competitive intelligence. *Decision Support Systems* 50, 4 (2011), 743–754.

[YN05] YOST B., NORTH C.: Single complex glyphs versus multiple simple glyphs. In *CHI ’05 Extended Abstracts on Human Factors in Computing Systems* (2005), CHI EA ’05, ACM, pp. 1889–1892.

[YNTT11] YATANI K., NOVATI M., TRUSTY A., TRUONG K. N.: Review Spotlight: A user interface for summarizing user-generated reviews using adjective-noun word pairs. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011), CHI ’11, ACM, pp. 1541–1550.

[YWC*16] YU Z., WANG Z., CHEN L., GUO B., LI W.: Featuring, detecting, and visualizing human sentiment in Chinese micro-blog. *ACM Transactions on Knowledge Discovery from Data* 10, 4 (May 2016), 48:1–48:23.

[ZCW*14] ZHAO J., CAO N., WEN Z., SONG Y., LIN Y.-R., COLLINS C.: #FluxFlow: Visual analysis of anomalous information spreading on social media. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (Dec. 2014), 1773–1782.

[ZDWX12] ZHAO J., DONG L., WU J., XU K.: MoodLens: An emoticon-based sentiment analysis system for chinese tweets. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2012), KDD ’12, ACM, pp. 1528–1531.

[ZGWZ14] ZHAO J., GOU L., WANG F., ZHOU M.: PEARL: An interactive visual analytic tool for understanding personal emotion style derived from social media. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology* (2014), VAST ’14, IEEE, pp. 203–212.

[ZKKT09] ZHANG J., KAWAI Y., KUMAMOTO T., TANAKA K.: A novel visualization method for distinction of web news sentiment. In *Proceedings of the 10th International Conference on Web Information Systems Engineering* (WISE ’09), Vossen G., Long D. D. E., Yu J. X., (Eds.), vol. 5802 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2009, pp. 181–194.

[ZLW13] ZHANG C., LIU Y., WANG C.: Time-space varying visual analysis of micro-blog sentiment. In *Proceedings of the 6th International Symposium on Visual Information Communication and Interaction* (2013), VINCI ’13, ACM, pp. 64–71.

[ZQLT16] ZHAO Y., QIN B., LIU T., TANG D.: Social sentiment sensor: a visualization system for topic detection and topic sentiment analysis on microblog. *Multimedia Tools and Applications* 75, 15 (2016), 8843–8860.

[ZTTL16] ZHAO S., TONG Y., LIU X., TAN S.: Correlating Twitter with the stock market through non-Gaussian SVAR. In *Proceedings of the Eighth International Conference on Advanced Computational Intelligence* (2016), ICACI 2016, pp. 257–264.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher’s web site:

Table S1: The summary of our survey.

Table S2: The summary of our survey (continued).

Table S3: The summary of our survey (continued).

Table S4: Citations counts for publications included in our survey (retrieved from Google Scholar on April 12, 2017).