Visualizing Contextual and Dynamic Features of Micropost Streams

Alexander Hubmann-Haidvogel, Adrian M.P. Braşoveanu, Arno Scharl, Marta Sabou, Stefan Gindl
MODUL University Vienna
Department of New Media Technology
Am Kahlenberg 1
1190 Vienna, Austria

{alexander.hubmann, adrian.brasoveanu, arno.scharl, marta.sabou, stefan.gindl}@modul.ac.at

ABSTRACT
Visual techniques provide an intuitive way of making sense of the large amounts of microposts available from social media sources, particularly in the case of emerging topics of interest to a global audience, which often raise controversy among key stakeholders. Micropost streams are context-dependent and highly dynamic in nature. We describe a visual analytics platform to handle high-volume micropost streams from multiple social media channels. For each post we extract key contextual features such as location, topic and sentiment, and subsequently render the resulting multi-dimensional information space using a suite of coordinated views that support a variety of complex information seeking behaviors. We also describe three new visualization techniques that extend the original platform to account for the dynamic nature of micropost streams through dynamic topography information landscapes, news flow diagrams and longitudinal cross-media analyses.

Categories and Subject Descriptors
H.5.2 [Information Interfaces and Presentation]: User Interfaces – interaction styles. I.3.6. [Computer Graphics]: Methodology and Techniques – Interaction Technique.

General Terms
Algorithms, Measurement, Design, Human Factors

Keywords
Social Media Analytics, Microposts, Contextual Features, News Flow, Dynamic Visualization, Information Landscape

1. INTRODUCTION
The ease of using social media channels has enabled people from around the world to express their opinions and propagate local or global news about virtually every imaginable topic. In doing so, they most often make use of short text messages (tweets, status updates) that we collectively refer to as microposts. Given the high volume, diversity and complex interdependency of social media-specific micropost streams, visual techniques play an increasingly important role in making sense of these novel data sources. Visual techniques can support analysts, journalists and marketing managers alike in taking the pulse of public opinion, in understanding the perceptions and preferences of key stakeholders, in detecting controversies, and in measuring the impact and diffusion of public communications. This is particularly true for domains that pose challenges through their global reach, the competing interests of many different stakeholders, and the dynamic and often conflicting nature of relevant evidence sources (e.g., environmental issues, political campaigns, financial markets).

To support such scenarios across application domains, we have developed a (social) media monitoring platform with a particular focus on visual analytics (Hubmann, 2009). The platform enables detecting and tracking topics that are frequently mentioned in a given data sample (e.g., a collection of Web documents crawled from relevant sources). The advanced data mining techniques underlying the platform extract a variety of contextual features from the document space. A visual interface based on multiple coordinated views allows exploring the evolution of the document space along the dimensions defined by these contextual features (temporal, geographic, semantic, and affective), and subsequent drill-down functionalities to analyze details of the data itself. In essence, the platform has the key characteristics of a decision support system, namely: 1) it aggregates data from many diverse sources; 2) it offers an easy to use visual dashboard for observing global trends; 3) it allows both a quick drill-down and complex analyses along the dimensions of the extracted contextual features. We briefly report on the overall platform in Section 3.

The platform has been originally designed to analyze traditional news media, but from early 2011 we have adapted it to support micropost analysis, taking advantage of the robust infrastructure for crawling, analyzing and visualizing Web sources. The multi-dimensional analysis enabled by the original design of the portal is well suited for analyzing contextual features of microposts. However, the visualization metaphors did not properly capture the highly dynamic nature of micropost streams, nor did they allow cross-comparison between social and traditional media sources. Our latest research therefore focuses on novel methods to support temporal analysis and cross-media visualizations. In sections 4, 5 and 6 we describe these novel visualizations.

2. RELATED WORK
With the rise of the social networks (Heer, 2005), understanding large-scale events through visualization emerged as an important research topic. Various visual interfaces have been designed for inspecting news or social media streams in diverse domains such as sports (Marcus, 2011), politics, (Diakopoulos, 2010; Shamma, 2009; Shamma, 2010), and climate change (Hubmann, 2009).

Published as part of the #MSM2012 Workshop proceedings, available online as CEUR Vol-838, at: http://ceur-ws.org/Vol-838
#MSM2012, April 16, 2012, Lyon, France.

Researchers have emphasized different aspects of extracting useful information including (sub-)events (Adams, 2011), topics (Hubmann, 2009), and video fragments (Diakopoulos, 2011). *Vox Civitas*, for example, is a visual analytic tool that aims to support journalists in getting useful information from social media streams related to televised debates and speeches (Diakopoulos, 2010). In terms of the number and type of social media channels that are visualized, most approaches focus primarily on Twitter, while streams from Facebook and YouTube are visualized to a lesser extent (Marcus, 2010). We regard these three channels as equally important and visualize their combined content.

To reflect the dynamic nature of social media channels, some visualizations provide real-time updates displaying messages as they are published, and also projecting them onto a map – e.g., *TwitterVision.com* or *AWorldofTweets.frogdesign.com*. Given the computational overhead, however, real-time visualizations are the exception rather than the norm, since most projects rely on update times anywhere between a few minutes and a few days.

Visual techniques render microposts along dimensions derived from their contextual features. Most frequently, visualization rely on temporal and geographic features, but increasingly they exploit more complex characteristics such as the sentiment of the micropost, its content (e.g., expressed through relevant keywords), or characteristics of its author. Indeed, user clustering as seen in *ThemeCrowds* (Archambault, 2011) or geographical maps (e.g., (Marcus, 2011), *TwitterReporter* (Meyer, 2011)) are must-have features for every system that aims to understand local news and correlate them with global trends. Commercial services such as *SocialMention.com* and *AlertRank.com* use visualizations to track sentiment across tweets. During the 2010 U.S. Midterm Elections, sentiment visualizations have been present in all major media outlets from *New York Times* to *Huffington Post* (Peters, 2010).

Fully utilizing contextual features requires the use of appropriate visual metaphors. In general, social media visualizations rely on one of the following three visual metaphors:

- *Multiple Coordinated Views*, also known as linked or tightly coupled views in the literature (Scharl, 2001), (Hubmann, 2009), ensures that a change in one of the views triggers an immediate update within the others. For example, the interface of *Vox Civitas* uses coordinated views to synchronize a timeline, a color-coded sentiment bar, a Twitter flow and a video window which helps linking parts of the video to relevant tweets (Diakopoulos, 2010). Additionally, (Marcus, 2011) use the multiple coordinated views in their system geared towards Twitter events and offer capabilities to drill down into sub-events and explore them based on geographic location, sentiment and link popularity.
- *Visual Backchannels* (Dork, 2010) represent interactive interfaces synchronizing a topic stream (e.g., a video) with real-time social media streams and additional visualizations. This concept has evolved from the earlier concept of *digital backchannel*, referring to news media outlets supplementing their breaking news coverages with relevant tweets – e.g., during political debates or sport games (Shamma, 2010). However, additionally to the methods described in *Hack the Debate* (Shamma, 2009) and *Statler* (Shamma, 2010), tools that use the visual backchannel metaphor display not only the Twitter flow that corresponds to certain media events such as debates, but also a wealth of graphics and statistics.
- *Timelines* follow the metaphor with the longest tradition, well suited for displaying the evolution of topics over time. Aigner et al. present an extensive collection of commented timelines (Aigner, 2011). The work by Adams et al. (Adams, 2011) is similar to our approach as it combines a color-coded sentiment display with interactive tooltips.

Beyond understanding micropost streams, a challenging research avenue compares the content of social media coverage with that of traditional news outlets. Cross-media analysis based on social sources is a relatively new field, but promising results have been published recently. In most cases comparisons are made between two sources such as Twitter and New York Times (Zhao, 2011), or Twitter and Yahoo! News (Hong, 2011). (Zhao, 2011) compares a Twitter corpus with a New York Times corpus to detect trending topics. For the New York Times, they apply a direct Latent Dirichlet Allocation (LDA), while for Twitter they use a modified LDA under the assumption that most tweets refer to a single topic. They use metrics including the distribution of categories, breadth of topics coverage, opinion topic and the spread of topics through re-tweets, and show that most Twitter topics are not covered appropriately by traditional news media channels. They conclude that for spreading breaking world news, Twitter seems to be a better platform than a traditional medium such as New York Times. Hong et al. compare Twitter with Yahoo! News to understand temporal dynamics of news topics (Hong, 2011). They show that local topics do not appear as often in Twitter, and they go on to compare the performance of different models (LDA, Temporal Collection, etc). (Lin, 2011) conducts a study on media biasing on both social networks and news media outlets, but is focused only on the quantity of mentions. While these studies highlight differences between social and news media, they typically lack visual support for monitoring diverse news sources.

3. ACQUISITION AND AGGREGATION OF CLIMATE CHANGE MICROPOSTS

Climate Change is a global issue characterized by diverse opinions of different stakeholders. Understanding the key topics in this area, their global reach and the opinions voiced by different parties is a complex task that requires investigating how these dimensions relate to each other. The *Media Watch on Climate Change* portal (www.ecoresearch.net/climate) addresses this task by providing advanced analytical and visual methods to support different types of information seeking behavior such as browsing, trend monitoring, analysis and search.

The underlying technologies have originally been developed for monitoring traditional news media (Hubmann, 2009) and have recently been adapted for use with social media sources, in particular micropost content harvested from *Twitter, YouTube* and *Facebook*. Between April 2011 and March 2012, the system has collected and analyzed an estimated 165,000 microposts from these channels. To support a detailed analysis of the collected microposts, we use a variety of visual metaphors to interact with contextual features along a number of dimensions: temporal, geographic, semantic and attitudinal. A key strength of the interface is the rapid synchronization of *multiple coordinated views*. It allows selecting the relevant data sources and provides trend charts, a document viewing panel as well as just-in-time information retrieval agents to retrieve similar documents in terms of either topic or geographic location. The right side of the interface contains a total of four different visualizations (two of which are being shown in Figure 1), which capture global views on the dataset. In addition to the shown semantic map (= information landscape; see Section 4) and tag cloud, users can also select a geographic map and an ontology graph. Any of these views can be closed, maximized, or opened in a separate pop-up window to allow a more thorough inspection (the views remain synchronized even when placed in different windows). While the *Media Watch on Climate Change* focuses on environmental coverage, the same technologies are currently being used for other domains as well, for example, for the Web intelligence platforms of the National Oceanic and Atmospheric Administration (NOAA), the National Cancer Institute (NCI), and the Vienna Chamber of Commerce and Industry (see www.weblyzard.com).

The portal's visualizations provide a good starting point for analyzing microposts along a variety of contextual features, in particular in the area of climate change. The system does not reflect the dynamic character of these micropost streams, however, and therefore misses a key benefit of social media – that of capturing events as they unfold. To overcome this limitation, we are currently developing the following set of novel visualizations that focus on the longitudinal and temporal analysis of micropost streams:

1. The dynamic topography information landscapes are an extension of the information landscapes paradigm. Instead of capturing the state of the information space at discrete moments in time, the topography is continuously updated as new microposts are being published (Section 4).
2. The news flow diagrams visualize microposts from multiple social media channels in real time, and reveal correlations in terms of the topics that they mention (Section 5).
3. The cross-media analysis charts allow longitudinal analyses of frequency and sentiment for any given topic and across data sources (e.g., between social media, news media, and the blogosphere; see Section 6).

4. DYNAMIC TOPOGRAPHY INFORMATION LANDSCAPE

Information Landscapes represent a powerful visualization technique for conveying topical relatedness in large document repositories (Krishnan, 2007). Yet, the traditional concept of information landscapes only allows for visualizing static conditions. We have made use of such static landscapes when visualizing traditional news media, which were less dynamic than social media sources and where it was sufficient to recompute the information landscape at weekly intervals.

For visualizing highly dynamic micropost streams, however, this is not a satisfying solution. What is required instead is a visual representation such as *ThemeRiver* (Havre, 2002) that conveys changes in topical clusters. Unfortunately, most of these representations lack the means to express complex topical relations and are therefore no substitute for the information landscape metaphor.

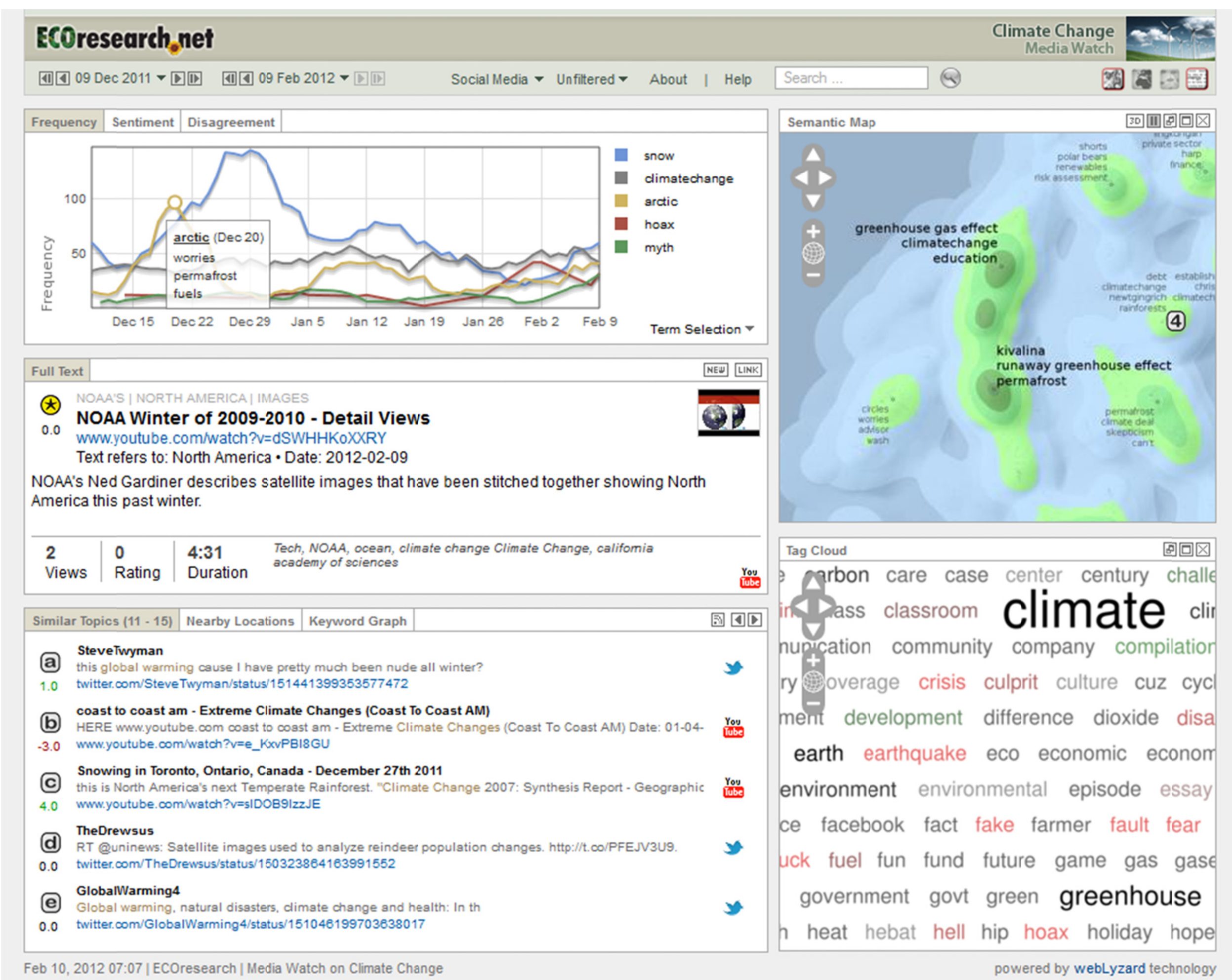

Figure 1. Screenshot of the Media Watch on Climate Change (www.ecoresearch.net/climate)

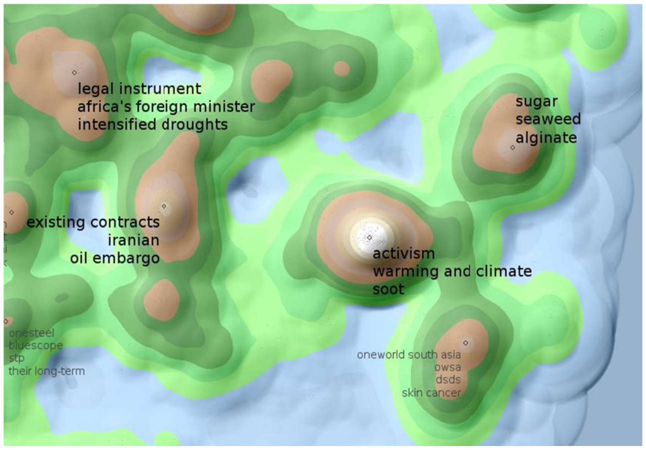

Figure 2. Information landscape on climate change coverage

In previous research, we have introduced dynamic topography information landscapes (Sabol et al., 2010) to address both topical relatedness and rapidly changing data. Dynamic topography information landscapes are visual representations based on a geographic map metaphor where topical relatedness is conveyed through spatial proximity in the visualization space with hills representing agglomerations (clusters) of topically similar documents. As shown in Figure 2, the hills are labeled with sets of dominant keyword labels (n-grams) extracted from the underlying documents to facilitate the users' orientation.

Micropost streams are characterized by the rapid emergence and decay of topics. The topical structure changes with each new posting. Dynamic information landscapes convey these changes as tectonic processes which modify the landscape topography accordingly. Rising hills indicate the emergence of new topics; shrinking hills a fading of existing ones. In the process of generating information landscapes, high-dimensional data is projected into a lower-dimensional space.

5. NEWS FLOW DIAGRAM

While the dynamic topography information landscape metaphor depicts the evolution of topic clusters within a collection of microposts without differentiating their origin, some scenarios require a comparative analysis of individual micropost streams. The two key problems related to the visualization of microposts originating from multiple social media sources is to show their provenance as well as the dynamic changes of topical associations between them. The *News Flow Diagram* concept addresses these issues by integrating several visual metaphors into a single display (see screenshot in Figure 3):

1. *Falling bar graphs* – Each micropost (Twitter message, Facebook status update, YouTube message) is represented internally through the title of the post, its time of publishing, its content, and a list of associated keywords. When a new micropost is posted we visualize, in real-time, its respective associated keywords through falling words. One document generates one falling word for each mentioned topic. The falling words will "hit" the lower part of the visualization and dissolve into the corresponding keyword bar, which increases in size accordingly. Figure 3 depicts how topics fall towards their respective keyword bars (e.g., "experiences" and "friends" in the upper diagram). A keyword bar collects all mentions of a certain topic in microposts from different social media channels and, therefore, its height correlates with the popularity of the topic in the social media outlets that are visualized. The falling bar metaphor was quite popular a few years ago due to the success of the *Digg Stack* visualization [Baer 2008].
2. *Multi-source stacked bars and color-coded sentiment bars*. Each falling word is color-coded to represent either its provenance or its associated sentiment value. Figure 3, for example, uses the color of the falling words to reflect their origin (Twitter: gray, Facebook: blue marine, YouTube: red). This color-coding is maintained in the keyword bars, each bar showing through its diversely colored portions the percentage of mentions of the corresponding keyword within the individual media sources. This allows inferring the most and least mentioned topics across sources. The same metaphor can be used to show sentiment values instead of provenance (not shown in Figure 3).
3. *Threaded arcs*. We use an adaptation of the threaded arcs display to convey associations between the keywords that appear in the same document. Dynamic link patterns conveyed through shifting arcs allow us to understand how the associations, initially displayed through falling bars, modify over time. Related keywords are highlighted to quickly notice them. Figure 3b shows the topic "ideas", which has appeared seven times, co-occurs most frequently with the two topics: "professor" and "response". These threaded arcs are only displayed when we click on a keyword bar.

Color-coding is an important part of this visualization. We use it to highlight various aspects of the data:

- *Sentiment coloring* – the color of the bars can represent the sentiment of a certain topic (see Figure 3a);
- *Source coloring* – words can also be displayed as stacked bars with specific colors that represent the sources in the stacked layout (provenance information); Figure 3b, for example, shows a situation where we have three sources (Facebook, Twitter, YouTube) and keywords from one source (YouTube) falling;
- *Arc coloring* – we use darker shades of gray for stronger relations between the terms (i.e., they co-occur more frequently), connecting the most related terms.

To demonstrate how associations evolve over time, we show the same word ("ideas") in both diagrams: Figure 3a uses color coding for sentiment information, Figure 3b for distinguishing the source (users can easily switch between both modes). In Figure 3a, "ideas" has a stronger connection with "response" than with "oxfam", but the word has only two hits. Figure 3b shows that after seven hits, "ideas" has a stronger connection with "professor", than with "response", and also the same weak link with "oxfam". Future versions of the visualization module will include information related to these connections in the tooltips, emphasizing the importance of interactivity and revealing the evolution of connections over time.

This visualization showcases the powerful mechanism of combining various visual metaphors with color-coding. We use the news flow diagrams to identify key topics (we only show the 50 most important terms), to describe the relations between them (co-occurrence of terms in a micropost are displayed through the falling bars), and to show the evolution of social media coverage over time (dynamic changes in the distribution of keywords/topics across various social media sources is represented through the lower arcs).

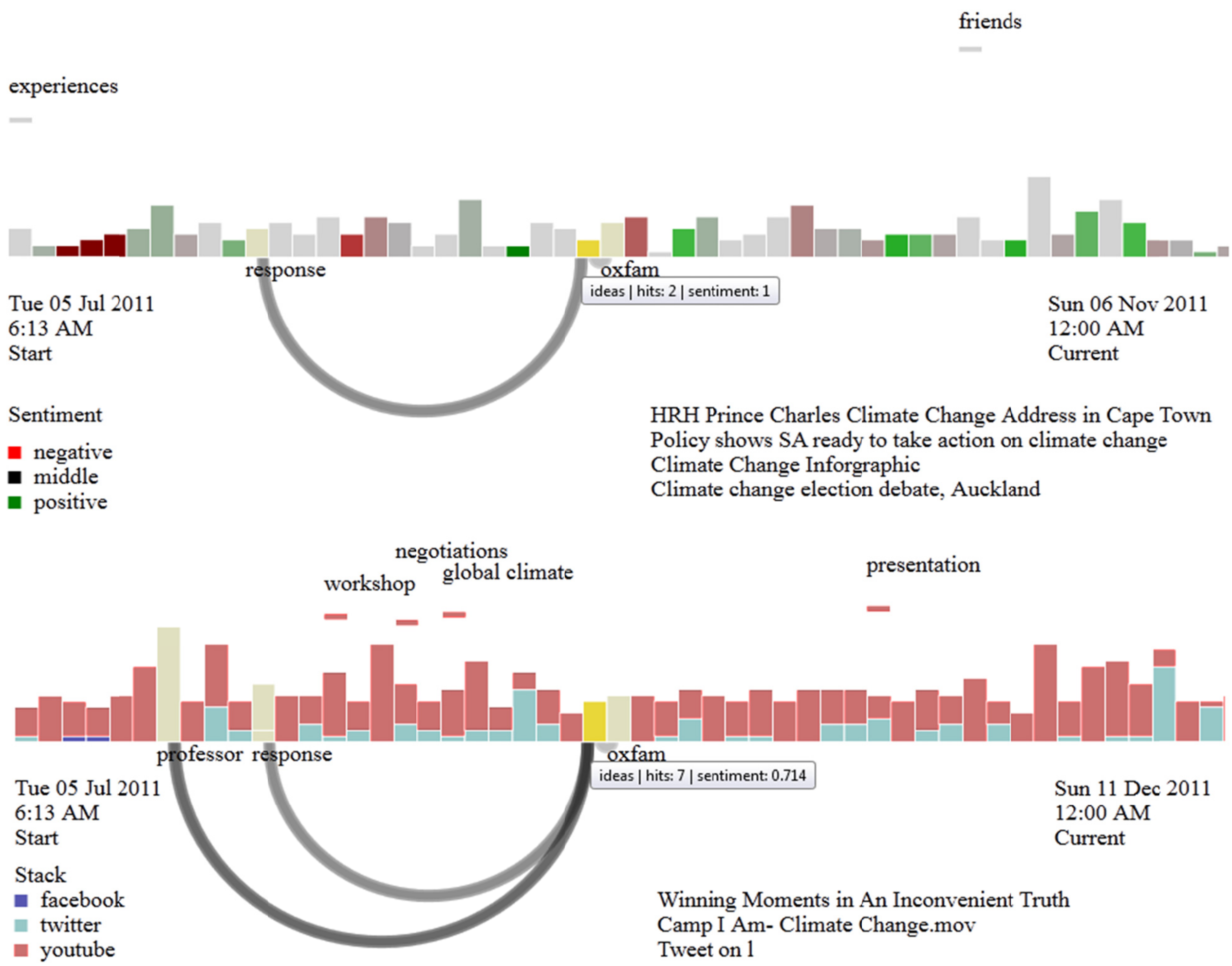

Figure 3. News Flow Diagram with color coding for showing sentiment (Figure 3a, above) and source (Figure 3b, below).

6. CROSS-MEDIA ANALYSIS

The Media Watch on Climate Change offers longitudinal analysis (i.e., monitoring over time) in terms of topic frequency, sentiment associated to a topic and disagreement over a topic. However, these trend charts are only plotted over a single data source (e.g., either news media or social media) and are available only for a set of pre-computed topics. Therefore they are neither suitable for social media streams where new topics emerge rapidly, nor do they allow comparing across different media sources.

To overcome these limitations, we are developing the new visualization shown in Figure 4, which allows (a) defining a topic to be monitored over time and (b) monitoring this topic across different media sources selected by clicking the appropriate check-boxes in the interface (e.g., traditional news media outlets, blogs, social networks such as Twitter, YouTube and Facebook). The visualization makes use of the data collection and charting frameworks of the portal to plot both frequency and sentiment related charts.

By plotting topic frequency (i.e., number of documents per day that mention that topic) over time, this visualization shows the impact of a topic on different media sources. For example, the screenshot in Figure 4 depicts a query for the topic "durban" and compares the amount of news coverage about the *17th Conference of the Parties to the United Nations Framework Convention on Climate Change* (COP17) held in Durban, South Africa, from 01 Nov to 31 Dec 2011 in traditional news media, Twitter postings, blogs, and NGOs. Coinciding with the beginning of the conference on the 28th of November, both samples show an increase in the coverage of this topic. The frequency then declines sharply after the end of the event, which is an effect more pronounced in the news media coverage. It also shows that coverage of the conference has been far more intense in news media than in micropost streams, except a short period of time in December.

In addition to frequency charts, we also visualize the sentiment of the documents mentioning a specific topic. A set of charts shows either positive or negative documents, average sentiment of documents for a day or the standard deviation of the sentiment over time. These charts help to understand the attitude expressed in different media outlets, e.g., which outlet has the most negative or positive documents, which outlet is characterized by the most controversies? A comparison of the average sentiment towards "COP17" in social media and news media channels showed a gradual shift from positive to negative in microposts, while news media sentiment remained positive during the entire duration of the event (see screenshot in Figure 4).

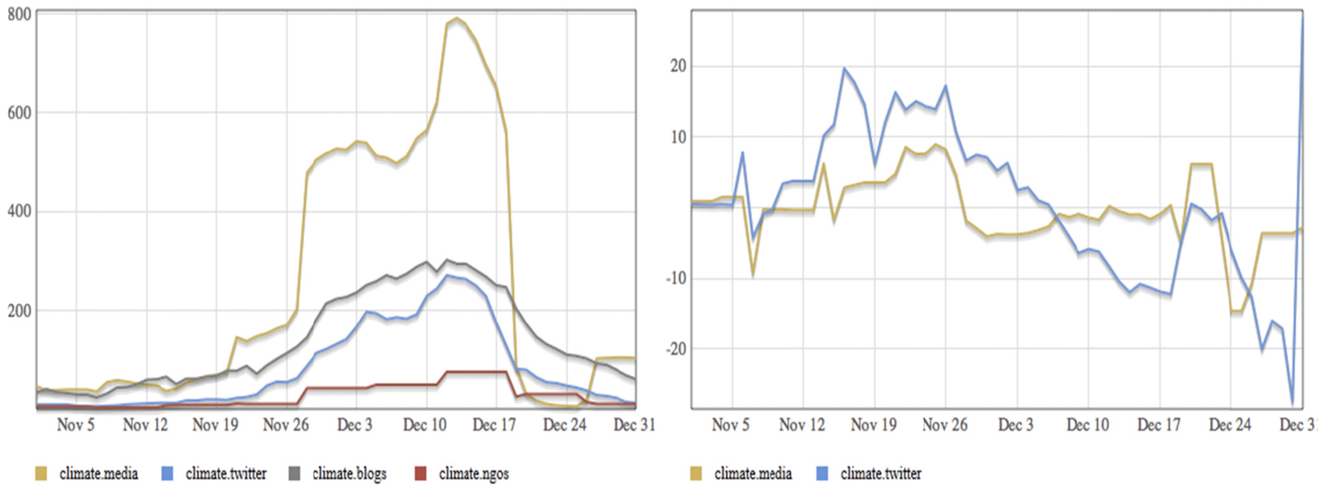

Figure 4. Cross-media analysis for November – December 2011; term frequency distribution for "Durban" (left) and average sentiment towards "COP17" (right).

An important issue of visualizing sentiment across media outlets is the meaningful computation of sentiment values for disparate documents. The sentiment detection algorithm cumulatively adds up the sentiment values of individual words in a document to compute an overall sentiment value for the document, which is then normalized based on the total number of tokens in the document. This allows the comparison of documents of different lengths, such as news articles and microposts.

The visualization can not only track user-specified topics, but can also assist the user in finding similar topics by providing a list of top terms associated with the query term. These associated terms are calculated using a combination of significant phrases detection and co-occurrence analysis on the document set (Hubmann, 2009), and are aggregated and ranked depending on documents matching the query term. A query for "COP17", for example, yields the terms "Durban", "UNFCC" and "Climate Change" as associated terms in Twitter microposts. Additional query term disambiguation is not required in this case, as the documents collected are already pre-filtered based on their relevance to the climate change domain.

7. CONCLUSION AND OUTLOOK

In this paper we describe recent work on making sense of microposts through visual means. Our earlier work on the *Media Watch on Climate Change* portal (www.ecoresearch.net/climate) focused on visual analytics over traditional news media and relied on extracting and visualizing a wealth of context features. This characteristic of the portal proved essential when adapting it to visualizing micropost streams from three main social media channels, as it enabled complex analysis along temporal, geographic, semantic and attitudinal dimensions in the challenging domain of climate change. Unlike many other social media visualizations, the presented approach relies on a robust infrastructure and combines data from multiple social media outlets.

While the contextual nature of the microposts has been fully capitalized upon, the existing visualizations fell short of conveying another key characteristic of microposts, namely their dynamic nature. This initiated research into the new visualizations described in this paper, including: (i) dynamic topographic information landscapes, which show through tectonic changes how major topic clusters evolve; (ii) the news flow diagrams which enable a fine-grained, comparative analysis across micropost streams, showing key topics being discussed and how they relate to each other; and (iii) cross-media analysis based on longitudinal datasets containing frequency and sentiment information.

Future work will focus on feature extraction from microposts and visualizations to depict contextual and dynamic characteristics of microposts. We are currently working on more robust methods for extracting contextual features from microposts, by further adapting our current methods to the particularities of these texts. Some of the features that we intend to introduce in future releases are related to interactive timelines and time series analysis.

Future research will also allow comparing timelines across topics and related to specific events. We will use timelines as a starting point for narrative visualizations (e.g., replaying the history of an event or a chain of events; identifying visual patterns that best describe a chain of events on social media). We will compare various media channels since our datasets and graphical tools are well suited for such an analysis. Finally, we will investigate novel ways of incorporating these individual visualizations to support the complex analytical scenarios of decision making tools.

8. ACKNOWLEDGMENTS

Key components of the system presented in this paper were developed within the DIVINE (www.weblyzard.com/divine) and Triple-C (www.ecoresearch.net/triple-c) research projects, funded by *FIT-IT Semantic Systems* of the Austrian Research Promotion Agency and the *Austrian Climate Research Program* of the Austrian Climate and Energy Fund (www.ffg.at; klimafonds.gv.at).

Adrian Brașoveanu was partially supported by the strategic grant POSDRU/88/1.5/S/60370 (2009) on "Doctoral Scholarships" of the Ministry of Labor, Family and Social Protection, Romania, co-financed by the European Social Fund – Investing in People.

9. REFERENCES

[Adams, 2011] Brett Adams, Dinh Phung, Svetha Venkatesh. 2011. Eventscapes: visualizing events over times with emotive facets. In *MM '11 Proceedings of the 19th ACM International Conference on Multimedia*, Scottsdale, AZ, USA (November 28 - December 01, 2011), 1477-1480.

[Aigner, 2011] Wolfgang Aigner, Silvia Miksch, Heidrun Schumann Christian Tominski. 2011. *Visualization of Time-Oriented Data.* Springer, 2011.

[Archambault, 2011] D. Archambault, D. Greene, P. Cunningham, and N. Hurley. ThemeCrowds: Multiresolution Summaries of Twitter Usage. In Proc. of the 3rd Workshop on Search and Mining User-generated Contents, Glasgow, UK, October 2011.

[Baer, 2008] K. Baer. Information Design Workbook. Graphic Approaches, Solutions and Inspiration + 30 Case Studies. Rockport Publishers, 2008.

[Diakopoulos, 2010] N. Diakopoulos, M. Naaman, F. Kivranswain. 2010. Diamonds in the Rough: Social Media Visual Analytics for Journalistic Inquiry, *IEEE Symposium on Visual Analytics Science and Technology (VAST)*.

[Dork, 2010] Marian Dork, Daniel Gruen, Carey Williamson, and Sheelagh Carpendale. A visual backchannel for large-scale events. TVCG: Transactions on Visualization and Computer Graphics, 16(6):1129-1138, Nov/Dec 2010.

[Havre, 2002] Havre, S., Hetzler, E., Whitney, P., and Nowell, L.: ThemeRiver: Visualizing Thematic Changes in Large Document Collections. IEEE Transactions on Visualization and Computer Graphics, 8(1):9–20, 2002.

[Heer, 2005] J. Heer and D. Boyd. Vizster: Visualizing online social networks. In *IEEE Symposium on Information Visualization*, 2005.

[Hong, 2011] Liangjie Hong, Byron Dom, Siva Gurumurthy, Kostas Tsioutsiouliklis. 2011. A Time-Dependent Topic Model for Multiple Text Streams. *KDD'11*, August 21–24, 2011, San Diego, California, USA.

[Hubmann, 2009] Alexander Hubmann-Haidvogel, Arno Scharl, and Albert Weichselbraun. 2009. Multiple coordinated views for searching and navigating Web content repositories. *Inf. Sci.* 179, 12 (May 2009), 1813-1821.

[Krishnan, 2007] Krishnan, M., Bohn, S., Cowley, W., Crow, V., and Nieplocha, J. 2007. Scalable visual analytics of massive textual datasets. 21st IEEE International Parallel and Distributed Processing Symposium. IEEE Computer Society.

[Lin, 2011] Yu-Ru Lin, James P. Bagrow, David Lazer. 2011. More Voices Than Ever? Quantifying Media Bias in Networks. In: *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, 2011.

[Marcus, 2011] Adam Marcus, Michael S. Bernstein, Osama Badar, David R. Karger, Samuel Madden, and Robert C. Miller. 2011. Twitinfo: aggregating and visualizing microblogs for event exploration. In *Proceedings of the 2011 annual conference on Human factors in computing systems* (CHI '11). ACM, New York, NY, USA, 227-236.

[Meyer, 2011] B. Meyer, K. Bryan, Y. Santos, Beomjin Kim. 2011. TwitterReporter: Breaking News Detection and Visualization through the Geo-Tagged Twitter Network. In: *Proceedings of the ISCA 26th International Conference on Computers and Their Applications*, March 23-15, 2011, Holiday Inn Downtown-Superdome, New Orleans, Louisiana, USA. ISCA 2011, 84-89.

[Peters, 2010] M. Peters. 2010. Four Ways to Visualize voter Sentiment for the Midterm Elections. *Mashable Social Media.* http://mashable.com/2010/10/29/elections-data-visualizations/.

[Sabol, 2010] Sabol, V., Syed, K.A.A., et al. 2010. Incremental Computation of Information Landscapes for Dynamic Web Interfaces. 10th Brazilian Symposium on Human Factors in Computer Systems (IHC-2010). M.S. Silveira et al. Belo Horizonte, Brazil: Brazilian Computing Society: 205-208.

[Scharl, 2001] Scharl, A. 2001. Explanation and Exploration: Visualizing the Topology of Web Information Systems, *International Journal of Human-Computer Studies*, 55(3): 239-258.

[Shamma, 2009] David A. Shamma, Lyndon Kennedy, Elizabeth F. Churchill. 2009. Tweet the Debates: Understanding Community Annotation of Uncollected Sources. In *The first ACM SIGMM Workshop on Social Media (WSM 2009)*, October 23, 2009, Beijing, China.

[Shamma, 2010] David A. Shamma, Lyndon Kennedy, Elizabeth F. Churchill. 2010. Tweetgeist: Can the Twitter Timeline Reveal the Structure of Broadcast Events? In *ACM Conference on Computer Supported Cooperative Work (CSCW 2010)*, February 6-10, 2010, Savannah, Georgia, USA.

[Zhao, 2011] Wayne Xin Zhao, Jing Jiang, Jianshu Weng, Jing He, Ee-Peng Lim, Hongfei Yan and Xiaoming Li. 2011. Comparing Twitter and Traditional Media using Topic Models. in *Advances in Information Retrieval - 33rd European Conference on IR Research*, ECIR 2011, Dublin, Ireland, April 18-21, 2011. Proceedings. Lecture Notes in Computer Science 6611, Springer 2011, 338-349.