World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:10, No:2, 2016

Visual Text Analytics Technologies for Real-Time Big Data: Chronological Evolution and Issues

Siti Azrina B. A. Aziz, Siti Hafizah A. Hamid

Abstract**—**New approaches to analyze and visualize data stream in real-time basis is important in making a prompt decision by the decision maker. Financial market trading and surveillance, large-scale emergency response and crowd control are some example scenarios that require real-time analytic and data visualization. This situation has led to the development of techniques and tools that support humans in analyzing the source data. With the emergence of Big Data and social media, new techniques and tools are required in order to process the streaming data. Today, ranges of tools which implement some of these functionalities are available. In this paper, we present chronological evolution evaluation of technologies for supporting of real-time analytic and visualization of the data stream. Based on the past research papers published from 2002 to 2014, we gathered the general information, main techniques, challenges and open issues. The techniques for streaming text visualization are identified based on Text Visualization Browser in chronological order. This paper aims to review the evolution of streaming text visualization techniques and tools, as well as to discuss the problems and challenges for each of identified tools.

Keywords**—**Information visualization, visual analytics, text mining, visual text analytics tools, big data visualization**.**

I. Introduction

Big Data is defined as a large amount of data from different sources in huge sizes, starting from terabytes to petabytes or zettabytes [1]. For text data, it is mostly used in social media to share breaking news [2], for servers to log any issues on a timely basis and others. The texts generated by the servers are in a semi-structured format, and through the social media or news blogs, data are in an unstructured format. Since the data is more than 1 Terabyte in size and that it requires a real-time visualization, new technologies have to be implemented in order to extract the information from the data. Scenarios such as real-time decision making in financial market trading, real-time situation discovery for the large-scale emergency response, early warning for natural disaster detection and server administration and log management are some of the important scenarios that require real-time analytic and visualization.

Visual text analytics enables knowledge discovery from the text via the usage of interactive graphical representation. [3] It helps the user to further understand the information by providing semantic mapping methods to the semi-structured or unstructured text data. With the additional visualization techniques, it provides the human brain with visual pattern recognition and spatial reasoning capabilities. [1]. This technique is very useful, especially when dealing with big data. In [4], researchers observed a trend, whereby small visual analytics software companies such as *Tableau* grew exponentially, and at the same time, big software vendors such as IBM and SAP started to acquire success in producing visual analytic tools. In [5], Harger and Crossno discussed the open source visual analytics tool. However, they only managed to find few specific tools that are related to visual analytics. Therefore, they have expanded their study to include information visualization, statistical packages, and graph analysis in the survey paper. In this paper, the researchers discussed three aspects of each toolkit: development environments, visualization functions, and analysis capabilities. In [6], Šilić, A, and Bašić, BD present related areas of research, types of data collections that are visualized, technical aspects of generating visualizations and evaluation strategies.

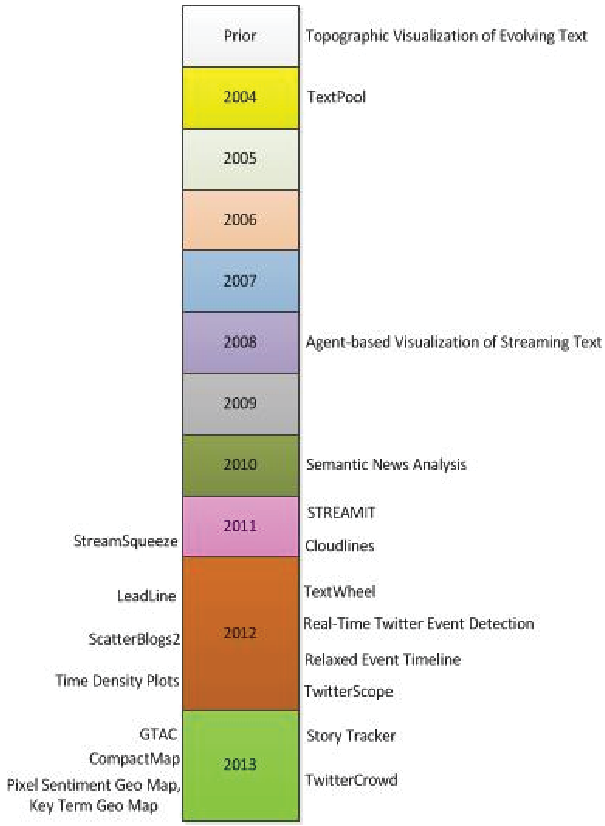

Fig. 1 Evolution of Visual Analytic technologies

Siti Azrina Abd Aziz and Siti Hafizah Ab Hamid are with Faculty of Computer Science, University of Malaya, Jalan Universiti, 50603 Kuala Lumpur, Wilayah Persekutuan Kuala Lumpur, Malaysia (e-mail: sitiazrina@siswa.um.edu.my, sitihafizah@um.edu.my)

This paper, however, does not focus specifically on real-time data processing. Rather, it focuses on all the techniques for text processing, either offline or online. In [7], Liu et al. discussed various techniques to visualize different types of data such as graph, geo-map, text and multivariate data. To the best of our knowledge, a chronological paper on visual analytic technologies focusing on real-time data stream does not exist yet.

This paper surveys real-time visual text analytic technologies for big data application. The paper then discusses the evolution of the techniques and tools. In addition, it provides the problems and challenges for each of the tools in real-time processing. The rest of the paper is structured as follows: Section 2 provides the methodology of the process. Section 3 discusses the evolution of tools and techniques, as well as the problems and challenges. Section 4 discusses the open issues and existing challenges of each technique and tools. Section 5 provides the discussion of the tools and Section 6 concludes the whole paper.

II. Methodology

In this paper, we used the Text Visualization Browser [8] to find the relevant techniques for streaming text visualization from the year 2002 to 2013. Based on the site, we filtered them based on the data processing streams and identified 20 techniques and tools for streaming text processing. Based on these techniques and tools, we did a further check on Google Scholar to find the relevant articles. We excluded 2 tools; the main reason is because the papers for these tools are not available in the Google Scholar site.

III. Evolution of Techniques and Tools

A. The Early Years

The first recognizable visual streaming text analytics technique was created in the year 2002. Topographic Visualization of Evolving Text [9] is using a probabilistic method based on latent variable models. It is an unsupervised topographic visualization of dynamically evolving, coherent textual information. The incoming data are either in the form of a stream of text or stream of words. This results in the ability to show time-coherent present in the data even though the technique does not really process text on a real-time basis.

TextPool technique and tool [10] processes text streams in real-time using Information Retrieval (IR) techniques. This tool was created to understand the “buzz” from various sources of information, such as closed-captioned cable channels, and internet-based channels, such as news feeds blogs or email. This tool presented the visualization using the graph with nodes that represent the term from the streaming text. The terms that are closely related will appear close to one another in the graph. This tool allows users to control the temporal context of how much recent stream the tool should visualize in a particular time. This method is quite straight forward, whereby the new incoming data will replace the old one. In terms of visualization, this means that the old data will be removed from the window and replaced with the newer set of data.

Agent-based Visualization of Streaming Text [11] presents visualization infrastructure that maps data element to agents. The behavior of each agent is based on the parameterized elements. This tool has three components: the visualizer, the analyzer, and the scraper. This tool can be applied to news aggregators, blogs, Twitter, Digg, and RSS feeds. However, this tool has weaknesses, such as the agent-based infrastructure is only partly utilized. Some agent properties, such as text size and color are not meaningful. The control stream filtering is only available through dialog in the scraper. Each time new data arrives in the stream, the infrastructure requires a complete refresh of the display.

Semantic News Analysis [12] uses the Europe Media Monitor(EMM) system, in which it streams the online news collected by the EMM to their system. Even though the user can search for the related articles by using EMM alone, there are two additional features that provide added values to the users – users may investigate the temporal and quantitative occurrences of entities in different languages. The borderline between historical and real-time data analysis has raised lots of issues in both textual mining and visualization research field.

B. The Recent Years

StreamSqueeze [13] presents a novel screen-filling visualization technique that focuses on the recent event. The recent event will appear in the same column and is depicted larger, with more details. However, one of the disadvantages of using this tool is that some information is lost due to the sorting and optimization requirement. Cloudlines [14] addresses the problem for multiple event-based in long time series data. Even though the screen is limited in terms of its display size and data overlapping issues, it is overcome by putting the more relevant recent data in the context of the past. This tool uses kernel density estimator; however, they acknowledge that the logic has an issue with tail regions of the distribution. STREAMIT [15] is based on the dynamic force-directed simulation into which documents are continuously inserted. It introduces the Dynamic Keyword Importance that helps to manipulate the importance of keyword for different visualization results. This tool uses the Force-Directed Placement (FDP) to visualize the dynamic document. However, this tool has an issue with visual cluttering, especially when dealing with a large document.

C. Visualizing Social Media Data

Leadline [16] is an interactive system that automatically identifies meaningful events in the news and social media data. This tool associates the topical themes with events that cause the changes. This is the major difference of the tool, as compared to the other previous tools. The key technical contributor of this tool is that it can automatically identify temporal peaks in topics as indicators of events. This tool uses Named Entity Recognition (NER) algorithm. However, the accuracy of the detected entities relies on the performance of NER. There is also plan to improve the tool by including a downward trend of an event, rather than just displaying the event’s upward trend. In 2012, STREAMIT [17] was expanded to include new features, such as topic modeling, automatic cluster discovery, and enhanced visualization. The system allows the user to manipulate the simulation speed. It also allows the user to pause the system and save cluster or document for further investigation. TextWheel [18] presents the multiple attributes of news articles and macro/micro relationship between news stream into one coherent analytical content. It makes use of other attributes of a news article, such as source, author, date and time and others, as compared to the previous approach that uses keyword-based searching and clustering of news. The data is visualized using Document Glyph. However, there is an issue with processing time if the news stream contains too many articles. Too many keywords overwhelm the keyword wheels and cause clutter in the display. To overcome this issue, it is better to use data mining technique to narrow down the size of the data. Real-time Twitter Event Detection [19] is an online method for detecting natural disasters or man-made catastrophes by analyzing Twitter Data. This tool is developed based on Cloudline. Event Visualizer [20] visualizes the data streams with logging data from a computer network. It visualizes the temporal and event of the streaming data. The strength of this approach is that it uses a combination of automatic algorithms to classify and score anomalous behavior and visual exploration. In the future, more sophisticated algorithms for burst and anomaly detection will be included.

Time Density Plot [21] uses sentiment, temporal density, and context coherent that comments on features for different targets. This tool is targeted towards business analysts who want to explore the customer feedback in company website to detect critical issues. Sentiment analysis is used to identify the feedback, either positive or negative from the customer on the company or product. This tool uses the Internet General Inquirer to classify the data. The data is visualized using pixel map calendars and time density plots.

D. Geographical Map Visualization

TwitterScope [22] was first developed in 2012. The tool groups similar messages into clusters displayed as “countries,” with keyword summaries, using semantic analysis, map generation techniques and graph clustering. In order to make the system more flexible, there is a plan to re-implement some the processes in Javascript so that the load can be shifted from the network to the local client. Using this method, it is believed that the data will be easier to understand, as compared to visualizing it using rivers of text. In 2013, TwitterScope [23] was enhanced, and a new technique called GMap was introduced. It applies the Prosecutes transformation to documents within a component and proposes a stable packing algorithm for components to achieve further stability. The tool focuses on the mental map preservation to avoid distortion of the layout for the purpose of mental stability. ScatterBlogs2 [24] provides analysts with an interactive task-tailored message filters based on the recorded messages of well-understood previous events. Analysts can use this tool to monitor customer’s interest on the certain topic. This tool uses supervised classification and query creation. The future work for this tool is to find ways of supporting the analysts to develop strategies to create beneficial filter combinations more quickly or semi-automatically. Story Tracker[25] uses the text mining technique to facilitate the analysis of similar topics from the news that split and merge over time. It uses a hybrid technique which combines the advantages of ordered list and ThemeRiver-like visualization. The tool visualizes the temporal and topical aspect of news. GTAC [26] extracts a structured representation of events from the Twitter streams. The tool integrates MCL clustering, Named Entity Recognition (NER) and NLP techniques to analyze the data. The data will be visualized using geographical and temporal patterns. However, since this tool is using NER algorithm, the accuracy of detected entities relies on the performance of the algorithm. TwitterCrowds [27] extends the description of ThemeCrowds and SentireCrowds visualization interface and introduces list equivalents of both. ThemeCrowds and SentireCrowds support the visualization of hierarchical clusters. The visualization systems take time series of multilevel clustering of Twitter users as input. For future work, the tool will adapt techniques to reduce the processing time from data collection to visualization. CompactMap [28] is an online visual interface that packs text clusters efficiently with stable updates to maintain the user’s mental map. The tool automatically retrieves tweets and summarizes topics for a user-specified search query, and visualizes the results in a stable view. This allows the users to track the topics of discussion over time. It achieves the spatiotemporal coherent layouts by dynamically matching clusters across time and removing cluster overlaps according to spatial proximity and constraints. This tool has been tested for Twitter data. Pixel Sentiment Geo Map and Key Term Geo Map [29] combine the feature-based sentiment and geo-term association to enable store managers to analyze web survey feedback. This tool uses survey feedback and tweet data. The tool uses the feature-based algorithm, sentiment analysis, sentence-based term, key term geo map, key term distribution map, pixel sentiment calendar and SOM that cluster nodes into related clusters. The future plan is for the tool to be able to detect geo-temporal sentiment patterns, trends, and influences in the customer feedback streams for the live alert.

IV. Discussion

We compared 18 techniques and tools for real-time visual analytic tools with respect to a set of 4 criteria. Based on the history of visual analytic technique and tool, below are important checkpoints in the evolution process.

A. Applied Domain

Based on the study, in the early years, the focus of the techniques and tools was mostly on news stream analysis. The source data is mostly from a news site or RSS feed. The motivation of having the tools is a large number of documents or news available in the news site. The tools will generally help the users to understand the time coherent of the topic in order to help the users to understand the news evolvement. However, with the rise of social network site, such as Facebook in the year 2005 and Twitter in 2006, there is a large amount of text data created on the Internet. That is the reason that in the recent years (the year 2006 onwards), more and more techniques and tools were created to perform analysis on the social media, especially Twitter. Based on this study, 70% of the total number of tools and techniques are created in the year 2012 onwards, and the focus is on social media. Other than social media, there are other domains, such as system log and web survey. System log domain is beneficial for the user to monitor the health of their overall system performance. However, there are only two out of seventeen tools or techniques that focus on this area. Web survey has also just recently become a focus in the domain. This domain is important for the company to understand customer's feedback and monitor the sentiment towards their product. Therefore, for future research, system log and web survey are two domains that can be further explored by researchers.

TABLE I
CHARACTERISTICS OF THE VISUALIZATION TOOL AND TECHNIQUE

Tool/techniques	Applied domain	Visualization method/technique	Visualization type	Analysis technique
Topographic Visualization of Evolving Text	News stream.	Time coherent - topic detection	2D topographic	Vector space model, latent variable model
TextPool	News feed, email, blog.	Topic based	Text collage	Information Retrieval (IR) technique.
Skimmer	News, blogs, Twitter, Digger, and RSS feeds	Frequency of words	Circle size by frequency of word occurrence	Clustering
Semantic News Analysis.	News stream.	Temporal events	Radial tree layout	Transform incoming data to internal XML format
StreamSqueeze	News stream and system log monitoring	Recent item will be shown larger	Screen-filling visualization	Clustering
Cloudlines	News stream.	Event episode in time series, sentiment and context coherence	Circular objects	Kernel density estimator
STREAMIT	News stream.	Document collection	Dynamic moving particles in 2D	Special keyword importance . Force-Directed Placement. Use GPU hardware acceleration
LeadLine	News stream, Twitter	Topical themes for event	Flow-like visual metaphor	Named entity recognition algorithm. Early event detection technique
TextWheel	News stream.	Keyword-based searching	Dynamic text glyph	Text mining, sentiment, word frequency
Event Visualizer	System log.	Event detection	Relaxed timeline	User defined rules
Time Density Plots	Web survey, RSS news feed	Temporal sentiment patterns.	Pixel map calendars, sequential time density plots	Internet General Inquirer
TwitterScope – Gmap	Twitter data	Keyword summary, semantic analysis	Dynamic map metaphor technique	Prosecute transformation and projection, graph packing algorithm
ScatterBlogs2	Real-time microblog	Train test data	Geomap, topic view	LDA topic view, Content lens
StoryTracker	News stream	Topic detection	The hybrid technique, which combines the advantages of connected ordered lists and ThemeRiver-like visualizations.	Document clustering
LeadLine	Twitter data	Event detection	Geospatial, entity and temporal	MCL clustering, NER and NLP technique
TwitterCrowds	Twitter data	Topic and Sentiment Analysis	Tag clouds	Clustering
CompactMap	Twitter data	Topic detection	Space-filling metaphor	Clustering
OpinionFlow	Web survey and twitter data	Key term and Sentiment analysis.	Geomap, temporal	Feature-based algorithm, sentiment analysis, sentence based term, Kernel density estimator

B. Visualization Methods or Techniques

Based on this study, there are a few categories of visualization methods or techniques that can be identified. Time-series techniques are the common techniques among the tools, such as time detection, topic detection, and event detection. All these techniques use the time attributes in order to classify the data. The latest technique is a semantic analysis that is used to identify the behavior of public users toward certain issues. Semantic analysis is important for companies and government agencies to understand their product's or service's satisfaction further. Also, there are general techniques being used by the tools, such as frequency of words and document collection. This technique mainly uses the source text itself to identify the pattern of the data.

C. Visualization Type

Based on this study, in the early years, the visualization type is quite simple. Visualization algorithms are being applied to the analyzed data to determine the distance between the data. The similar data will basically appear closer than the dissimilar one. Next, the data will be plotted on the screen. This approach is quite simple and straight forward. However, it does not provide much information to the users. Techniques such as text collage, screen-filling visualization, and map metaphor have an issue with visual cluttering, especially when it involves a large amount of document. Some tools also use animation technique that can help to display historical data up to the latest event. In the recent years, researchers include geo map visualization to provide better insights to the users.

D. Analysis Technique

Analysis technique is another important aspect for visualization tools or techniques. From this study, we can see that the analysis techniques are related to the visualization methods or techniques. Most of these techniques are text mining techniques, such as clustering and named entity recognition.

V. Open Issues and Challenges

A. Visual Clutter

Since the incoming data comes on a real-time basis, the data visualization needs to be updated when any new data arrives. However, the tool needs to ensure that the new data can be visualized in the context of old data. The limited screen also causes high overlapping of data. Visual clutter can decrease the user satisfaction of using a particular tool.

B. Response Time

For Big Data, since the data comes in high volume, it needs an effective algorithm that can perform a fast analysis so that there will be no obvious delay in the data visualization. Fast response time is highly required for the real-time visual analytic tool to ensure the analyzed data is always up-to-date. This will affect the decision or surveillance result for the users.

C. Accuracy of Visualized Data

Before any data can be visualized, it will go through the process of analysis. The accuracy of the visualized data mainly depends on the algorithm, such as named entity recognition (NER) [16] and placement algorithm [22]. Kernel density estimator also has an issue with the tail regions of the distribution. Sorting and optimizing data also cause loss of information. Since most scenario that requires real-time analytic are critical scenarios, having an accurate data will also affect the analytics of the data.

VI. Conclusion

In this paper, we discussed the visual text analytic techniques and tools for real-time scenarios that are related to the Big Data. Based on our study, the tools can be divided into early years and recent years. Early years tools were mostly built on the existing offline data stream techniques. This is because, during early years, there is not much importance in having real-time analytic tools. With the emergence of big data and scenarios such as financial trading, situational awareness and surveillance that requires prompt decision and early warnings, more tools were created since the year 2012. Open issues such as limited screen space for visualization, response time and accuracy of visualized data are some of the common issues for the recent tools.

References

[1] P. Chandarana and M. Vijayalakshmi, "Big Data analytics frameworks," *2014 Int. Conf. Circuits, Syst. Commun. Inf. Technol. Appl.*, pp. 430–434, Apr. 2014.

[2] P. O. Box and A. Ain, "Real-Time Big Data Analytics : Applications and Challenges," pp. 305–310, 2014.

[3] C. Rohrdantz, D. Oelke, M. Krstajic, and F. Fischer, *Real-time visualization of streaming text data: tasks and challenges*, no. October. 2011.

[4] L. Zhang, A. Stoffel, M. Behrisch, S. Mittelstadt, T. Schreck, R. Pompl, S. Weber, H. Last, and D. Keim, "Visual analytics for the big data era—A comparative review of state-of-the-art commercial systems," in *Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on*, 2012, pp. 173–182.

[5] J. R. Harger and P. J. Crossno, "Comparison of open-source visual analytics toolkits," in *IS&T/SPIE Electronic Imaging*, 2012, p. 82940E–82940E–10.

[6] A. Šilić and B. Bašić, "Visualization of text streams: A survey," *Knowledge-Based Intell. Inf. …*, pp. 1–12, 2010.

[7] S. Liu, W. Cui, Y. Wu, and M. Liu, "A survey on information visualization: recent advances and challenges," *Vis. Comput.*, Jan. 2014.

[8] K. Kucher and A. Kerren, "Text Visualization Browser: A Visual Survey of Text Visualization Techniques," *Poster Abstr. IEEE VIS 2014*, 2014.

[9] A. Kabán and M. A. Girolami, "A Dynamic Probabilistic Model to Visualise Topic Evolution in Text Streams," *J. Intell. Inf. Syst.*, vol. 18, no. 2–3, pp. 107–125, Mar. 2002.

[10] C. Albrecht-Buehler, B. Watson, and D. A. Shamma, "TextPool: Visualizing Live Text Streams," *IEEE Symp. Inf. Vis.*, 2004.

[11] J. R. Benson, P. Lafleur, D. Crist, B. Watson, and N. Carolina, "Agent-based Visualization of Streaming Text," *Inf. Vis. Conf.*, 2008.

[12] M. Krstajic, F. Mansmann, A. Stoffel, M. Atkinson, and D. a. Keim, "Processing online news streams for large-scale semantic analysis," *2010 IEEE 26th Int. Conf. Data Eng. Work. (ICDEW 2010)*, pp. 215–220, 2010.

[13] F. Mansmann, M. Krstajic, F. Fischer, and E. Bertini, "StreamSqueeze: a dynamic stream visualization for monitoring of event data," p. 829404, Dec. 2011.

[14] M. Krstajić, E. Bertini, and D. A. Keim, "CloudLines: compact display of event episodes in multiple time-series.," *IEEE Trans. Vis. Comput. Graph.*, vol. 17, no. 12, pp. 2432–9, Dec. 2011.

[15] J. Alsakran, Y. Chen, Y. Zhao, J. Yang, and D. Luo, "STREAMIT: Dynamic visualization and interactive exploration of text streams," *2011 IEEE Pacific Vis. Symp.*, pp. 131–138, Mar. 2011.

[16] W. Dou, X. Wang, D. Skau, W. Ribarsky, and M. X. Zhou, "LeadLine: Interactive visual analysis of text data through event identification and exploration," *2012 IEEE Conf. Vis. Anal. Sci. Technol.*, pp. 93–102, Oct. 2012.

[17] J. Alsakran, Y. Chen, D. Luo, Y. Zhao, J. Yang, W. Dou, and S. Liu, "Real-time visualization of streaming text with a force-based dynamic system.," *IEEE Comput. Graph. Appl.*, vol. 32, no. 1, pp. 34–45, 2012.

[18] W. Cui, H. Qu, H. Zhou, W. Zhang, and S. Skiena, "Watch the Story Unfold with TextWheel," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 2, pp. 1–17, Feb. 2012.

[19] M. Krstajic, C. Rohrdantz, M. Hund, and A. Weiler, "Getting there first: Real-time detection of real-world incidents on twitter," pp. 2–5, 2012.

[20] F. Fischer, F. Mansmann, and D. A. Keim, "Real-time visual analytics for event data streams," in *Proceedings of the 27th Annual ACM Symposium on Applied Computing - SAC '12*, 2012, p. 801.

[21] C. Rohrdantz, M. C. Hao, U. Dayal, L.-E. Haug, and D. A. Keim, "Feature-Based Visual Sentiment Analysis of Text Document Streams," *ACM Trans. Intell. Syst. Technol.*, vol. 3, no. 2, pp. 1–25, Feb. 2012.

[22] E. R. Gansner, Y. Hu, and S. C. North, "Interactive Visualization of Streaming Text Data with Dynamic Maps," *J. Graph Algorithms Appl.*, vol. 17, no. 4, pp. 515–540, 2013.

[23] E. Gansner, Y. Hu, and S. North, "Visualizing Streaming Text Data with Dynamic Maps," *arXiv Prepr. arXiv1206.3980*, Jun. 2012.

[24] H. Bosch, D. Thom, F. Heimerl, E. Püttmann, S. Koch, R. Krüger, M. Wörner, and T. Ertl, "ScatterBlogs2: real-time monitoring of microblog messages through user-guided filtering.," *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 12, pp. 2022–31, Dec. 2013.

[25] M. Krstajic, M. Najm-Araghi, F. Mansmann, and D. A. Keim, "Story Tracker: Incremental visual text analytics of news story development," *Inf. Vis.*, vol. 12, pp. 308–323, 2013.

[26] T. Kraft, D. X. Wang, J. Delawder, W. Dou, and W. Ribarsky, "Less After-the-Fact: Investigative visual analysis of events from streaming twitter," in *2013 IEEE Symposium on Large-Scale Data Analysis and Visualization (LDAV)*, 2013, pp. 95–103.

[27] D. Archambault, D. Greene, and P. Cunningham, "TwitterCrowds: Techniques for Exploring Topic and Sentiment in Microblogging Data," pp. 1–19, Jun. 2013.

[28] X. Liu, Y. Hu, S. North, and H.-W. Shen, “CompactMap: A mental map preserving visual interface for streaming text data,” in *Big Data, 2013 IEEE International Conference on*, 2013, pp. 48–55.

[29] M. C. Hao, C. Rohrdantz, H. Janetzko, D. A. Keim, U. Dayal, L. e. Haug, M. Hsu, and F. Stoffel, “Visual sentiment analysis of customer feedback streams using geo-temporal term associations,” *Inf. Vis.*, vol. 12, no. 3–4, pp. 273–290, Jun. 2013.