Visual analysis of online social media to open up the investigation of stance phenomena

Kucher, Kostiantyn; Schamp-Bjerede, Teri; Kerren, Andreas; Paradis, Carita; Sahlgren, Magnus

Published in:
Information Visualization

DOI:
10.1177/1473871615575079

2016

Info Vis

Article

Information Visualization
2016, Vol. 15(2) 93–116

DOI: 10.1177/1473871615575079
ivi.sagepub.com
SAGE

Visual analysis of online social media to open up the investigation of stance phenomena

Kostiantyn Kucher[1], Teri Schamp-Bjerede[2], Andreas Kerren[1], Carita Paradis[2] and Magnus Sahlgren[3]

Abstract

Online social media are a perfect text source for stance analysis. Stance in human communication is concerned with speaker attitudes, beliefs, feelings and opinions. Expressions of stance are associated with the speakers' view of what they are talking about and what is up for discussion and negotiation in the intersubjective exchange. Taking stance is thus crucial for the social construction of meaning. Increased knowledge of stance can be useful for many application fields such as business intelligence, security analytics, or social media monitoring. In order to process large amounts of text data for stance analyses, linguists need interactive tools to explore the textual sources as well as the processed data based on computational linguistics techniques. Both original texts and derived data are important for refining the analyses iteratively. In this work, we present a visual analytics tool for online social media text data that can be used to open up the investigation of stance phenomena. Our approach complements traditional linguistic analysis techniques and is based on the analysis of utterances associated with two stance categories: sentiment and certainty. Our contributions include (1) the description of a novel web-based solution for analyzing the use and patterns of stance meanings and expressions in human communication over time; and (2) specialized techniques used for visualizing analysis provenance and corpus overview/navigation. We demonstrate our approach by means of text media on a highly controversial scandal with regard to expressions of anger and provide an expert review from linguists who have been using our tool.

Keywords

Visual analytics, visualization, text visualization, interaction, time-series, stance analysis, sentiment analysis, text analytics, visual linguistics, online social media, text and document data

Introduction

The vast amount of digital data available online provides unprecedented opportunities for automated analyses. For example, text data of all kinds make it possible for researchers in the field of linguistics to employ a bottom-up approach to understand various aspects of language: while the traditional way of manual text investigation involved static corpora, linguists nowadays can analyze text data that reflect global events and ongoing language evolution. The research on specific language phenomena benefits from text data collected from web sources such as online social media (Twitter, Facebook, blogs, forums, etc.). Those

[1]Department of Computer Science, Linnaeus University, Växjö, Sweden
[2]Centre for Languages and Literature, Lund University, Lund, Sweden
[3]Gavagai AB, Stockholm, Sweden

Corresponding author:
Kostiantyn Kucher, Department of Computer Science, Linnaeus University, Vejdes Plats 7, SE-351 95 Växjö, Sweden.
Email: kostiantyn.kucher@lnu.se

texts are typically created by multiple authors who are engaged in discussions or refer to each other's messages in which they express their thoughts and opinions.

This presents an opportunity for researchers who are interested in stance analysis. *Stance* is a relatively broad concept in linguistics related to (inter-)subjectivity expressed in text or human conversation, for example, attitudes, feelings, perspectives, or judgments. Note that stance is not just another concept for subjectivity. It is beyond subjectivity in that the process of taking stance itself is evaluative and interactional. Stance could be viewed as a concept that includes sentiment, certainty, and so on as its subcategories. Analyzing these subcategories leads toward better understanding of stance.

Research on stance includes both theoretical efforts (related to the definition and the knowledge about the nature of this phenomenon) and practical efforts (related to collecting evidence and explaining the means of taking stance), and it can lead to various text analytics applications. The practical tasks require processing large quantities of textual data that are infeasible for manual investigation, for example, providing a temporal overview of stance usage in social media, retrieving the corresponding text data relevant to stance phenomena, or analyzing the occurrences of stance expressions. Therefore, stance researchers are interested in automated ways of text processing that can be offered by researchers from the field of computational linguistics or natural language processing (NLP).

However, many linguists face difficulties when trying to interpret the output of NLP algorithms. For NLP experts, it is equally challenging to gain insight into the underlying text data and to provide useful feedback in order to refine their automatic analyses. In fact, NLP researchers would also benefit from a technique that could improve their understanding of the computational processes associated with the state-of-the-art NLP algorithms (e.g. it is difficult to interpret the state of a large artificial neural network just by weight matrices). This predicament can be resolved by introducing a visual analytics (VA) approach to provide linguistics researchers with interactive visualizations for analyzing large text data and for presenting the NLP experts with feedback at the same time. Our research project StaViCTA (Advances in the description and explanation of Stance in discourse using Visual and Computational Text Analytics (project web page: http://cs.lnu.se/stavicta/)) addresses this challenge and aims to produce a refined theory of stance, efficient interactive visualization, and computational techniques for its analysis, as well as solutions for specific applications. Due to the early stage of research in stance analysis, the project itself follows an iterative progress plan. Therefore, we consider sentiment analysis, including certainty or uncertainty, as underlying aspects of linguistic stance in order to support the construction of the model in general.

In this work, we focus on the exploration of social media documents (in English) and the collection of a training dataset which later will be used to develop appropriate machine learning (ML) approaches. The composed training data consist of text chunks, called *utterances*, that are associated with specific expressions of stance (see Figure 1). These utterances can be used for both NLP purposes and manual linguistic investigation; we denote them by *stance markers*. This collection of relevant stance markers is the basis for a refined theory and sophisticated NLP models for stance analysis in general.

Here, we present our tool called uVSAT that can help stance researchers to identify candidate documents that may contain stance expressions, analyze the document texts, and export the new stance markers (as introduced in our previous poster abstract[1]). uVSAT supports the research task of how we can study the use and patterns of stance meanings and stance expressions in human communication over time in order to investigate what stance markers and stance markings are used when, why, how, where, and in what type of dialogic sequences related to the contexts where they occur. Our effort described in this article is meant to complement the existing techniques for stance analysis based on manual close reading and traditional linguistic tools by introducing a VA approach to this problem, while not providing a completely automatic stance analysis yet. The main contributions of the VA approach presented in this article include the following:

- A web-based *VA solution* for investigating stance phenomena based on sentiment analyses of document texts and time-series;
- An *interactive history diagram* for document set queries that facilitates the analysis provenance;
- Interactive *aggregation charts* that provide document set overview, navigation, and comparison functionality with regard to stance types or specific stance markers.

The remainder of this article is organized as follows: the next section provides the background of stance analysis from the perspective of linguistics and NLP. The subsequent section covers the related work in text visualization, including work dedicated to sentiment analysis visualization. After this, we explain the system architecture and data model as well as user tasks supported by uVSAT. Then, we describe in

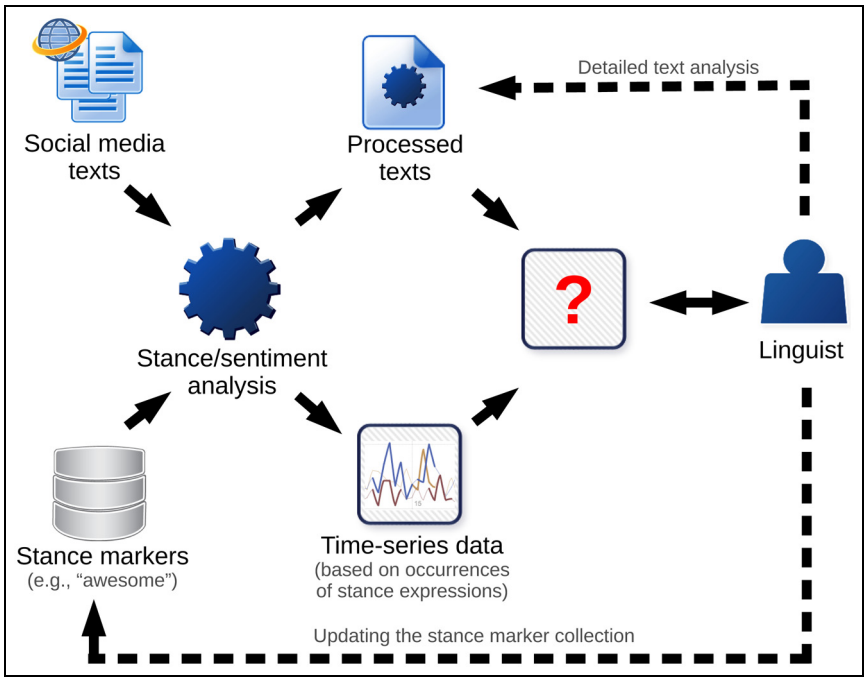

Figure 1. The diagram gives an overview of the underlying research problems from the user perspective. To succeed with the analysis of stance, linguists require means to analyze and interact with the output of NLP algorithms as well as means of further manual investigation. These means are still missing in the analysis loop and are indicated by the red question mark. The dashed edges denote the user operations that depend on the results of interactive visual analysis.

detail our visualization and interaction approaches for this tool. The subsequent section discusses a use case from the linguistics domain based on exploration of data with regard to anger sentiment as a subcategory of stance. The penultimate section provides the results of a domain expert review and our reflections about the tool. Finally, we summarize the contributions and future work in the last section.

Background

Our research on visual stance analytics is by nature tightly connected to the domains of linguistics and NLP. Since the problem of stance analysis is not widely discussed in the VA community (as opposed to sentiment analysis), we present the theoretical background of stance and its relation to sentiment in this section.

Stance and sentiment model

Stance is a topical area of interest in linguistics because the interactive nature of communication between individuals is considered vital. The function of taking stance in the communicative situation is to convey the speaker's viewpoint of what is talked about and to regulate the exchange between the dialog partners. Communication here works on more than the pure understanding of words. Words are always understood in the light of the contexts and the situations where they are used.[2,3] In doing so, language is used to recontextualize human experiences into written and spoken forms. Its social role is to affect the state of mind of other people and to negotiate meanings in order to bring about cognitive changes.[4,5] Language users construe their expressions to communicate their particular perspective and viewpoint of what is talked about. As the following scheme[6] demonstrates, this process of taking stance is evaluative and fundamentally interactional, a type of ongoing negotiation:

1. An utterance proposed by X;
2. Y's engagement (mental processing or interpretation or positioning) as to the utterance in context;
3. Y's response to X's utterance;
4. X's engagement (mental processing or interpretation or positioning) as to the utterance in context;
5. X's response to Y's utterance;
6. Repeat 2–6.

Ours is a broad understanding of the process of taking stance, as it is critical to address the subtle but important differences in how people create discourse—imbuing it with their personal word choices as distinct acts of taking stance. This encompasses expressions of subjectivity, ranging from individual words to larger chunks of text. These items

express speaker's (1) sentiments, (2) attitudes, and (3) beliefs, covering meanings of certainty, volition, evidence, emotion, valence, degree, and so on. Following Du Bois,[7] we divide the process of taking stance into three parts: (1) speaker evaluation of what is talked about, (2) speaker positioning (epistemicity), and (3) alignment in communication, that is, establishment of agreement or disagreement. Stance has been studied under different headings and scope, such as evaluation,[8,9] sentiment,[10] and appraisal,[11] and, of course, under the title stance itself.[5,12–14] Yet, at the present time, there is no conclusive and universally accepted definition of linguistic stance.

As stated above, subcategories of stance include sentiment, certainty or uncertainty, as well as other subcategories that are not well-defined yet. For this article, we have limited the scope of our understanding of stance to sentiment and certainty or uncertainty. These subcategories are generally considered to describe the feelings and assessments of an utterance; as such, they can encapsulate an evaluative statement that is deemed to be a stance act. Our approach is based on the expectation that the occurrences of such expressions lead to occurrences of other stance expressions—we denote the particular analyzed subcategories by *stance types* throughout this article to simplify the notation. From the computational perspective, this approach could be described as "multidimensional" sentiment analysis.

Sentiment analysis

From an operational point of view, stance includes phenomena such as subjectivity, sentiment, belief, trust, and uncertainty. Some of these phenomena, such as sentiment and subjectivity, have enjoyed considerable attention in the NLP community (for instance, see the works of Pang and Lee,[15] Liu,[16] and Lin et al.[17]), while others, such as belief, trust, or uncertainty, have remained comparatively peripheral (but there is a number of efforts[18,19] to analyze uncertainty and speculation, respectively). Sentiment analysis in particular has become a staple in NLP, both in research and in commercial applications, with a large number of vendors offering solutions for social media monitoring where sentiment analysis is an important part of the analytics suite.

As with any research area that gains popularity in a research community, there has been a wide variety of approaches suggested in the literature. Examples range from simple keyword matching[20] over standard machine learning techniques[15,21] to the use of topic modeling algorithms and latent variable models[22–24] to deep learning architectures.[25,26] State-of-the-art approaches to sentiment analysis now approach, and in some cases even exceed, 90% accuracy on standardized benchmark test suites.[21,27,28]

Sentiment analysis is normally considered as a classification problem over two or three classes, where *positive* and *negative* define the basic polarity, and *neutral* is used to describe a lack of attitudinal content. From the perspective of stance analysis, this is a very simplistic ontology of emotions that is likely to be too restricted to be useful for analyzing and describing complex interpersonal processes of taking stance. Current research on sentiment analysis is beyond the standard positive–negative dichotomy and operates over a wider spectrum of emotions, such as Ekman's[29] six basic emotions (the so-called Big Six): *anger*, *fear*, *happiness*, *surprise*, *disgust*, and *sadness*,[30] or some other multi-class taxonomy of sentiments.[25,31] Another example of a more complex sentiment palette is the RepTrak model used in the RepLab evaluation campaign that includes eight different categories designed specifically for reputation classification.[32]

As opposed to some of more complex approaches based on ML, we opt for a simplistic approach to sentiment classification for the purposes of the visualization tool in order to preserve transparency and simplicity. As previously noted, we have chosen to address stance through subcategories. More specifically in uVSAT, these are based on Ekman's Big Six emotions, employing the NLP solution of simple lexical matching over lists of attitude terms (which we call stance markers as already mentioned in the "Introduction" section). The main goal at this stage of the project is to facilitate experiments to further improve our understanding of stance in general and our analysis techniques in particular. While our method of sentiment analysis is simple, such a lexical-based approach is still widely used by visualization and VA solutions,[33,34] especially the ones aiming for high performance when processing large amounts of input data.[35] There are also several examples of combining both lexical-based and machine learning–based approaches for sentiment analysis that reports similar[36] or even surprisingly good[37] results when using the lexical approach.

Related work

Our tool uVSAT was designed to visualize and interact with large text data sources as well as the results of automatic text processing which include time-series. There have recently been multiple works dedicated to text visualization and analytics of social media. Survey articles by Alencar et al.,[38] Gan et al.,[39] Kerren et al.,[40] and Kucher and Kerren[41] demonstrate a variety of techniques used for the visualization of single documents, document collections (corpora),

and text-related data streams. In this section, we will discuss several groups of works relevant to our research from various aspects.

Time-dependent text visualization

A good number of such works address temporal aspects to visualize events, topic competition or evolution, or other time-dependent data. While some of them introduce novel metaphors for visual encoding, multiple techniques combine well-known representations such as line plots, river metaphors, or animated force-directed graphs. Havre et al.[42] introduce ThemeRiver, the original technique for temporal data visualization based on a river metaphor that is designed to depict topic evolution in document collections. Dou et al.[43] combine trees, text tags, and rivers in their HierarchicalTopics system to visualize the temporal evolution of topics in corpora. Xu et al.[44] combine line plots, stacked charts, and word clouds to depict topic competition in social media document collections. To support the real-time monitoring of streaming Twitter data backed up with automatic text classification, Bosch et al.[45] use timeline, word clouds, glyphs, and maps in the ScatterBlogs2 system. For the work in this article, we decided to choose simple visual representations (line plots, text tags, and bubble charts) for the data currently available to us, although we plan to design more specialized visual encodings for other tasks in the future.

Sentiment visualization

While specific problems (and the corresponding analysis techniques) such as topic modeling and event detection have been very popular in text visualization, the interest for sentiment analysis and visualization is also arising in the VA community. Liu et al.[46] and Oelke et al.[47] describe visualizations for opinion mining of reviews. Wanner et al.,[48] Cui et al.,[49] and Rohrdantz et al.[50] present approaches for visual sentiment analysis that supports temporal data. Görg et al.[37] describe the fluid integration of sentiment analysis as well as other computational text analyses with interactive visualizations in their system Jigsaw. Online social media data are used for visual sentiment analysis by Wanner et al.,[51] Zhang et al.,[52] and Hao et al.[53] SentiView, introduced by Wang et al.,[54] not only facilitates temporal sentiment analysis but also augments it with relation analysis based on graph representation—this is relevant to our long-term research goals involving intersubjectivity and stance analysis. The recent work of Zhao et al.[33] describes PEARL, a VA system for multidimensional personal emotion or sentiment visualization of Twitter posts over time, and uses an approach similar to ours (based on lexical matching of *emotional words* pertaining to eight emotion categories and three additional emotion dimensions)—however, our work focuses on the analysis and visualization of data related to multiple posters and sources, and we are interested in categories beyond emotions or sentiment. In general, most of the discussed works involve sentiment analysis as a means rather than the object of research. Our approach, in contrast to theirs, focuses on the analysis of sentiment to bootstrap the research on visual stance analysis. This leads us to the involvement of experts in linguistics as users and the discussion of existing visualization approaches related to the domain of linguistics.

Visualization for linguistic research

InfoVis and VA techniques have been used to facilitate tasks such as the analysis of corpora (e.g. Compus by Fekete and Dufournaud,[55] CorpusSeparator by Correll et al.,[56] Text Variation Explorer by Siirtola et al.,[57] and those techniques proposed by Regan and Becker[58]), the analysis of relations or reuse (e.g. ShakerVis by Geng et al.[59] and techniques proposed by Jänicke et al.[60]), and lexical analysis (e.g. the study by Rohrdantz et al.[61]). An additional category of tasks that is worthy of mention is related to semantics: while numerous text visualization techniques use topic modeling, experts in computational linguistics use visualization to facilitate their research on this subject. For instance, Kabán and Girolami[62] visualize their own model of dynamically evolving text collections. Another task related to stance analysis is discourse analysis. Existing work on visualization of discourse includes the graph-based approach by Brandes and Corman,[63] Conceptual Recurrence Plots by Angus et al.,[64] and several recent works that focus on discourse in online social media: Lingoscope by Diakopoulos et al.[65] or ConVis by Hoque and Carenini.[66]

VA for sentiment research

Finally, the work that is most relevant to our approach in this article is dedicated to sentiment visualization which facilitates the research on sentiment for linguists. Gregory et al.[67] conduct visual sentiment analysis of document collection with regard to *affect bearing* words. Their approach involves eight affect categories (positive, negative, virtue, vice, pleasure, pain, power cooperative, and power conflict) and uses IN-SPIRE for visualization purposes. The recent work of Makki et al.[68] focuses on sentiment lexicon refinement from reviews dataset which involves user input via interactive visualization. Their sentiment analysis is based on

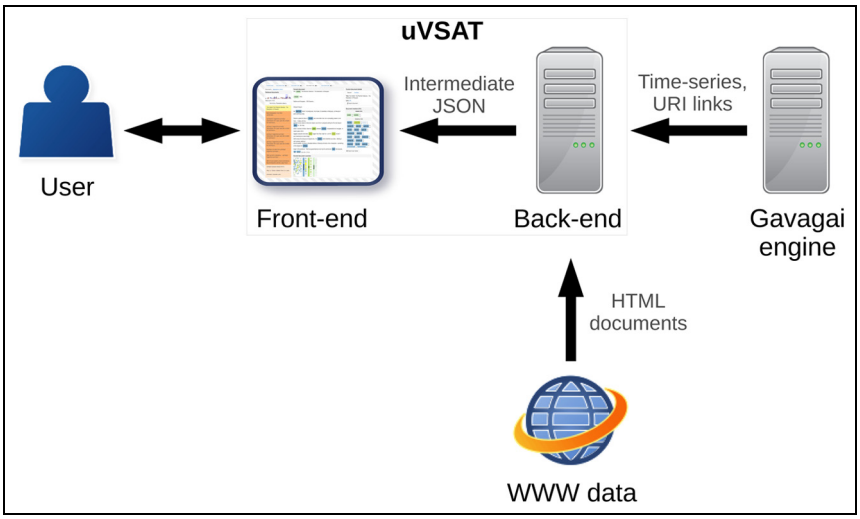

Figure 2. The architecture of uVSAT comprises front-end and back-end tiers that communicate with external servers.

a standard positive–negative dichotomy. The two major differences between these works and our proposed approach in uVSAT are the involvement of online social media text data (which is dynamic with regard to analysis sessions and available for temporal analysis) and the choice of sentiment categories (which is a base for the further analysis of stance).

To the best of our knowledge, the problem of stance analysis and visualization has not been addressed by work in VA or information visualization. Therefore, we would like to raise the awareness of the InfoVis and VA communities in this article by building on the discussed work in text visualization for sentiment analysis and existing work on visual text analytics for linguists.

Overall architecture and data

Before we can discuss the overall architecture of our VA approach, we have to briefly present the different members of the StaViCTA project in order to motivate our designs. The visualization group at the Department of Computer Science, Linnaeus University, is responsible for VA research and the development of the VA approaches needed in the project and presented in this work. A domain expert group in linguistics at the Centre for Languages and Literature, Lund University, is in charge of task identification, stance theory construction, evaluation, and so on. Finally, a group at the company Gavagai has broad knowledge in NLP and develops automatic analysis techniques and tools for the project. Gavagai monitors and processes online media (e.g. newswire, weblogs, forums, and social media such as Twitter and Facebook) for media monitoring and text analytics purposes.

System architecture and workflow

Figure 2 displays the overall architecture of uVSAT that is implemented as a web application. The back-end consists of a (visualization) server application implemented in Java that communicates with the Gavagai computing server, fetches the HTML content from URI links, processes the text data, and communicates the results in JSON format to the client(s). The front-end is implemented in JavaScript with D3[69] and Rickshaw[70] libraries, and it only requires a modern web browser. While the major and cost-intensive computational analyses are processed by the Gavagai and visualization servers, several minor analyses (which do not require intense computations for large amounts of data) are implemented on the client side.

Data model

uVSAT has been designed to use time-series data from external providers through a RESTful API,[71] as well as to fetch and process corresponding HTML data from respective web servers. Currently, we use time-series data only from our collaboration partners at Gavagai (although we plan to support other data sources in the future). Gavagai analyzes text data from multiple sources, but for the purposes of the system presented in this article, they use the data fetched from various blogs and forums.

As mentioned in the "Background" section, we focus on the simplest possible type of stance analysis, that is, counting the occurrences of sentiment terms in documents that mention specific target terms. This simple approach allows our partners to support the analysis of large amounts of text data, up to 15 million

documents per day. Here, a *target* can be anything of interest: a person, a brand, a company, a location, an event, or even something abstract such as a concept or an idea—as long as it can be defined by a set of keywords (also denoted by *target terms* in the context of our tool). Our present set of targets T includes the following

$$T = \{diet, weapons, Hobbit, Coca{-}Cola, Pepsi\}$$

To detect documents associated with stance, we consider specific markers relevant to sentiment and (un)certainty from several available sources (WordNet-Affect,[72] GeneralInquirer,[73] and Compass DeRose[74]), while refining those marker lists is one of the purposes of uVSAT (since the sources above do not differentiate stance from sentiment, etc.). Our choice of analyzed stance types (also denoted by *observers* in the context of our tool) includes the Big Six emotions (see the "Background" section) as well as two other categories

$$O = \{anger, joy/happiness, fear, sadness, disgust, surprise, certainty, uncertainty\}$$

As an example, *weapons* is a monitored target which is defined by a list of 3771 keywords, harvested from the Wikipedia lists of weapons.[75] Whenever one of these keywords is mentioned in open online media, the entire utterance containing the keyword is analyzed for occurrences of stance markers. Here, *utterance* is simply defined as a sequence of text defined by delimiter symbols, for instance, the text fragment

> I am so sick of people who sell such rifles and so sick of people who buy this distasteful weapon.

contains two occurrences of the stance marker "sick of" and one occurrence of "distasteful," generating a *polarization value* of 3 for the target *weapons* for observer *disgust*.

To summarize the description of n targets, m observers, and their possible combinations, we can describe the hierarchical structure of the data as $\{(T_i, \{O_{i1}, \ldots, O_{ij}\}) | 1 \leqslant i \leqslant n, \ 1 \leqslant j \leqslant m\}$ for targets $T_i \in T$ and the corresponding observers $O_{ik} \in O$, for instance, (*Hobbit*, {*disgust*, *anger*, …}).

The occurrence counts are aggregated for each target–observer combination (T_i, O_{ik})—for example, *Hobbit*/*disgust* or *Hobbit*/*anger* (note that we equivalently use the notations (T_i, O_{ik}) and T_i/O_{ik})—over a specific time frame which is presently set to 1 h. Thus, all occurrence counts for a specific stance type within this time frame $[t_1; t_2]$ are summed, resulting in an hourly value v for each combination. These values are then retrieved and visualized by uVSAT as time-series. Because of this aggregation step (which is necessary to reduce the complexity and computational demands), the time-series data describe the general tendencies with regard to stance but do not directly provide any details about the distribution of specific markers. Therefore, further exploration of the original text documents is required from the users.

The Gavagai API also provides URIs to the documents used to calculate the polarization values (taking (T_i, O_{ik}, t_1, t_2) as arguments and returning sets of URIs), although the corresponding HTML content has to be downloaded and processed on our side. Unfortunately, the total amount of available data makes it infeasible for the VA tool to prefetch everything. Therefore, we limit ourselves to queries for specified sets of target–observer combinations across interactively selected time intervals (although we plan to support streaming data in the future).

Requirement analysis

After the introduction of the fundamentals and research gaps of visual stance analytics including a short discussion of the origin and structure of available datasets, we are able to take a closer look at the actual analysis challenges and most important tasks that uVSAT should address. They are based on extensive discussions with our collaboration partners in linguistics and computer linguistics.

Analysis challenges

We have designed uVSAT to facilitate users with answering the following questions:

Q1. How do the calculated values for targets or observers change over time? What are the overall temporal trends?
Q2. How to identify "interesting regions" in multiple time-series which span over long intervals of time? How to reduce the visual complexity with regard to noisy data?
Q3. What are the original documents associated with the values for targets or observers? How to identify the most interesting documents with regard to stance analysis?
Q4. How are markers distributed in a particular document?
Q5. How are specific markers distributed in the retrieved sets of documents? How to identify the documents with a large number of markers or the documents which contain a lot of unique marker types?
Q6. How to handle a long analysis session involving multiple time intervals and document sets? How to recover a previously discarded document set? How to

navigate quickly to a previously analyzed document set?
Q7. Are there any relationships between analyzed document sets?
Q8. How to use particular marker, document or document set analysis results for further investigation?

Analytical tasks

These questions and problems can be mapped to the following categories of high-level (analytical) tasks:

T1. Time-series analysis: compare the values for various targets and observers (Q1, Q2), explore trends (Q1, Q2), and identify interesting regions for further investigation (Q2);
T2. Document sets navigation: query for the documents associated with selected observers or time intervals (Q3), keep track of related queries (Q7), and navigate the queries history (Q6);
T3. Document sets analysis: explore the retrieved document sets (Q3) and reveal the general trends by using data aggregation (Q5);
T4. Document navigation: query for specific documents either explicitly (Q6) or while navigating enclosing document sets (Q3) and aggregated data (Q5);
T5. Document analysis: explore the text content and stance marker distribution in a selected document (Q4) and export the static content for manual investigation (Q8);
T6. Stance marker collection: export the selected utterances (or parts of them) as new markers (Q8).

In the following section, we discuss our visualization approach in detail, justify the design decisions, and refer back to the above-listed research questions and tasks.

Visualization approach

The graphical user interface (GUI) of our tool offers a tab-oriented design with two types of tabs (cf. Figures 3 and 4): a single timeline view tab that is used to work with an arbitrary number of timeline plots, and multiple document view tabs that are opened by the user when fetching the document URIs for selected time intervals. As the timeline view is the entry point of all visual analyses supported by our approach, we start our discussion with this view.

Timeline view

The timeline view tab (cf. Figure 3) provides the users with the interfaces for exploring time-series data for selected targets or observers and specified time intervals. Note that fetching the input data to be analyzed—that is, the initial selection of specific targets, observers, and time ranges—from the Gavagai server is done via a simple dialog box as explained in our use case (cf. the corresponding section). In this section, we concentrate on overall design aspects including visual representation and interaction possibilities.

Color coding considerations. Before we address the particular representations, we have to explain the color coding scheme used for the timeline view as well as document views. As mentioned in subsection "Data model," the analyses supported by our tool involve the combinations of targets T_i and specific observers O_{ik}. So, the resulting hierarchical data structure for one specific target might be (*diet*, {*anger*, *joy*, …}), for instance. The time-series data fetched from our partners are organized this way with the focus on target–observer combinations, and our initial choice of the color coding was based on the decision to provide a unique color for each combination. However, this approach had two issues: first, the sheer number of combinations (45 entries in our present set of target–observer combinations) made it difficult to use a color scheme that would facilitate the users' perception of the data and, second, this color scheme was not related to the scheme for document views (described below), so the users could easily lose the mental map when switching between the view tabs.

The analyses employed by document views (see the corresponding subsection below) concentrate on the observers, that is, stance types, and do not differentiate between observers related to various targets. This had an implication that the color coding for document views was initially based on ColorBrewer,[76] and it contained separate colors for observers and targets.

Afterward, we have changed the color coding used for the timeline view in accordance to the TreeColors approach.[77] To generate the colors, we have *inverted* our hierarchy to the form $\{(O_j, \{T_{j1}, \ldots, T_{ji}\}) | 1 \leqslant j \leqslant m, 1 \leqslant i \leqslant n\}$, for instance, (*joy*, {*diet*, *Hobbit*, …}), and then used the TreeColors package. The resulting color coding aims to assign different observers distinct color hues although it is not perfect since there are still too many of those. The colors assigned to target–observer combinations pertaining to the same observer have rather similar hues. This, on one hand, makes it simple to spot such similar combinations. On the other hand, although, it makes it difficult to discern such plots—this is partially alleviated by interaction techniques such as details on hover and filtering. Overall, the main benefit of this approach is that it allows of using the same color hues for

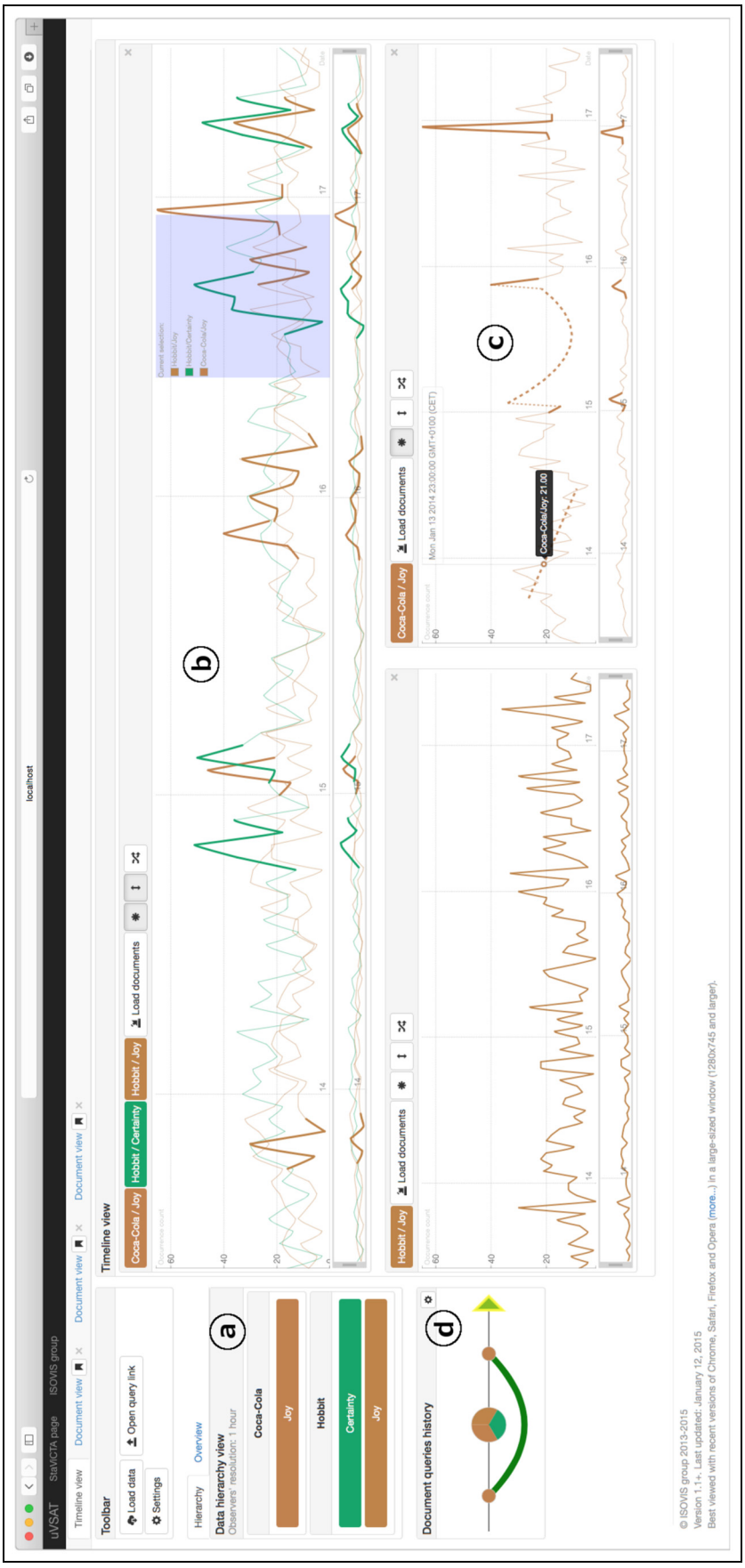

Figure 3. The screenshot of our tool shows the *timeline view*. Users start by loading time-series data associated with the two targets *Hobbit* and *Coca-Cola* and two observers *certainty* and *joy*. Then, they can explore the data and select time intervals for specific target–observer combinations (see the blue-shaded area) in order to further analyze the corresponding documents.
The screenshot demonstrates (a) the data hierarchy view, (b) timeline plots, (c) trend lines, and (d) the history diagram.

Figure 4. The screenshot displays our *document view* for a selected time interval (cf. Figure 3) and the target–observer combinations *Coca-Cola/joy*, *Hobbit/joy*, and *Hobbit/certainty*. Here, users can explore the resulting set of documents (on the left) and are able to analyze the contents of specific documents (middle) with regard to the occurrence of stance markers summarized on the right-hand side.
The screenshot demonstrates (a) the document links list, (b) the current document view, (c) the document marker view, and (d) the current document overview.

observers across the timeline and document views that helps to preserve the users' mental map.

Data hierarchy view. After the input data have been loaded, the users are provided with the data hierarchy view displayed in Figure 3(a) that shows the hierarchical structure of the available target–observer combinations. Users can also open a tab with iconic "overview plots" (cf. Figure 10) for *all* fetched time-series which are similar to regular timeline plots with highlighted *regions of interest* (*ROIs*) (see below). These overview plots support a simple way to compare the time-series and to find more general patterns in the data (research questions Q1 and Q2). As soon as interesting target–observer combinations are found, the user may want to investigate these data in detail and drag-and-drop the entries from the data hierarchy view onto the main part of the tab. Then, uVSAT displays the timeline plots for the chosen combinations. For instance, in Figure 3, a user has selected three views where several target–observer combinations are visualized.

Timeline plots. uVSAT uses a standard line plot representation for time-series data (cf. Figure 3(b)) and supports usual interaction techniques for such plots (research question Q1). We have chosen this visual representation as our domain experts are already familiar with it. In addition, line plots can be easily extended with additional graphical features. Details on hover, plot overview, and scroll and zoom are provided by default by the Rickshaw component. Users are also able to filter the plots with regard to visible target–observer combinations by switching on and off the corresponding labels. Our tool supports multiple plots displayed on the same canvas (users can drag-and-drop additional items from the data hierarchy view) or separately (users can drag the plot containers to change the timeline view layout). For the comparison of several plots displayed side by side, users can control the automatic vertical scaling—by default, plots are scaled to fit the containers. This functionality was explicitly wished by our domain experts.

ROI highlighting. To facilitate the search for ROIs, our tool also supports automatic ROI highlighting (research question Q2). Currently, we use a basic ad hoc algorithm for marking the ROIs based on outlier or differential analysis. As a first step of the algorithm, time-series points x_i are marked, which differ significantly (with regard to threshold parameters θ_1 and θ_2) either from the mean value μ_x (standard deviation σ_x is used for comparison) or from the preceding point (judging by the first derivative x_i')

$$A = \left\{ x_i : |x_i - \mu_x| > \theta_1 \sigma_x \vee |x_i'| > \theta_2 \max_j \left(|x_j'| \right) \right\}$$

Since the source time-series data are in general noisy, A will result in multiple regions of small size (comprising only one or several points). Therefore, in the second step, we smooth the results by marking neighboring points as parts of ROI, which will result in contiguous regions

$$ROI = A \cup \{ x_i : (x_{i-1} \in A) \vee (x_{i+1} \in A) \}$$

ROIs are highlighted by thick line segments (cf. Figure 3(b)). The algorithm parameters θ_1 and θ_2 can be adjusted by the user, which can be used to partially alleviate the problem of noisy data or to increase or reduce the number of highlighted regions to focus on.

Trend analysis. Users have several options of conducting trend analyses over selected time intervals for specified observers (cf. Figure 3(c)). uVSAT supports linear and quadratic time-series trend analysis based on polynomial regression (calculated with the ordinary least squares (OLS) method). We implemented two variations: one can choose to either render trends as overlay plots (cf. Figure 5(a)) or to substitute selected timeline plot segments with trend lines (cf. Figure 5(b)) to reduce the visual complexity of the displayed data (research questions Q1 and Q2). Trend lines are easily distinguishable by the use of dashed lines. Even information about the predicted value change at the current trend rate and a button for removing trend lines are available on hover.

Document URI links queries. As soon as the user is more interested in the concrete documents whose frequencies are represented by the different time plots, he or she can select time intervals for specific sets of observers and load the corresponding URI links to the documents (research question Q3). In this case, a new

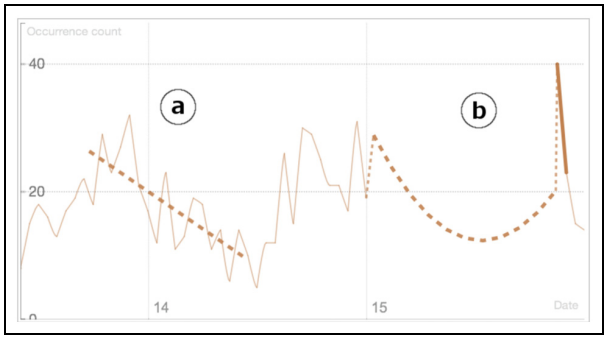

Figure 5. Trends can be displayed as either (a) overlay plots or (b) instead of original plot segments.

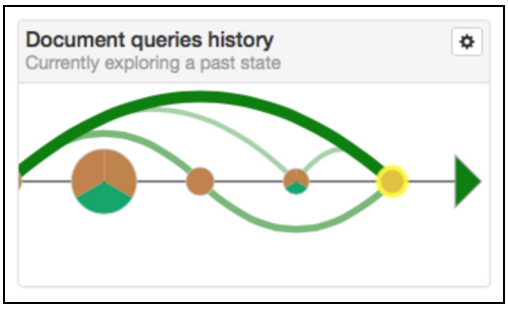

Figure 6. The history diagram allows users to keep track of document queries and navigate between interface states.

document view tab is created and a thumbnail of the line plot used for the query is displayed in this new view in order to preserve the mental map. An example of this thumbnail can be seen in Figure 4 in the left upper corner.

History diagram

Since the workflow of uVSAT involves multiple document view tabs that also may be closed by a user during the analysis process, the need for overview and control of such user actions arises. Our interactive history diagram (cf. Figure 3(d) and Figure 6) provides an overview of the document URI queries sequence, their results, and relations to each other (research questions Q6 and Q7).

In this diagram that supports the so-called analysis provenance,[78] nodes represent URI queries and edges represent the detected relations between corresponding query results (this partially resembles the visualization approach described by Cernea et al.[79]). The size of every node is proportional to the number of URI links retrieved for the corresponding query. Nodes are represented by glyphs similar to pie charts (although only qualitative information about relevant observers is used), following the same color coding of observers as the timeline plots. The currently selected node is highlighted in yellow. Since the diagram is used for history navigation, it also contains a dedicated node (depicted by a triangle) that represents the up-to-date interface state. Edges connect only nodes whose query results contain common subsets of URI links. The size of common subsets (i.e. Jaccard similarity of link sets[80]) is mapped to edge opacity, thickness, or both of these attributes (selected as a user setting). The layout of the history diagram is based on arc diagrams by Wattenberg:[81] nodes are simply aligned along a horizontal axis in the order of corresponding queries, and edges are rendered as curved arcs. We apply a random-order greedy heuristic described by He et al.[82] to decrease the number of edge crossings when allocating edges to the upper or lower part of the drawing.

The interactive history covers the following functionalities: every time a user issues a URI links query that leads to the creation of a new document view tab, the state of this new tab and the timeline view tab are saved and a corresponding node is added to the history diagram. When the user clicks on a history node, the timeline view tab state is restored, a document view tab with corresponding state is either created or brought into focus (if currently present), and the user actions temporarily stop affecting the history state (e.g. issuing a new query will not add the resulting state to history)—we have chosen such behavior to keep the history sequential. When the user clicks on the triangle, the previously saved up-to-date state is restored. Under circumstances, this can lead to some document view tabs getting closed.

Document views

A document view tab (cf. Figure 4) basically consists of two areas. The left (smaller) area provides information about all documents fetched based on the selection described at the end of subsection "Timeline view." Thus, it shows the aforementioned line plot thumbnail used for the query as well as a link list (cf. Figure 4(a)) to HTML documents (blog posts, forum messages, etc.) that were marked as associated with a specific target–observer combination. Users can filter the list by URI domain and sort it by the timestamp value or by polarization value (as reported by the Gavagai server). Polarization values are also used for the color coding of list entries (research question Q3).

By selecting a link from the list, the corresponding document content is fetched, processed at the (visualization) server side, and rendered at the client side. If the content is not available at this time, the corresponding list entry is marked. The document data at this stage are raw HTML which affects the analysis. This is because the source code comments and metadata (such as keywords) often contain text irrelevant to the document content. To direct the user's focus on textual document data, uVSAT renders the HTML content as plain text by using the Jericho library.[83] All data and analysis results related to the single focus document are shown in the second area on the right-hand side of the document list. This area integrates four subviews: the current document view, the current document details view (not further discussed here), the document marker view, and the current document overview.

It should be noted that uVSAT also provides an opportunity to copy the query link for a given

document view tab and to use it in later analysis sessions by opening a tab with identical contents (research question Q6).

Current document view. Figure 4(b) displays the text representation of a document. The stance markers and target terms are highlighted and support brushing in coordination with the other views (research question Q4). The motivation for the color coding for document view tabs was described above: it uses a scheme with eight colors for stance markers and a separate scheme with five colors based on ColorBrewer for target terms since targets share stance markers associated with observers (types of stance), for example, the word "commendable" is a marker of *joy* for both *Hobbit* and *Coca-Cola*. To distinguish target terms from stance markers, the former are marked by a striped background pattern.

Document marker view. Information about stance markers (and their occurrence counts) as well as target terms detected in the current document is summarized in the document marker view (cf. Figure 4(c)). The stance markers for each observer are sorted by their counts to facilitate user investigations (note that target terms occurrences do not affect the statistics since such terms are not directly related to expressions of stance). The users can navigate the document with regard to markers or terms occurrences and to filter them (research question Q4).

Current document overview. To give users an overview of marker or term distributions in the current document (and an additional means of navigation), uVSAT provides several visual representations displayed in Figure 4(d). First of all, a two-dimensional (2D) overview is visualized by mapping the current positions of all markers or terms onto a canvas (they are represented by circles and diamonds, respectively). The current viewport is displayed as a rectangle. This overview supports navigation by clicking on a plot item or the canvas. Additionally, a separate one-dimensional (1D) overview for each observer and target is visualized by projecting the positions of corresponding markers or terms onto a vertical axis. Such overviews help the users to immediately perceive the distributions over the document length since the 2D overview can become cluttered in case of numerous markers or terms. 1D overviews support document navigation by clicking on plot items. Seeing such distributions is especially interesting for our domain experts because it is important for a better understanding of stance in discourse (research question Q4), for instance, if a marker for a specific stance type mostly occurs in the context of another marker.

Aggregation charts

While the techniques discussed above allow the users to analyze a selected document in detail and provide an indication of interesting documents (by polarization values), the document sets retrieved for certain queries may contain thousands of documents, and the users will benefit from a method that helps them to select documents that are interesting for further stance marker investigation (research question Q5). uVSAT addresses this problem with a technique that we call *aggregation charts*: it provides an informative overview and means of navigation for the current document set with regard to detected markers and observers (cf. Figures 7 and 8).

The visual representation is based on basic bubble charts described by Viégas et al.[84] Every item in the chart represents a single document which corresponds to the target; the color coding is based on the nominal target values. A single item is visually represented by a glyph consisting of two nested circles. The size of the outer circle is proportional to the total number of corresponding stance markers detected in the document, and the size of the inner circle (filled with a more saturated color) is proportional to the number of unique marker types detected in the document. For instance, a document with 100 occurrences of a marker "good" and 100 occurrences of a marker "bad" has only two unique marker types: "good" and "bad."

The aggregated data used for these charts can be organized in two ways: by observer and by stance marker. In the former case, a separate chart is visualized for each observer associated with the document set. In the latter case, one individual chart is visualized for each unique marker type (belonging to present observers) that has been detected in at least one document.

Figures 7 and 8 display examples of aggregation charts visualized for a document set based on 1517 URIs retrieved for the target–observer combinations *Coca-Cola*/*joy*, *Hobbit*/*joy*, and *Hobbit*/*certainty*. In Figure 7, the charts are organized by observer: the left chart contains items pertaining to both *Coca-Cola* and *Hobbit*; however, the right one does not contain items for *Coca-Cola* since no corresponding target–observer combination was available. This figure also shows the details for a chart item displayed on hover. An example of aggregation charts organized by stance markers is displayed in Figure 8. There are multiple charts sorted by the corresponding document numbers in decreasing order, and the user can browse these charts with a specific marker in mind. Details for the first chart (marker: "good") are provided in a tooltip. Here,

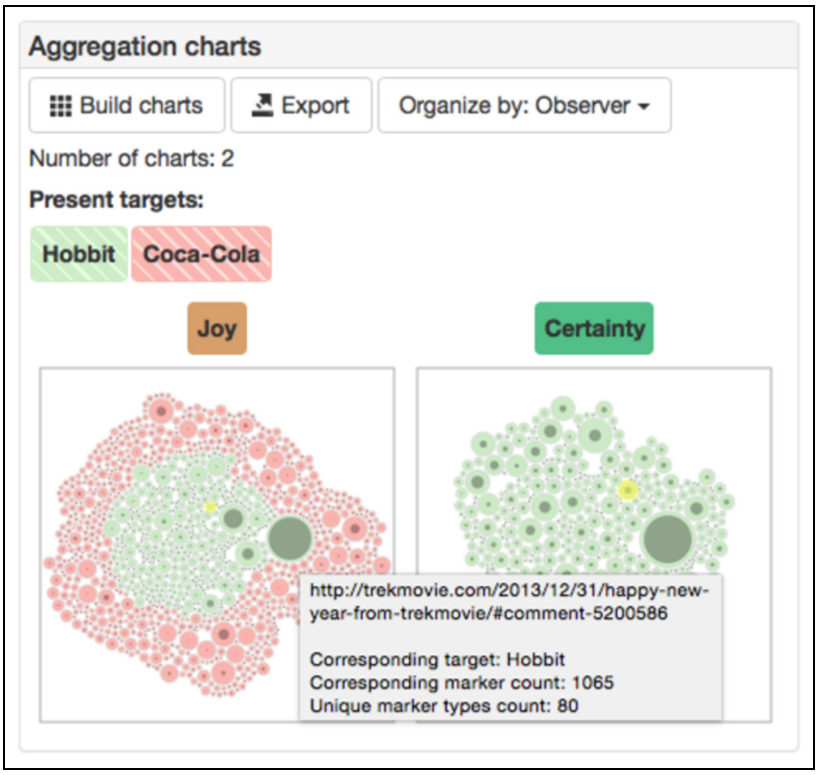

Figure 7. Aggregation charts *organized by observer* allow users to explore the distribution of documents with respect to the corresponding observer.

Figure 8. Aggregation charts *organized by marker* allow users to reverse the flow of analysis: they can concentrate on document distributions with regard to a specific interesting stance marker.

the currently selected document is highlighted (yellow) in all charts.

Aggregation charts facilitate the quick perception of the distribution of observers or stance markers in all documents, the identification of documents with a large number of stance markers or unique marker types, the navigation to such documents, and the analysis of document properties concerning other observers or stance markers (by brushing the corresponding chart item).

Marker and document export

One aim of our visualization tool is to identify and collect relevant stance markers from a larger number of analyzed documents (research question Q8). uVSAT supports the export of new stance markers from document view tabs by selecting a portion of text in the current document view (depicted in Figure 4(c)), assigning it with arbitrary tags, and exporting it to a JSON file. This approach allows us to collect a dataset of stance markers not restricted by the categories currently used for observers. Moreover, we are able not only to collect stance markers as short phrases (1-grams,[85] 2-grams, or similar) but also to collect larger utterances which provide context for stance analysis.

Our tool also supports the export of currently viewed documents and aggregation charts as static HTML pages. In the former case, the document view with highlighted stance markers and target terms, document details, hierarchical markers view, and document overview (essentially, all the data pertaining to the current document on a document view tab) are exported. In the latter case, all aggregation charts that are currently available are exported together with the corresponding document set query (used observers, selected time interval, etc.). This feature allows users to store static data for further manual investigation or referencing, which can be especially helpful for researchers in linguistics.

Use case: linguistics research

The use case described here is one in which a linguist has chosen to analyze negative sentiments of stance (focusing on *anger*) in blogs, within a limited 1-week time frame. This example illustrates how researchers in linguistics benefit from our tool when conducting stance analysis. The event chosen was the highly controversial Coca-Cola commercial presented during Super Bowl XLVIII[86] (3 February 2014 CET). The aims of the analysis are the following:

A1. Analyze the overall usage of stance-related sentiments for the scandal time span;
A2. Identify the document with the largest number of markers of *anger*;
A3. Identify the most frequently used *anger* markers;
A4. Analyze how such markers are used in the previously identified document;
A5. Finalize the choice of the detected document for further linguistic research.

For performing an accurate analysis, data revealing information about the communicative forces and the attitudes to the ideas discussed at different points in time as well as possible relationships between those attitudes must be made available to the researcher. By using uVSAT, the linguist is able to analyze these aspects of the social media data which would be impossible for manual stance analysis.

Timeline data analysis

First, the researcher uses the *Load data* dialog box and selects all *Coca-Cola* observers for the time interval 30 January 2014 12:00—6 February 2014 12:00 CET in order to obtain a very broad return of data (cf. Figure 9). The time-series calculated for corresponding observers are loaded from Gavagai API.

By viewing the *hierarchy* and *overview tabs* (cf. Figure 10), the researcher verifies that all of the chosen observers have been loaded and confirms that there are sufficient data to be analyzed.

The researcher immediately notices the spike of activity on multiple plots around early hours of 3 February CET, which corresponds to the late evening of 2 February EST—the time when the advertisement was aired in the United States (aim A1).

Then, the researcher creates timeline plots by dragging-and-dropping the observer items onto the *timeline view*. Using the slider control, the researcher concentrates on the time span 3 February 2014 01:00–3 February 2014 19:00 CET. To confirm a conjecture that some of the observers have extremely low counts in the current time span (aim A1), the researcher filters them out. The remaining observers are *certainty*, *joy*, *uncertainty*, and *anger* (see Figure 11). To start analyzing the textual data, the researcher issues a request for corresponding URIs.

Identifying the document of interest

The resulting URI set comprises 3424 document links. While the researcher could explore this dataset manually, it would take a significant amount of time to achieve aim A2. At this point, the researcher decides to build the aggregation charts for the current document set and to investigate the charts organized by observer. For this, the text document data are

Figure 9. The dialog box used to select the time intervals and target–observer combinations to load time-series data. Note that there are additional observer types (*frequency*, *positivity*, and *negativity*) provided by Gavagai by default that are not associated with concrete stance markers (therefore, they are beyond the focus of our research).

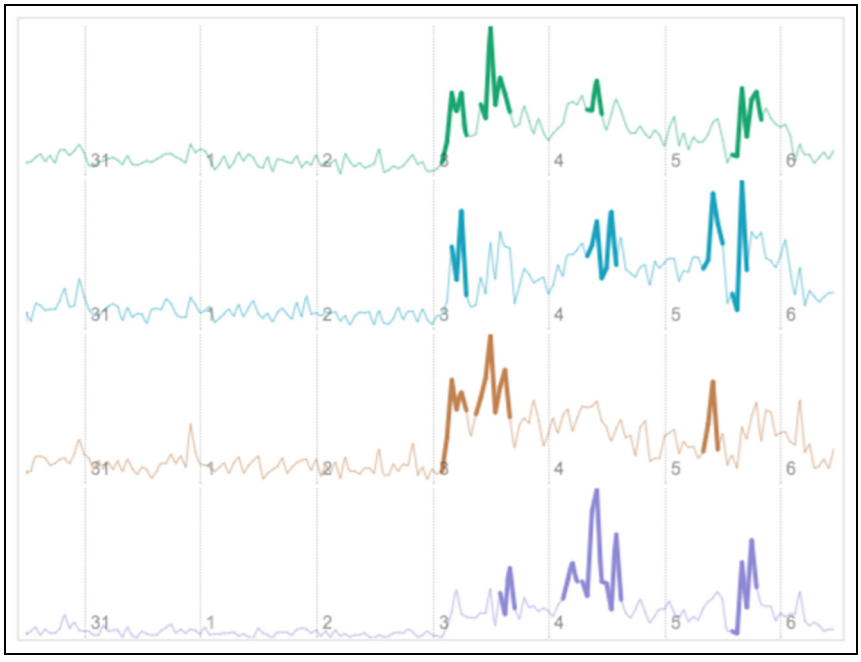

Figure 10. Part of the timeline overview: the plots for observers are ordered by mean value in descending order, *certainty* being the first. Note the spike around 3 February, when the scandal occurred.

fetched from respective web servers and processed by uVSAT.

The aggregation chart for *anger* (cf. Figure 12) comprises 1948 documents which in total contain 154 unique markers of *anger*. The researcher immediately identifies two candidate documents with the largest number of corresponding markers which are represented by glyphs with the largest diameters (also, with large shaded areas which means large number of unique marker types). By hovering on these glyphs, the researcher finds out that one of them contains 142 occurrences of *anger* markers (39 unique types) and another one contains 193 occurrences (41 unique types). The researcher selects the latter glyph by clicking and loads the corresponding document.

The loaded document of interest (depicted in Figure 13) is a blog post[87] with a heated discussion in commentaries. To concentrate on the analysis of *anger* markers, the researcher filters out all markers of other observers. The current document overview plots at the bottom of the screenshot clearly show that the markers of *anger*, as well as the target terms of *Coca-Cola*, are

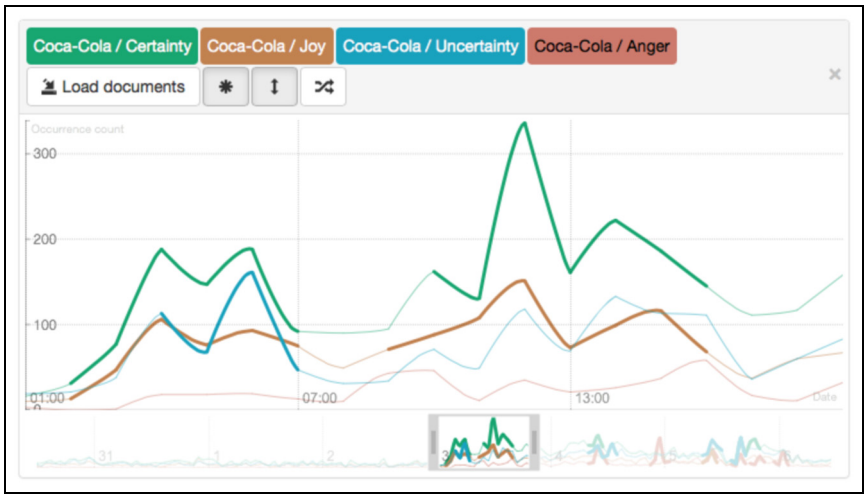

Figure 11. Timeline view: four observers for target *Coca-Cola* that are used for detailed analysis are *certainty*, *joy*, *uncertainty*, and *anger*.

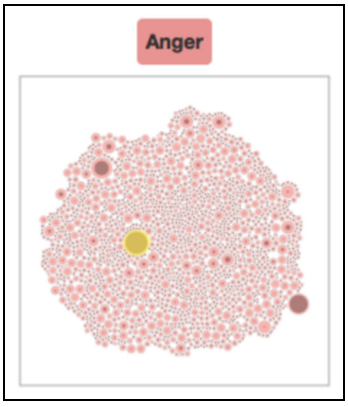

Figure 12. The aggregation chart for *anger* provides an opportunity to identify the document with the largest number of corresponding stance marker occurrences. There seem to be two candidate documents which are represented by large glyphs (also with large shaded area). By hovering on these glyphs, the one with larger count of markers (in this case, 193 occurrences) is identified and later used for detailed analysis.

evenly distributed throughout the entire document. To refine the analysis, the researcher needs to concentrate on specific markers.

Identifying the markers of anger

The aggregation charts for the current document set can be organized by stance marker instead of observer. The researcher selects this option and explores the resulting set of 605 aggregation charts (one per each unique stance marker type). Since the charts are ordered by marker occurrences number in descending order, the researcher quickly identifies several most frequent markers of *anger*, thus achieving aim A3 (see Table 1).

Table 1. Stance markers of *anger* in the documents.

Marker	Corresponding documents	Unique markers in documents
Hate	579	123
Angry	347	113
Offended	265	92
Outrage	232	91
Fit	206	107

The most frequently used stance markers of *anger* in the document set related to the use case. These data have been discovered by investigating the details when hovering over aggregation charts' labels.

Final document analysis

After identifying the most frequent markers of *anger* using the aggregation charts (here: "hate," "angry," "offended," etc.), the researcher concentrates on the previously selected document and filters out all the other markers. It turns out that some of the identified markers are also among the most frequent markers of *anger* in the document as well (cf. Table 2).

The researcher reviews the current document overview once more (cf. Figure 14) and concludes that the identified markers are also distributed throughout this document. As the observer *anger* has the marker "hate" prolifically used, the analyst investigates further, addressing the linguistic characteristics that are employed by users who have posted these. The linguist

Figure 13. Document view for a selected document with majority of stance markers filtered out. Besides the *Coca-Cola* target terms, only the instances of all markers of *anger* are displayed.

Table 2. Stance markers of *anger* in the selected document.

Marker	Occurrences in document	Rank in document
Hate	40	1
Offended	25	2
Angry	16	3
Outrage	4	8
Fit	3	9

The number of occurrences and ranks of the previously identified stance markers of *anger* in the current document.

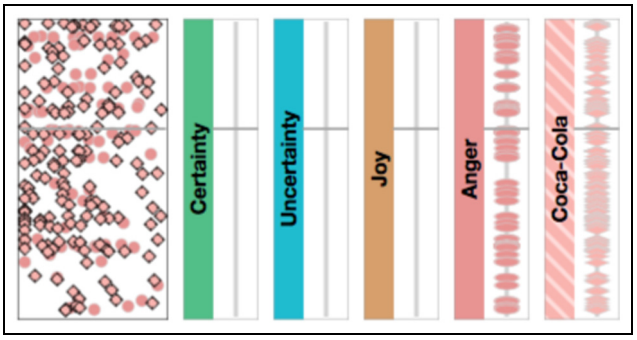

Figure 14. The overview for the previously selected document with only five marker types of *anger* displayed. Note that even after filtering the other *anger* markers (cf. Figure 13), numerous instances of these five marker types remain and they seem to be distributed throughout the whole document.

now proceeds with a close analysis of the document giving critical attention to the markers "hate," "offended," and "angry," thus achieving aim A4. The researcher's conclusion is that the identified document is interesting for further manual linguistic analysis with regard to the flow of the conversation, and so on, as well as for preparation of an ML training dataset. By exporting the document from uVSAT, the linguist achieves aim A5.

Summary

By using uVSAT, the researcher has been able to achieve his or her analysis aims, that is, exploring the data related to the case, analyzing the stance-related phenomena of *anger* and exporting the analyzed text data. By being able to interpret the ROIs on the timeline view, the researcher was able to limit a great amount of documents to an amount for a more detailed review. The tool's ability to visualize multiple markers simultaneously in the document overview positively guided the investigation. By viewing the aggregation charts, the researcher's decisions were visually supported, and he or she was able to draw the conclusions about stance phenomena in the dataset. The potential for employing these different refinement possibilities lets the researcher review statistical plots that are dynamic and updated as new postings are incorporated into the document view. The analysis features provided by the document view complements the manual stance analysis based on close reading. Overall, the patterns constructed by uVSAT create an ample opportunity for the researcher to employ user-based data en masse.

On a final note, the linguist began with one specific study area. After using uVSAT, the researcher concluded that the data have also revealed three other possible areas of interest: (1) directionality and frequency of the *anger* markers, that is, who the poster intends as the recipients and how often they appear and respond; (2) instances of how posters modify their use of *anger*, that is, intensifiers or attenuators; and (3) if *anger* is negated so as to create a positive meaning. The tool has provided several new potentials for future lines of research that could have gone unnoticed if traditional linguistic investigations were used.

Expert reviews and discussion

In this section, we present the results of two domain expert reviews as well as performance issues. Based on these findings, we discuss some lessons learned during the development and testing phase of uVSAT.

Domain expert reviews

For the time being, our research partners at Lund University have been the primary users of uVSAT. They are familiar with standard tools for corpus analysis (e.g. AntConc, BYU-BNC, WORDSMITH, or Google Ngram Viewer) and manual text analysis. As a kind of project preparation, we introduced basic visualization concepts and techniques to them at the beginning of our collaboration. Their suggestions and feedback during the design and development stage of uVSAT are summarized in the following with regard to general analysis workflow, visualization and interaction techniques, and possible improvements for the tool.

General analysis workflow. The experts have been very enthusiastic about the opportunity to analyze a large number of online social media documents in detail with regard to stance and sentiment in an interactive way. They have noted that their usual tools of

choice in most cases require text preprocessing and employ static or rarely updated corpora, as opposed to our approach:

> The uVSAT tool can accommodate the time factor and help the analyst sift through large amounts of data where important chunks could easily be overlooked. Using the uVSAT tool, which is visually driven to reveal patterns, the researcher can track these and follow how language is being shaped by current digital communications.

The experts have also appreciated the fact that uVSAT is implemented as a web application which does not require a specific OS or installation or update procedures.

Interactive visualization approach. The feedback on the design of both timeline and document views has been positive. The experts have approved of the features facilitating the time-series analysis, in particular, they have liked that ROI highlighting is turned on by default. The experts have commended the usage of color coding to highlight the ROIs as well as the markers or terms. They have also approved our decision to convert HTML documents into plain text in order to concentrate on the text content in the document view tabs. The experts have also been very positive about the aggregation charts as a means of overview, pattern detection, and navigation:

> Aggregation charts give extremely comprehensive views that are easily understood by this user. These images result in giving the researcher a direct visual confirmation of the number of markers, which then can be scrolled through, chosen and loaded.

The ability to export stance markers as well the content for further manual investigation was also commented on:

> This gives the user a pro-active involvement in the ongoing improvement of the tool that is neither confusing nor time-consuming.

Possible improvements. One of the experts' suggestions during the development was related to the comparison of several timeline plots. We have addressed it by providing an ability to control the layout of the timeline view and to disable the automatic vertical scaling which allows the user to compare the plots situated side by side. The feedback also included some complaints related to the tool performance (see below in the next subsection) as well as a wish for additional functionality related to document set overview (e.g. clustering the documents in aggregation charts by the URL domain). We have also learned that the trend analysis feature is only rarely used since it currently focuses on already-available time-series data—therefore, we are planning to extend this feature by supporting predictive trend analysis to increase its level of utility.

Summary. The experts have stated that uVSAT is a useful addition into their arsenal of stance analysis techniques. They are using it to explore and analyze the social media data and complement it with manual stance analysis as well as by processing the exported data with other software tools, for example, for concordance analysis. They have also started to collect the ML training dataset, thus achieving the general design goals. In general, the domain experts have concluded the following:

> For a linguist, uVSAT is a viable tool for working with stance analysis.

Performance and scalability

In this subsection, we discuss certain aspects that affect the user experience when trying to apply uVSAT for the analysis of rather large datasets: data transmission delays, data processing delays, and user interface responsiveness.

We currently store neither time-series data nor document text data on our visualization server. Hence, uVSAT issues request for time-series data, URIs, and HTML content from external servers on demand. This leads to delays while retrieving the source data. Additional delays occur while transmitting the data between the front-end and back-end components and, finally, while processing the data at the server side.

We address the networking delay by conducting some types of analyses (such as ROI highlighting or trend computations) on the client side. It currently seems, although, that the performance bottleneck is the step of fetching the HTML content from numerous external servers which may have varying connection speed, performance, access frequency limitations, and even availability. We plan to introduce a local database for caching the external data (as well as some processing results), although it can lead to validity concerns (see subsection "Lessons learned").

As for the UI responsiveness: D3 and Rickshaw use SVG for rendering which may require significant computational resources (and leads to UI lags). On a 2013 MacBook Pro computer with Intel Core i7 processor

(2.3 GHz), sensible UI delays start to occur when re-rendering plots with a total of about 3000 points. This is partially addressed with a style of workflow involving preliminary analysis of time-series overview and focusing on selected time intervals.

Lessons learned

Our current visualization approach involves multiple coordinated views based on standard representations. Its main advantage (as opposed to a more complex integrated view) is the ease of user adoption: the primary users of our tool are researchers in linguistics who do not tolerate abundant details or unintuitive visual representations. The corresponding disadvantage, however, is the necessity of large display area to lay out all the views in sufficient size. We plan to address this issue in the future by developing novel visual representations for stance-related and time-dependent text data, having the domain particularities in mind.

The fact that our source data originate in online social media also has certain consequences: the text documents may be edited or deleted at any time. This presents us with a trade-off between data validity and performance. By fetching online data on user's demand (as uVSAT currently does), every document is analyzed in its up-to-date state (or it is marked as unavailable), but it requires computational resources (and it is also related to inevitable networking delays). Otherwise, if the data are cached while the original data are modified, it would invalidate the detailed analysis of document contents. To address this issue, we plan to involve uncertainty tackling techniques. Another possibility would involve storing the versioned source documents—while in practice, it would require significant resources, in theory, it could provide an analysis opportunity with regard to additional temporal dimension.

Conclusion and future work

In this article, we have introduced the problem of stance analysis of online social media texts that requires a joint multidisciplinary effort of researchers in linguistics, NLP, and VA. We have described an analysis approach for stance analysis based on sentiment or certainty considerations and presented our tool uVSAT for visual stance analysis that supports the interactive exploration of time-series data associated with online social media documents, including the text content of such documents. While uVSAT does not provide completely automatic stance analysis, it facilitates the linguists by complementing manual stance analysis of text documents based on close reading with a VA approach that allows the researchers to use massive datasets originating from social media.

The contributions of this article include the description of a VA tool that contains multiple approaches for analyzing temporal and textual data as well as exporting stance markers in order to prepare a stance-oriented training dataset. We also presented special visualization techniques developed for our tool: the history diagram (for document set query analysis provenance) and the aggregation charts (for document set overview, navigation, and comparison).

We already used uVSAT for the purposes of the StaViCTA project, and we provided feedback from the linguistics experts in this article. Using uVSAT, our researchers in linguistics have been able to collect stance markers that are now being used to define stance categories other than sentiment and certainty or uncertainty (e.g. concessions and judgment). The tool is currently being used for collecting documents that form the training dataset for our researchers in NLP as well as for actual stance analysis conducted by the linguists. We are convinced that our tool will be useful for other interested researchers.

Future work includes additional overview and navigation techniques for document sets, support for local database caching, streaming data, uncertainty tackling (with regard to missing time-series data as well as unavailable web documents), and arbitrary time-series data sources. In order to provide our tool to others, we will develop our own (more lightweight) analysis engine to become independent from Gavagai. We also plan to conduct a larger study to evaluate the effectiveness of single techniques such as history diagram and aggregation charts.

Acknowledgements

The authors would like to thank the anonymous reviewers for their valuable feedback.

Funding

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).

References

1. Kucher K, Kerren A, Paradis C, et al. Visual analysis of stance markers in online social media. In: *Poster abstracts of IEEE visual analytics science and technology (VAST '14)*, Paris, 9–14 November 2014.
2. Grice HP. Meaning. *Philos Rev* 1957; 66(3): 377–388.
3. Jaffe A. *Stance: sociolinguistic perspectives*. Oxford: Oxford University Press, 2009.

4. Warglien M and Gärdenfors P. Semantics, conceptual spaces, and the meeting of minds. *Synthese* 2013; 190(12): 2165–2193.
5. Paradis C. Meanings of words: theory and application. In: Hass U and Storjohann P (eds) *Handbuch Wort und Wortschatz* (Handbücher Sprachwissen-HSW, Band 3). Berlin: Mouton de Gruyter, in press.
6. Schamp-Bjerede T, Paradis C, Kerren A, et al. Hedges and tweets: certainty and uncertainty in epistemic markers in microblog feeds. In: *Conference presentation: 47th annual meeting of the Societas Linguistica Europaea (SLE '14)*, Poznań, 11–14 September 2014.
7. Du Bois JW. The stance triangle. In Englebretson R (ed.) *Stancetaking in discourse: subjectivity, evaluation, interaction*. Amsterdam and Philadelphia, PA: John Benjamins, 2007, pp. 139–182.
8. Thompson G and Hunston S. Evaluation: an introduction. In Hunston S and Thompson G (eds) *Evaluation in text: authorial stance and the construction of discourse*. Oxford: Oxford University Press, 2000, pp. 1–27.
9. Hunston S. Flavours of corpus linguistics. *Paper given at Charles University, Prague 2012 and at Corpus Linguistics* 2011, Birmingham, 2014.
10. Mohammad S. Portable features for classifying emotional text. In: *Proceedings of the 2012 conference of the North American chapter of the association for computational linguistics: human language technologies (HLT '12)*, Montreal, Canada, 3–8 June 2012., pp. 587–591. Stroudsburg, PA: Association for Computational Linguistics.
11. Martin JR and White PR. *The language of evaluation*. Basingstoke: Palgrave Macmillan, 2003.
12. Englebretson R. Stancetaking in discourse: an introduction. In Englebretson R (ed.) *Stancetaking in discourse: subjectivity, evaluation, interaction*. Amsterdam and Philadelphia, PA: John Benjamins, 2007, pp. 1–25.
13. Biber D. Stance in spoken and written university registers. *J Engl Acad Purp* 2006; 5(2): 97–116.
14. Biber D and Finegan E. Styles of stance in English: lexical and grammatical marking of evidentiality and affect. *Text: Interdiscip J Stud Discourse* 1989; 9(1): 93–124.
15. Pang B and Lee L. Opinion mining and sentiment analysis. *Found Trends Inf Retr* 2008; 2(1–2): 1–135.
16. Liu B. Sentiment analysis and subjectivity. In Indurkhya N and Damerau FJ (eds) *Handbook of natural language processing*. London: Chapman & Hall, 2010, pp. 627–666.
17. Lin C, He Y and Everson R. Sentence subjectivity detection with weakly-supervised learning. In: *Proceedings of fifth international joint conference on natural language processing (IJCNLP '11)*, Chiang Mai, Thailand, 8–13 November 2011, pp. 1153–1161. Stroudsburg, PA: Association for Computational Linguistics.
18. Ganter V and Strube M. Finding hedges by chasing weasels: hedge detection using Wikipedia tags and shallow linguistic features. In: *Proceedings of the ACL-IJCNLP 2009 conference short papers (ACLShort '09)*, Singapore, 4 August 2009, pp. 173–176. Stroudsburg, PA: Association for Computational Linguistics.
19. Velldal E. Predicting speculation: a simple disambiguation approach to hedge detection in biomedical literature. *J Biomed Semant* 2011; 2(5): 1–14.
20. Esuli A and Sebastiani F. SENTIWORDNET: a publicly available lexical resource for opinion mining. In: *Proceedings of the 5th conference on language resources and evaluation (LREC'06)*, Genoa, 24–26 May 2006, pp. 417–422. Paris, France: European Language Resources Association.
21. Wang S and Manning CD. Baselines and bigrams: simple, good sentiment and topic classification. In: *Proceedings of the 50th annual meeting of the association for computational linguistics: short papers—volume 2 (ACL'12)*, Jeju Island, Korea, 8–14 July 2012, pp. 90–94. Stroudsburg, PA: Association for Computational Linguistics.
22. Lin C and He Y. Joint sentiment/topic model for sentiment analysis. In: *Proceedings of the 18th ACM conference on information and knowledge management (CIKM '09)*, Hong Kong, 2–6 November 2009, pp. 375–384. New York: ACM.
23. Maas AL, Daly RE, Pham PT, et al. Learning word vectors for sentiment analysis. In: *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies—volume 1 (HTL '11)*, Portland, OR, 19–24 June 2011, pp. 142–150. Stroudsburg, PA: Association for Computational Linguistics.
24. Wang L and Wan Y. Sentiment classification of documents based on latent semantic analysis. In: Lin S and Huang X (eds) *Advanced research on computer education, simulation and modeling*, vol. 176 (Communications in computer and information science). Berlin, Heidelberg: Springer, 2011, pp. 356–361.
25. Socher R, Pennington J, Huang E, et al. Semi-supervised recursive autoencoders for predicting sentiment distributions. In: *Proceedings of the conference on empirical methods in natural language processing (EMNLP '11)*, Edinburgh, Scotland, 27–31 July 2011, pp. 151–161. Stroudsburg, PA: Association for Computational Linguistics.
26. Tang D, Wei F, Yang N, et al. Learning sentiment-specific word embedding for Twitter sentiment classification. In: *Proceedings of the 52nd annual meeting of the association for computational linguistics: long papers—volume 1 (ACL '14)*, Baltimore, MD, 23–25 June 2014, pp. 1555–1565. Stroudsburg, PA: Association for Computational Linguistics.
27. Kalchbrenner N, Grefenstette E and Blunsom P. A convolutional neural network for modelling sentences. In: *Proceedings of the 52nd annual meeting of the association for computational linguistics: long papers—volume 1 (ACL'14)*, Baltimore, MD, 23–25 June 2014, pp. 655–665. Stroudsburg, PA: Association for Computational Linguistics.
28. Le Q and Mikolov T. Distributed representations of sentences and documents. In: *Proceedings of the 31st international conference on machine learning (ICML'14)*, Beijing, China, 21–26 June 2014, pp. 1188–1196.
29. Ekman P. An argument for basic emotions. *Cognition Emotion* 1992; 6(3–4): 169–200.

30. Balabantaray RC, Mohammad M and Sharma N. Multi-class Twitter emotion classification: a new approach. *Int J Appl Inf Syst* 2012; 4(1): 48–53.
31. 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 (WISDOM '12)*, Beijing, China, 12 August 2012, pp. 1–8. New York: ACM.
32. Amigó E, De Albornoz JC, Chugur I, et al. Overview of RepLab 2013: evaluating online reputation monitoring systems. In: Forner P, Müller H, Paredes R, et al. (eds) *Information access evaluation. Multilinguality, multimodality, and visualization*. Berlin, Heidelberg: Springer, 2013, pp. 333–352.
33. Zhao J, Gou L, Wang F, et al. PEARL: an interactive visual analytic tool for understanding personal emotion style derived from social media. In: *Proceedings of IEEE symposium on visual analytics science and technology (VAST '14)*, Paris, France, 9–14 November 2014, pp. 203–212. New York: IEEE.
34. Wanner F, Stoffel A, Jäckle D, et al. State-of-the-art report of visual analysis for event detection in text data streams. *Comput Graph Forum* 2014; 33(3): 1–15.
35. Keim DA, Krstajić M, Rohrdantz C, et al. Real-time visual analytics for text streams. *Computer* 2013; 46(7): 47–55.
36. Nguyen VD, Varghese B and Barker A. The royal birth of 2013: analysing and visualising public sentiment in the UK using Twitter. In: *Proceedings of IEEE international conference on big data (BigData '13)*, Santa Clara, CA, 6–9 October 2013, pp. 46–54. New York: IEEE.
37. Görg C, Liu Z, Kihm J, et al. Combining computational analyses and interactive visualization for document exploration and sensemaking in Jigsaw. *IEEE T Vis Comput Gr* 2013; 19(10): 1646–1663.
38. Alencar AB, de Oliveira MCF and Paulovich FV. Seeing beyond reading: a survey on visual text analytics. *Wiley Interdiscip Rev: Data Min Knowl Discov* 2012; 2(6): 476–492.
39. Gan Q, Zhu M, Li M, et al. Document visualization: an overview of current research. *Wiley Interdiscip Rev: Comput Stat* 2014; 6(1): 19–36.
40. Kerren A, Kyusakova M and Paradis C. Chapter 5. From culture to text to interactive visualization of wine reviews. In: Marchese FT and Banissi E (eds) *Knowledge visualization currents: from text to art to culture*. Oxford: Springer, 2012, pp. 85–110.
41. Kucher K and Kerren A. Text visualization browser: a visual survey of text visualization techniques. In: *Poster abstracts of IEEE information visualization (InfoVis '14)*, Paris, France, 9–14 November 2014.
42. Havre S, Hetzler E, Whitney P, et al. ThemeRiver: visualizing thematic changes in large document collections. *IEEE T Vis Comput Gr* 2002; 8(1): 9–20.
43. Dou W, Yu L, Wang X, et al. Hierarchical topics: visually exploring large text collections using topic hierarchies. *IEEE T Vis Comput Gr* 2013; 19(12): 2002–2011.
44. Xu P, Wu Y, Wei E, et al. Visual analysis of topic competition on social media. *IEEE T Vis Comput Gr* 2013; 19(12): 2012–2021.
45. Bosch H, Thom D, Heimerl F, et al. ScatterBlogs2: real-time monitoring of microblog messages through user-guided filtering. *IEEE T Vis Comput Gr* 2013; 19(12): 2022–2031.
46. Liu B, Hu M and Cheng J. Opinion observer: analyzing and comparing opinions on the web. In: *Proceedings of the 14th international conference on World Wide Web (WWW '05)*, Chiba, Japan, 10–14 May 2005, pp. 342–351. New York: ACM.
47. Oelke D, Hao M, Rohrdantz C, et al. Visual opinion analysis of customer feedback data. In: *Proceedings of IEEE symposium on visual analytics science and technology (VAST '09)*, Atlantic City, NJ, 12–13 October 2009, pp. 187–194. New York: IEEE.
48. Wanner F, Rohrdantz C, Mansmann F, et al. Visual sentiment analysis of RSS news feeds featuring the US presidential election in 2008. In: *Proceedings of the IUI workshop on visual interfaces to the social and the semantic web (VISSW '09)*, Sanibel, FL, 8 February 2009.
49. Cui W, Qu H, Zhou H, et al. Watch the story unfold with TextWheel: visualization of large-scale news streams. *ACM Trans Intell Syst Technol* 2012; 3(2): 20:1–20:17.
50. Rohrdantz C, Hao MC, Dayal U, et al. Feature-based visual sentiment analysis of text document streams. *ACM Trans Intell Syst Technol* 2012; 3(2): 26:1–26:25.
51. Wanner F, Weiler A and Schreck T. Topic Tracker: shape-based visualization for trend and sentiment tracking in Twitter. In: *Proceedings of the 2nd IEEE workshop on interactive visual text analytics "Task-Driven Analysis of Social Media" (IEEE VisWeek '12)*, Seattle, WA, 15 October 2012.
52. Zhang C, Liu Y and Wang C. Time-space varying visual analysis of micro-blog sentiment. In: *Proceedings of the 6th international symposium on visual information communication and interaction (VINCI '13)*, Tianjin, China, 17–18 August 2013, pp. 64–71. New York: ACM.
53. Hao MC, Rohrdantz C, Janetzko H, et al. Visual sentiment analysis of customer feedback streams using geo-temporal term associations. *Inform Visual* 2013; 12(3–4): 273–290.
54. Wang C, Xiao Z, Liu Y, et al. SentiView: sentiment analysis and visualization for internet popular topics. *IEEE T Hum Mach Syst* 2013; 43(6): 620–630.
55. Fekete JD and Dufournaud N. Compus: visualization and analysis of structured documents for understanding social life in the 16th century. In: *Proceedings of the fifth ACM conference on digital libraries (DL '00)*, San Antonio, TX, June 2–7 2000, pp. 47–55. New York: ACM.
56. Correll M, Witmore M and Gleicher M. Exploring collections of tagged text for literary scholarship. *Comput Graph Forum* 2011; 30(3): 731–740.
57. Siirtola H, Säily T, Nevalainen T, et al. Text variation explorer: towards interactive visualization tools for corpus linguistics. *Int J Corpus Linguist* 2014; 19(3): 417–429.

58. Regan T and Becker L. Visualizing the text of Philip Pullman's trilogy "His Dark Materials." In: *Proceedings of the 6th Nordic conference on human-computer interaction: extending boundaries (NordCHI '10)*, Reykjavik, Iceland, 16–20 October, 2010, pp. 759–764. New York: ACM.
59. Geng Z, Cheesman T, Laramee RS, et al. ShakerVis: visual analysis of segment variation of German translations of Shakespeare's Othello. *Inform Visual*. Epub ahead of print 23 July 2013. DOI: 10.1177/1473871613495845.
60. Jänicke S, Geßner A, Büchler M, et al. Visualizations for text re-use. In: *Proceedings of the international conference on information visualization theory and applications (IVAPP '14)*, Lisbon, Portugal, 5–8 January 2014, pp. 59–70. Lisbon: SciTePress.
61. Rohrdantz C, Niekler A, Hautli A, et al. Lexical semantics and distribution of suffixes: a visual analysis. In: *Proceedings of the EACL 2012 joint workshop of LINGVIS & UNCLH (EACL '12)*, Avignon, 23–24 April 2012, pp. 7–15. Stroudsburg, PA: Association for Computational Linguistics.
62. Kabán A and Girolami MA. A dynamic probabilistic model to visualise topic evolution in text streams. *J Intell Inf Syst* 2002; 18(2–3): 107–125.
63. Brandes U and Corman SR. Visual unrolling of network evolution and the analysis of dynamic discourse. *Inform Visual* 2003; 2(1): 40–50.
64. Angus D, Smith A and Wiles J. Conceptual recurrence plots: revealing patterns in human discourse. *IEEE T Vis Comput Gr* 2012; 18(6): 988–997.
65. Diakopoulos N, Zhang AX, Elgesem D, et al. Identifying and analyzing moral evaluation frames in climate change blog discourse. In: *Proceedings of the international conference on weblogs and social media (ICWSM '14)*, Ann Arbor, MI, 1–4 June 2014, pp. 583–586. Palo Alto, CA: Association for the Advancement of Artificial Intelligence.
66. Hoque E and Carenini G. ConVis: a visual text analytic system for exploring blog conversations. *Comput Graph Forum* 2014; 33(3): 221–230.
67. Gregory ML, Chinchor N, Whitney P, et al. User-directed sentiment analysis: visualizing the affective content of documents. In: *Proceedings of the workshop on sentiment and subjectivity in text (SST '06)*, Sydney, Australia, 22 July 2006, pp. 23–30. Stroudsburg, PA: Association for Computational Linguistics.
68. Makki R, Brooks S and Milios EE. Context-specific sentiment lexicon expansion via minimal user interaction. In: *Proceedings of the international conference on information visualization theory and applications (IVAPP '14)*, Lisbon, Portugal, 5–8 January 2014, pp. 178–186. Lisbon: SciTePress.
69. D3—data-driven documents, http://d3js.org/ (accessed 19 August 2014).
70. Rickshaw: a JavaScript toolkit for creating interactive time-series graphs, http://code.shutterstock.com/rickshaw/ (accessed 19 August 2014).
71. Fielding RT and Taylor RN. Principled design of the modern web architecture. *ACM T Internet Techn* 2002; 2(2): 115–150.
72. Strapparava C and Valitutti A. WordNet affect: an affective extension of WordNet. In: *Proceedings of the 4th international conference on language resources and evaluation (LREC '04)*, Lisbon, Portugal, 24-30 May 2004., vol. 4, pp. 1083–1086. Paris, France: European Language Resources Association.
73. GeneralInquirer, http://www.wjh.harvard.edu/~inquirer/ (accessed 19 August 2014).
74. The Compass DeRose guide to emotion words, http://www.derose.net/steve/resources/emotionwords/ewords.html (accessed 19 August 2014).
75. Wikipedia. Lists of weapons– Wikipedia, The Free Encyclopedia. Available at:http://en.wikipedia.org/wiki/Lists_of_weapons (accessed 13 March 2015).
76. ColorBrewer 2.0—color advice for cartography, http://colorbrewer2.org/ (accessed 20 August 2014).
77. Tennekes M and de Jonge E. Tree colors: color schemes for tree-structured data. *IEEE T Vis Comput Gr* 2014; 20(12): 2072–2081.
78. Kerren A and Schreiber F. Toward the role of interaction in visual analytics. In: *Proceedings of the winter simulation conference (WSC '12)*, Berlin, 9–12 December 2012, pp. 420:1–420:13. New York: IEEE.
79. Cernea D, Truderung I, Kerren A, et al. WebComets: a tab-oriented approach for browser history visualization. In: *Proceedings of the international conference on information visualization theory and applications (IVAPP '14)*, Lisbon, Portugal, 5–8 January 2014, pp. 439–450. Lisbon: SciTePress.
80. Hamers L, Hemeryck Y, Herweyers G, et al. Similarity measures in scientometric research: the Jaccard index versus Salton's cosine formula. *Inform Process Manag* 1989; 25(3): 315–318.
81. Wattenberg M. Arc diagrams: visualizing structure in strings. In: *Proceedings of IEEE symposium on information visualization (INFOVIS '02)*, Boston, MA, 28–29 October 2002, pp. 110–116. New York: IEEE.
82. He H, Sýkora O and Vrt'o I. Crossing minimisation heuristics for 2-page drawings. *Electron Notes Discrete Math* 2005; 22: 527–534.
83. Jericho HTML Parser, http://jericho.htmlparser.net/ (accessed 20 August 2014).
84. Viégas F, Wattenberg M, van Ham F, et al. ManyEyes: a site for visualization at Internet scale. *IEEE T Vis Comput Gr* 2007; 13(6): 1121–1128.
85. Manning CD and Schütze H. *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press, 1999.
86. Broderick R. Coca-Cola's Multilingual Super Bowl Ad Inspired A Complete Meltdown Online. Available at: http://www.buzzfeed.com/ryanhatesthis/coca-colas-multilingual-super-bowl-ad-inspired-a-racist-mel#v3khrr (accessed 13 March 2015).
87. Aravosis J. Racists explode at Coke for Super Bowl ad singing "America the Beautiful" in foreign languages. Available at:http://americablog.com/2014/02/bigots-pod-coke-super-bowl-ad-singing-national-anthem.html (accessed 13 March 2015).