FinaVistory: Using Narrative Visualization to Explain Social and Economic Relationships in Financial News

Yeuk-Yin CHAN and Huamin QU
Department of Computer Science and Engineering
The Hong Kong University of Science and Technology
yychanae@connect.ust.hk, huamin@cse.ust.hk

Abstract—When reading financial news, although there are critics explaining the fluctuation of Economic indexes in articles everyday, the news often assert bias on authors' favorite opinions. On the other hand, the amount of financial news published these days is staggering with diversified opinions on the same issues. Unless with acumen for the whole environment, audience will find hard to stay objective and identify useful information from the mass media. If computer is granted ability to analyze all news and generate a narrative that addresses all concerns related to the news, all kinds of readers can obtain a more compelling story. Nowadays, Narrative Visualization is a popular technique for news media to deliver reader driven stories, and a potential field for application side. In this paper, we first identify financial news with two features: polarity associated with each news and, factors impacting the indexes or being impacted by the indexes. Then we apply these features to explain the price fluctuation and explain the Socio-economic relationships between price and different topics. Finally, two case studies are conducted to demonstrate the effectiveness and usefulness of FinaVistory in helping users with different purposes to understand the stories behind Europe unemployment after 2008 financial crisis and oil price in 2014.

Index Terms—Narrative Visualization, FinaVistory, Social and Economic Effects, Text Mining

I. Introduction

Data visualization techniques nowadays provide mainstream news organizations new ways to deliver journalism information from large amount of data. Text and statistics data now can be turned into diagrams and charts to convey messages. On top of that, the goal of visualization is to provide a backdrop of stories, as well as to provide powerful vehicle for discovering stories [1].

At the same time, financial analysts research macroeconomic and microeconomic conditions along with company fundamentals to make business, sector and industry recommendations. When they receive information from the financial news they also need to make judgment on different opinions in order to stay objective on the issue [2].

In the field of data visualization, there are visualization techniques to demonstrate a good use for satisfying a single purpose such as visualizing frauds, explaining market volatility or detecting trader sentiments. However, in case of Economic research on financial news, the purposes are much more versatile and open end. For example, economists may want to find out the economic relationships between oil price and inflation while some others may want to find out the political causes affecting oil price. Visual analytics here are meant to bring out a whole story rather than just giving explanation to a particular concern. Therefore, narrative visualization plays a good role in this area. Once a general intended narrative is complete, the visualization opens up to a reader-driven stage where the user is free to interactively explore the data, thus fulfill their individual purposes [1].

In this paper, we identify the key settings of financial news: reasons for Economic index changes and impacts on/from it. To address these settings we classify each news article into three attributes: positive news (news about index increase), negative news (news about index decrease) and impact news (news about impacts on/from index change). Then we generate several visualization results across the narrative based on the classifications to address the concerns for the key settings. To verify our results, we incorporate economic equations and actual outcome to tell the stories that people are most interested in around the main topics in our demonstration.

In summary, the major contributions of this paper are:

1. Introduce a way of standard and automatic narrative visualization for satisfying the understanding of financial and economic information and relationship with different motivations included but not limited to investing, economic and social purposes, without the need of human processing unlike the current practice.

2. Identify the attributes of financial news and methodologies to extract useful information from these attributes.

3. Examine popular main topics in the field of Economics and Finance under a simple work flow to emphasize simplicity in analyzing complex environment.

II. Related Work

In general, current research works have demonstrated various visualization techniques to visualize financial information. For example, a bubble motion chart and radar word frequency chart are used to uncover relationships between news headlines that can convey meaning about trader sentiment, and their influence on stock prices [3]. WireVis, a visualization system

Fig. 1. Sample interface of FinaVistory visualizing oil price news in 2014 and EU Unemployment during 2008-14. Apart from the main theme showing the percentage difference and time domain of the price, users can explore different reasons behind the fluctuation, impacts on or from the change and relationships with different Economic indicators.

of categorical, time-varying data from financial transactions has been used by Bank of America to detect transaction frauds [4]. Moreover, financial news topics generated by topic modeling can be used to explain market volatility [5]. While these research works focus on delivering insights for particular financial activities, our visualization work focuses on giving more versatile insights including but not limited to social, economic and financial scenarios.

To be more specific, there are research works focusing on visualizing news articles, story telling on financial information and text visualization.

A. News Visualization

Previous work has done for visualization of large amount of new stories. Examples are Contexter [6], a system using manual extraction rules (i.e. extracting capitalized words) and statistics (i.e. comparing shared documents) to visualize a network of news topics, Newsmap, a hierarchical knowledge map generated by SOM categorization can be used as a tool for browsing business intelligence and medical knowledge hidden in news articles [7], and TextWheel , which uses a combination of keyword wheels, document transportation belt, and dynamic glyphs to convey messages behind the dynamic natures of news streams [8]. The approaches demonstrated are price-performance effective and are good examples to emphasize the importance of data preprocessing. Compared with these research works, our belief on financial news is that they contain unique attributes apart from general news. Thus our works aim at providing tailor-made methods for financial news processing and visual representations.

B. Narratives and Financial Activities

Based on Efficient Market Hypothesis, stock prices rapidly and perfectly incorporate financial information in a short time. Therefore, what actually causes the discourse of market

are more probably about behavioral finance, in which how media frame the investment decisions can condition financial decisions [9]. Studies have also been done on the effects of financial communications and conceptualize the effects in terms of impression management [10]. Thus narratives play an important role in directing investor sentiment, consequently affecting the volatility and expectation in financial market also. These works, however, focuses on exploring the narrative components in financial scenarios only, in which to make people better understand the meanings behind, we explore the potentials of narrative visualization in our work.

C. Data Visualization on Relationship and Keyword

For visualizing relationships or specific topics on a time, many tools have been developed. Wordle [11] [12] and Tag Cloud [13] are popular tools visualizing the significance of certain keywords by using different front sizes, colors, etc. On the other hand, force-directed layout [14] can be used to visualize modularity clusters [15]. However, since these individual tools are only generalized to cope with a broad usage of visualization, we now aim at providing a comprehensive sets of visual presentation to let users understand the matters in a holistic way.

III. OVERVIEW

A. Research Problems

Our motivation is that, there is little attention focusing on combining all the above-mentioned research field. However, there are already plenty of financial findings, well-defined narrative structures and visual tools available nowadays, and researchers are looking at the opportunities to extend visual analysis applications to enable storytelling too [1]. Thus we combine the knowledge of these fields to design a system to answer questions that can be possibly arisen from fluctuations

of the Economic indexes. In particularly the domain of narrative visualization, we are answering these questions:

1. How to establish a narrative flow of visualizations to allow multi-messaging from different ordering and combinations of features in the system?

2. What is the essential information that needs to be extracted from financial news to explain all the socio-economic problems and concerns?

B. Our System

We developed a visualization system, FinaVistory, to address our research problems. The interfaces can be seen in the above Fig. 1 using Oil Price in 2014 and EU Unemployment after Financial Crisis as examples.

(i) Header "...Overview" indicating the main topic and the time frame within the analysis. These give an overview for user of what they are investigating and the respective timeline.

(ii) Stacked bar chart showing the time distribution of positive and negative news. Positive news means the article conveys a message of increasing index while negative news means the opposite. From the bar chart, users can understand two things: the proportion of positive and negative news and the total volume of news in a time. With the aid of time series of price index on the right, users can understand the relationship between price fluctuation and news volume and positive or negative news.

(iii) A Word Cloud and display of top 10 most relevant news. These serve the purpose of giving more insights to the user. The font size of keywords represents their appearance frequency within the news documents during that given period. The color of the word is classified based on its relative frequency among the 10 articles. If it has the highest proportion on one article its color will belong to the one assigned to that article. When user clicks on the keyword, the system will direct the user to a new page of Google News search with keywords consisted of the price and the word within the selected time period, so that user can understand the news related to the keyword and the price at that time.

Fig. 2. Network of Topics of Oil Price. Links are shown when user hovers to a specific node. Closer nodes belong to a more complex network.

(iv) A network visualization of impacts using force-directed layout diagram (Fig.2) with pie charts summarizing the statistics of the nodes. The nodes represent the topics generated from topic modeling of all impact articles (news mentioning the impact on or from the price change), with links appeared when users hover on the specific nodes, showing the similarity between topics. A more vivid line means a higher similarity

between the nodes. Nodes with more complex network will be shown in a more inside cluster in the diagram too.
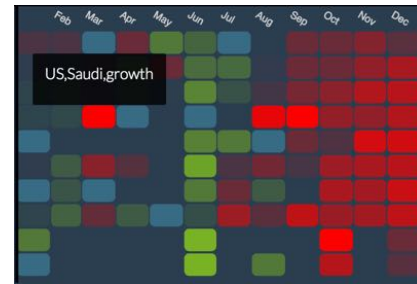

Fig. 3. Summary Plot showing the polarity of news retrieved from the topics user chooses across different time with topic shown when user hovers over a block.

(v) A summary plot showing the relevance of topics to price change in different time. (Fig.3) Users can observe the time evolution of topics in relation to the change of price. Color intensity represents the relevance. This feature aims at giving users an aggregated overview of relationships between social topics and the index that cannot be presented by quantitative means.
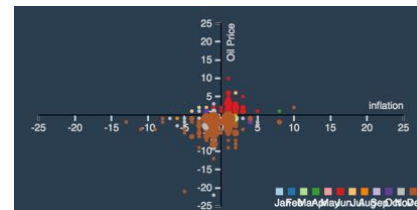

Fig. 4. Bubble chart showing the polarity of opinions on a user-chosen topic alongside with the main topic (oil price) across different time.

(vi) A bubble chart (Fig.4) showing the relationship of user input variable and the price. To create a better flexibility for users and to leverage users' domain knowledge, the chart allows user to input their own variables, then the system will search in the text, compute the polarity of the variables and draw out the relationship between it and the price. Different colors represent different time and different sizes represent number of points overlapped in the coordinates.

IV. DESIGN AND IMPLEMENTATION

A. Narrative Structure

The system narrative structure design focuses on utilizing the under-utilization areas of narrative visualization found from the findings from the design space analysis of narrative visualization [1]. Firstly, it consists of ordering strategies corresponding to distinct genres of visual narration. The ways how the system arranges its components is able to construct a clear flow of logic to understand relationships. Secondly, the consistent interactive designs used by visualizations allows sets of visual techniques be reused in different scenaio and topics. Apart from that, it focuses on the utilization of strategies

to engage the user in the interactive functionality and common storytelling techniques across narrative visualizations.

To be specific, FinaVistory adapts the following attributes:

1. Random Access: There is no path suggested at all for guiding users to browse each component. For example, users can first have a look of the price change, then understand the reasons from the keywords in respective time, and lastly see the impacts of the change from the network. Or else users can also first select a couple of social topics from the network, then visualize them in the summary plot to see the relationship between them and the relevance of the price change, then compare the news with their Economics knowledge through the bubble chart.

2. Standard interactivity: Like other visualization systems, FinaVistory provides highlighting, details filtering, selection, search navigation buttons and explicit instruction for each graphs and charts.

3. Messaging: One area of under-utilization of narrative structure is the imbalanced proportion of texts and graphs. To cope with the problem, FinaVistory provides accompanying article, multi-messaging, introductory text, synthesis and visualization description to aid storytelling. For example, users are first introduced the overview of price changes under the header, and then select respective word cloud to understand the correct and wrong reasons for changes and browse their most relevant news below. Also each graph or chart is given an instructive text around it.

B. Extracting Useful Information from Financial News: News Classification and Topic Modeling

The main purpose of classifying news into relevant categories is to screen out irrelevant articles that will possibly distort development of results in the system. For oil price we import around 7000 articles in 2014 containing 'oil price' as keyword and for unemployment around 6900 articles in the years 2009-2014 containing 'Europe' and 'unemployment' from ProQuest database, in which the publishers include but are not limited to The Financial Times, The Wall Street Journal, The Guardian, and New York Times.

Since these news corpus only mean to have contained the keyword in the article, in order to find out whether a news article has mentioned the reasons about price change, the system calculates a increase/decrease semantic value for each article. It extracts three words before and after the keywords in the articles, and plus one for each verb means 'increase' or minus one for those meaning 'decrease', determined by the auxiliary list in Table 1.

TABLE I
INCREASE/ DECREASE CALCULATION AUXILIARY LIST

Increase	Decrease
benefit, boost, climb, gain, growth, heating, help, high, hike, improve, increase, jump, positive, push, raise, rebound, recover, rise, soar, spike, spur, strong, support, surge	collapse, crash, cut, decline, drop, fall, low, plummet, plunge, slide, slump, slow, tumble, weak

We choose three words only because of the English Grammar 'Subjective-Verb-Objective" sentence structure that, if a verb appear within three word distance of the keyword, it must be related to the keyword, and if a noun appears within this distance, they are related to each other also being as compound noun.

Therefore in short, for each article, the indicator function $P_i(t)$ for polarity value for article t in regard to the main topic i would be:

$$P_i(t) = N(p, i, t) - N(n, i, t) \qquad (1)$$

,where $N(p,i,t)$ is the number of 'positive' words that appears in three words around the main topic i mentioned in the article t and $N(n,i,t)$ is the number of 'negative' words that appears in three words around the same main topic in the same article. The polarity of the indicator function represents whether the article is towards a 'positive' or 'negative' stance and the absolute value means the degree of the polarity.

After the above calculation, we then remove all the articles that the polarity value is equal to 0, in which less than half of articles are remained for our cases. For the remaining associated news records collected, we decompose them into thematic features that reveal the issues related to the topic. This is done by applying a simple topic modeling technique called Latent Dirichlet Allocation (LDA), a generative probabilistic model for collections of discrete data such as text corpora [16]. It is the simplest topic model and uses the Dirichlet prior in order to ensure sparsity in the underlying multinomial distribution [5]. And since the results from the topic model is excellent, we did not find a need to apply a more sophisticated topic model. The system removes common stop words from the classified articles, runs LDA by setting the number of topics to 100 and keyword to 3, and combine the topics with same keywords but different ordering.

C. Data Visualization Techniques

Fig. 5. Word Cloud example. Larger word size mean a higher frequency of appearance. Words with same color come from same passage also.

1) Word Cloud: The word cloud presentation is generated by a open source HTML5 Word Cloud [17]. It assigns sizes for each word based on their frequencies and try to find a place to fit the words in the canvas. In order to utilize the color visualization we assign colors to relationships between keywords. Colors represent each 10 passages being analyzed by Word Cloud, and the keyword belongs to a particular one if its relative frequencies to that is the highest among the passages.

2) Network Visualization: The network visualization of topics is generated by force directed layout diagram. Once the forces on the nodes and edges of a graph have been defined, the forces are applied to the nodes pushing them further apart and the tension on the edges pull them together. This is repeated iteratively until the system comes to a mechanical equilibrium state. We use this method to visualize the similarity between topics generated by LDA. As LDA does not give out relationships between topics, we use their co-occurrence in news corpus C_{uv} as similarity measure, which is:

$$C_{uv} = max(\frac{\{U_1, U_2, ...U_{N_u}\} \cap \{V_1, V_2, ...V_{N_v}\}}{\{U_1, U_2, ...U_{N_u}\}},$$
$$\frac{\{U_1, U_2, ...U_{N_u}\} \cap \{V_1, V_2, ...V_{N_v}\}}{\{V_1, V_2, ...V_{N_v}\}}) \quad (2)$$

, where each *U* represents the document that contains topic u and each *V* represents the document that contains topic v. Intuitively, if a topic appears in many documents that also include another topic, two topics can be highly correlated.

The co-occurrence will be visualized as the link's distance and strength in the network, as shown Fig. 6:

Fig. 6. Overview of Topic Network for oil price in 2014 with some of the topics highlighted.

3) Summary Plot: Summary Plot is used to group the 'positive' and 'negative' news into relevant topics generated by LDA across the time. As each article contains a polarity value, each block in the summary plot reveals the average polarity of the collection of articles related to a particular topic in the specific month. As shown in Fig. 3, a bright red block for 'US, Saudi, Growth' means this topic is highly relevant to oil price decrease in March.

4) Bubble Chart: The bubble chart provides option for visualizing the relationship between keywords and main topic in which the keywords have polarity implication also. User can type in their own keyword and the system will calculate its polarity from the classified news records based on equation (1) which is used to calculate the value of the main topic as well. The calculation will result in a set of coordinates representing $(P_k(t), P_i(t))$ for each document t. We then combine and increase the size of each overlapping coordinate so the visualization result will look like the one in Fig. 4.

V. CASE STUDIES

In order to investigate in the usefulness of the system, we use it to analyze oil price, one of the hottest finance topics in 2014, and unemployment in Europe after 2008 Financial Crisis, one of the popular topics in Economics research field.
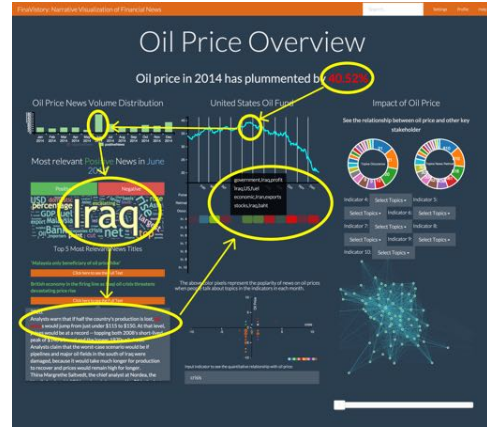
A. Visual Narratives of Oil Price in 2014

Fig. 7. Narrative flow for geopolitical cause of price drop

1) Social Perspective: Geopolitical Cause of Price Drop: When user examines the time series of oil price, they can see oil price started to plummet in June, and when they check the news in that month, messages from Word Cloud show that "Iraq" is the most popular topic about 'increasing oil price' in that month. Then by future browsing the relevant news articles we can see sentences like "Analysts warn that if half the country's production is lost, oil prices would jump from just under \$115 to \$150." or "The success of the insurrection in Iraq has changed that. Crude oil prices have surged to their highest in nine months." , in which it further proves that geopolitical risk has long been pushing the oil price at that time.

However, when we put the these kinds of topics to the summary plot, we can see a dramatic change of polarity started in August, meaning the public stance switch in the belief towards geopolitical risks. It became apparent that supply disruptions from conflict in the Middle East had unwound, or did not materialize as expected.
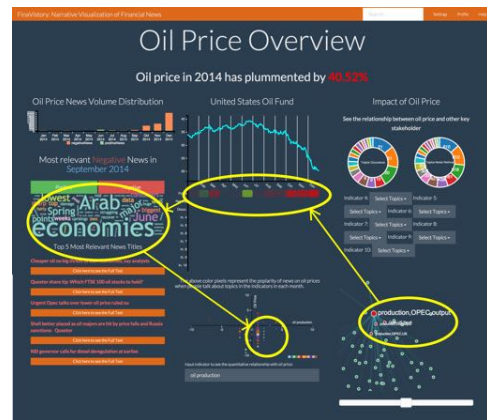

Fig. 8. Narrative flow for economic cause of price drop

2) Economics Perspective: Cause of Price Drop: Demand and Supply: For commodity pricing economists generally

would like to know the reasons in terms of supply and demand. A little background for oil pricing is that most of the time it is coordinated by Organization of the Petroleum Exporting Countries (OPEC). To get a deeper look from this angle we can examine the topic network to see its related parties. From the network we see a small cluster of topics {(production, OPEC, US),(OPEC, US, Saudi),(production, Saudi, economy)} etc., which give us hints that OPEC's issue is related to US and Saudi Arabia oil production. We can then put all these topics to the summary plot. Notice that these topics shows a relatively strong negative signal in October, November and December, thus we can suspect that the issues about dropping oil prices happened then. Furthermore, when we type "oil production" in the bubble chart, we can also see a tendency of "increasing oil production, decreasing oil price" in these three months.

To sum up, Saudi Arabia and US maintained a high level of oil production and OPEC failed to coordinate the balanced production in October, November and December, thus the increase of oil production decreased the oil price.
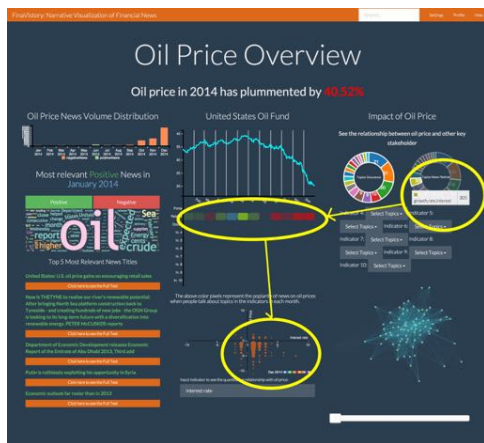

Fig. 9. Narrative flow for financial implication of price drop

3) Financial Implication: Monetary and fiscal policies: In the donut chart we can notice a topic of 'growth, interest, rate', which gives us a hint that oil price fluctuation had affected economic growth that in which the interest rate as well. By putting it in the bubble chart it is noticed that interest rate had a strong tendency of increasing along with decreasing oil price in late 2014. In some examples in the top relevant news, we can see statment like *"...... expected interest-rate hikes to come in the third quarter; they have now pushed out that estimate to the fourth quarter of 2015, largely because falling oil prices will keep inflation soft......"* , which gives an idea that since falling oil price exert a reduce inflation growth, central banks could respond with additional monetary policy loosening, which can support growth.
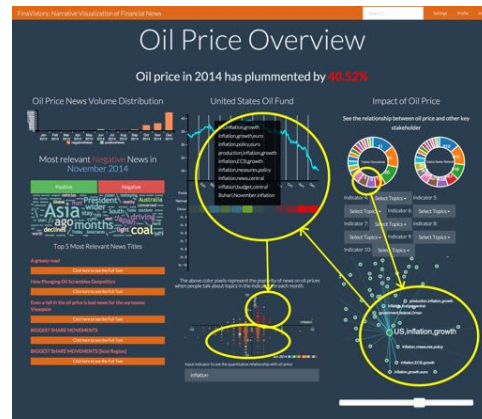

Fig. 10. Narrative flow for economic implication of price drop

4) Economic Implication: Decreasing Impact of Inflation: There is a historical relationship that oil price swings and inflation have been positively correlated, but it is an empirical observation that the extent and impact are changing over time. When we check all the topics are related to inflation in the summary plot, we can see that people most mention it with oil price change only in June, November and December, and with advanced economies like US and Europe as well. In addition, when we plot the relationship between inflation and oil price in the bubble chart, we can see that only the result in June satisfied the historical relationship. This verify the recent research that oil's impact on activities has been declining in over the years [18] [19][20].

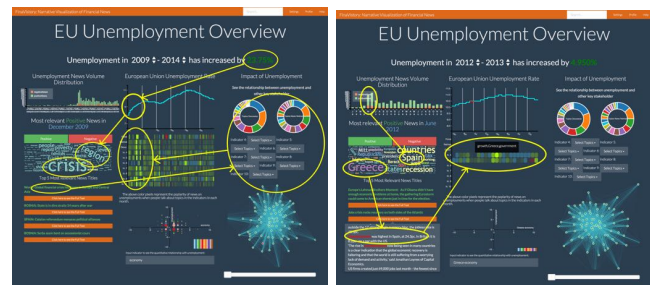*B. Visual Narratives of EU Unemployment in 2008-14*

Fig. 11. Narrative flow for social cause of EU unemployment

1) Social Perspective: Overall Rise after Financial Crisis and Staggering Unemployment in Greece and Spain in 2012: From the time series of unemployment we can observe an increasing trend in 2009, 2012 and 2013. Looking at the news in 2009 we can receive the impression of recession coming from Financial Crisis, and by plotting all keywords about Europe economy we can find out that the adverse effects of crisis mainly happened in 2009. On the other hand, if we focus on the years of 2012 and 2013, we observe a peak of 'increasing unemployment' news around the late 2012. And from the word cloud we can recognize keywords like "Greece", "Spain", "euro" and "crisis", giving a hint that these countries had not rebounded or even become worse in these situations.

The statements in news such as *"Euro-area unemployment reached the highest on record as a deepening economic slump and budget cuts prompted companies from Spain to Italy to reduce their workforces."* or *"Spiraling unemployment, biting austerity measures and political uncertainty have led to an upsurge in Greeks quitting the country for sunnier economic climes."* can further consolidate this kind of beliefs.
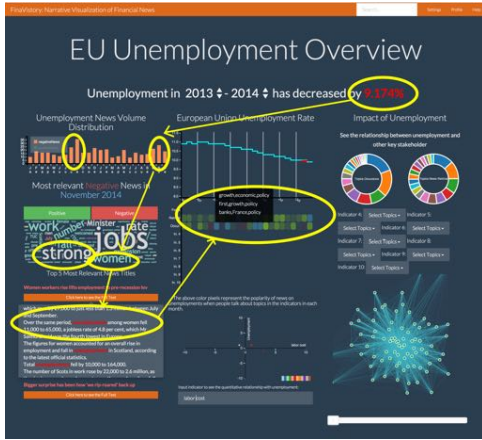

Fig. 12. Narrative flow for social cause of improvement

2) Social Perspective: Improvement starting from 2013: As observed from the time series the unemployment rate started to decrease from 2013. The information from the Word Cloud across the months show some positive signals like 'recovering', 'jobs' and 'strong'. Also by putting the topics related to fiscal policy into the summary plot, we can see some months like September and December 2013 there are stimulation policies agreement progresses and positive issues like women worker rise, and the decrease of unemployment had proven their effectiveness.
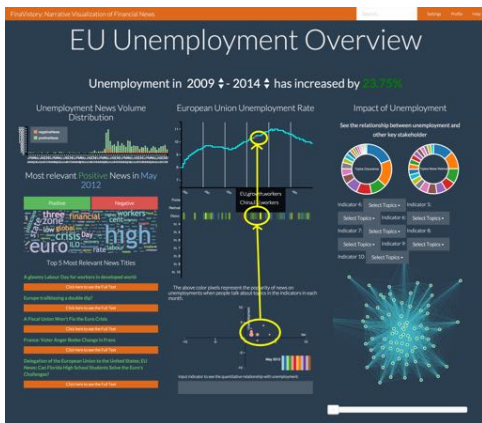

Fig. 13. Narrative flow for economic implication between tax and unemployment

3) Economic Implication: Tax and Unemployment: By plotting tax to bubble chart we can see that there is a relationship that increasing tax is correlated to increasing unemployment over the years. With a economic background this can be

explained by the fact that higher tax leads to higher tax wedge-the proportional difference between the costs of a worker to his or her employer and the amount of net earnings that the worker receives. Higher tax either increase a worker's cost that decrease employers' incentive to hire or decrease the labor supply if the employer is successful in shifting the tax burden to the labor. Therefore, for example in May 2013, when there was a bloom in discussion of increasing tax, the public reaction tended to be about increasing unemployment and topics about workers showed a relatively strong signal of increasing unemployment as well. This on the other hand give us a clue that tax cut could be a good way to solve the underlying unemployment situations [21].

VI. FURTHER DISCUSSION

The system has demonstrated multiple narrative flow and visualization messages to address significant events related to oil price in 2014 and Europe unemployment after financial crisis. Under different settings users can identify various reasons behind the issues and implications behind these kinds of Finance and Economics matters from financial, social and economical perspectives.

To consolidate our research work for future improvement, we asked prof. Yan YU, associate professor from the Department of Economics, HKUST Business School, to validate and discuss our results. We found out that, overall, although the results are reasonable and correct, there are still some shortcomings that need to be overcome in order to visualize financial news in all areas and provide a more sophisticated results:

1. The system should be able to identify scopes for keywords in different context. For example, Greece and Spain unemployment can be separated out from Europe unemployment.

2. The current polarity calculation cannot be applied to less popular financial news like individual stock prices due to the lack of indicating words around the main topics. Other approaches of polarity calculation should be investigated in finding out the polarity of each news articles.

3. We find out our system reveals more options and better narratives for oil price than Europe unemployment. It is because instead of Europe, individual countries like Germany, Greece and Spain should be investigated one by one instead. Possible methodologies can be used to focus on when the main topics need to be split into smaller sub topics.

4. Topic modelling methods collect topics throughout all the time. It might become over-aggregated and miss out some important features that only happens in a short time. Methods determining the time frame of topic extraction is necessary.

VII. CONCLUSION AND FUTURE WORK

In this paper, we have presented FinaVistory, a visual narrative system to generate narratives with visualization to help reader understand financial news from different perspectives. Compared with previous visual analytics systems, our system is unique in two ways. First, instead of presenting one whole

picture satisfying one purpose, we split different parts of visualizations into different areas to address different kinds of intentions of reading news, and the outcomes are different reader based combinations of visual analysis, which are up to the readers to judge and understand also. Secondly, our system does not aim at focusing one particular type of readers, instead it tries to fulfill all sorts of needs from people with different concerns, which is inspired by the 'long tail' theorem[22][23].

In the future, we are interested in combining a significant amount of narrative visualization, linking them all together to achieve full functionality for reading all topics of financial news, with a KISS (Keep It Simple Stupid) layout that readers do not need to understand the whole Economic and Financial situations by reading newspaper articles one by one anymore.

References

[1] Edward Segel and Jeffrey Heer. Narrative visualization: Telling stories with data. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1139–1148, 2010.

[2] Christina Granville. Becoming a financial analyst, 2006.

[3] Elizabeth Wu and Peter Phillips. Financial markets in motion: Visualising stock price and news interactions during the 2008 global financial crisis. *Procedia Computer Science*, 1(1):1765–1773, 2010.

[4] Remco Chang, Mohammad Ghoniem, Robert Kosara, William Ribarsky, Jing Yang, Evan Suma, Caroline Ziemkiewicz, Daniel Kern, and Agus Sudjianto. Wirevis: Visualization of categorical, time-varying data from financial transactions. In *Visual Analytics Science and Technology, 2007. VAST 2007. IEEE Symposium on*, pages 155–162. IEEE, 2007.

[5] Ryohei Hisano, Didier Sornette, Takayuki Mizuno, Takaaki Ohnishi, and Tsutomu Watanabe. High quality topic extraction from business news explains abnormal financial market volatility. 2013.

[6] Marko Grobelnik and Dunja Mladenic. Visualization of news articles. *Informatica (Slovenia)*, 28(4):375–380, 2004.

[7] Thian-Huat Ong, Hsinchun Chen, Wai-ki Sung, and Bin Zhu. Newsmap: a knowledge map for online news. *Decision Support Systems*, 39(4):583–597, 2005.

[8] Weiwei Cui, Huamin Qu, Hong Zhou, Wenbin Zhang, and Steve Skiena. Watch the story unfold with textwheel: Visualization of large-scale news streams. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(2):20, 2012.

[9] Amos Tversky and Daniel Kahneman. Rational choice and the framing of decisions. *Journal of business*, pages S251–S278, 1986.

[10] Jayne Godfrey, Paul Mather, and Alan Ramsay. Earnings and impression management in financial reports: the case of ceo changes. *Abacus*, 39(1):95–123, 2003.

[11] Fernanda B Viegas, Martin Wattenberg, and Jonathan Feinberg. Participatory visualization with wordle. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1137–1144, 2009.

[12] Kyle Koh, Bongshin Lee, Bohyoung Kim, and Jinwook Seo. Maniwordle: Providing flexible control over wordle. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1190–1197, 2010.

[13] Fernanda B Viégas and Martin Wattenberg. Timelines tag clouds and the case for vernacular visualization. *interactions*, 15(4):49–52, 2008.

[14] Andreas Noack. Modularity clustering is force-directed layout. *Physical Review E*, 79(2):026102, 2009.

[15] Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.

[16] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.

[17] Tim Chien. Html 5 word cloud, 2013.

[18] Alessandro Cologni and Matteo Manera. The asymmetric effects of oil shocks on output growth: A markov–switching analysis for the g-7 countries. *Economic Modelling*, 26(1):1–29, 2009.

[19] Mark A Hooker. Are oil shocks inflationary?: Asymmetric and nonlinear specifications versus changes in regime. *Journal of Money, Credit, and Banking*, 34(2):540–561, 2002.

[20] N Rodriguez-Alvarez, X Bosch-Lluis, A Camps, A Aguasca, M Vall-llossera, E Valencia, I Ramos-Perez, and H Park. Review of crop growth and soil moisture monitoring from a ground-based instrument implementing the interference pattern gnss-r technique. *Radio Science*, 46(6), 2011.

[21] Christopher A Pissarides. The impact of employment tax cuts on unemployment and wages; the role of unemployment benefits and tax structure. *European Economic Review*, 42(1):155–183, 1998.

[22] Volker Hoyer, Katarina Stanoesvka-Slabeva, Till Janner, and Christoph Schroth. Enterprise mashups: Design principles towards the long tail of user needs. In *Services Computing, 2008. SCC'08. IEEE International Conference on*, volume 2, pages 601–602. IEEE, 2008.

[23] Erik Brynjolfsson, Yu Jeffrey Hu, and Michael D Smith. From niches to riches: Anatomy of the long tail. *Sloan Management Review*, 47(4):67–71, 2006.