SONETA: A Social Media Geo-Trends Analysis Tool

Ioannis Ioannidis[1], Elefterios Papachristos[1], Christos Katsanos[1, 2], Nikolaos Karousos[2], Christos Fidas[1], Nikolaos Avouris[1]

[1] HCI Group, University of Patras, Patras, Greece
[2] Hellenic Open University, School of Science and Technology, Patras, Greece
`ioannidg@upatras.gr, {epap, ckatsanos}@ece.upatras.gr,`
`karousos@eap.gr, fidas@ece.upatras.gr, avouris@upatras.gr`

Abstract. Social media are used for expressing thoughts and opinions. Finding information hidden inside the vast amount of social media data is an overwhelming task especially for Small and Medium Enterprises (SMEs) that have limited resources. In this paper, we present an early prototype of SONETA, a social media geo-trend analysis tool that supports users, and specifically SMEs, by helping them detect social media trends in specific geographic areas. First, its architectural design and development process is delineated, followed by the presentation of a study that evaluated SONETA in terms of its usefulness and usability. Three SMEs participated in the SONETA evaluation study. Evaluation results showed that SONETA can be used in order to help SMEs in engaging social media more effectively.

Keywords: social media tools, geo-location analysis, visualization, Twitter

1   Introduction

From the beginning of Web 2.0 and subsequently the rise of Social Networking Sites much has changed in the way that the Web works. Social media transformed users from passive consumers to active generators of information. People participating in social media comment about virtually any aspect of their lives, they share personal moments, they discuss breaking news, and they express their thoughts and feelings about products or people. The wealth of information that is created by the millions of users that participate on a daily basis in social media poses an immense research opportunity for a variety of disciplines.

Social media research has been used in a variety of contexts, for example it has been successfully used in journalism in order to discover breaking news faster than it was previously possible with traditional media [1]. A variety of tools have been created to aid journalists in their work as for example applications that can help them identify and highlight interesting and trustworthy sources in the Twitter stream [2]. Emergency situations and crisis management is another field in which social media have been used with great success. For example social media has been used in order to support emergency planning, risk and damage assessment as well as for enhancing awareness about emergency situations. Social media data have also been proven to be

valuable in helping emergency operators to navigate through areas with heavy traffic [3]. In Japan it has been demonstrated that earthquake alerts could be created faster by using data derived from social media than official announcements from government agencies [4].

Much attention has been recently devoted to examine whether the business sector can gain value from social media [5-8]. For many years corporations and institutions relied on traditional structures and were content with a model in which the image of a company was controlled by their marketing and public relation department. The companies have gradually lost some of this control since social media enables consumers to talk freely to each other about products or services and express their opinions or complains [9]. Although companies cannot control or intervene in conversations about their products or services that happen in various social media platforms they nevertheless can gain extremely valuable feedback by listening carefully to what their customers have to say about them [10].

Listening to customer conversations in order to gain feedback about their own or their competitors' products as well as monitoring the reputation of their brand is not an easy task for companies considering the large amount of data that is generated on social media daily. Twitter alone is being used by 190 million users many of which are posting multiple times per day [1]. Taking into account the multitude of social media platforms it becomes apparent that in order gain valuable information companies have to examine large amounts of unstructured data. However, many companies are reluctant or unable to develop strategies and allocate resources in order to engage effectively with social media [9]. Social media engagement requires time, as well as commitment in personnel and financial resources, which can be difficult even for large companies, but even more difficult for Small and Medium Enterprises (SMEs) [6]. Although, social media present SMEs with excellent opportunities to compete with larger companies, SMEs are severely limited by their inability to acquire the necessary skills, hire additional employees or external help in order to learn use them effectively.

The need for new methods and technologies that could help companies harness the power of social media has led to the profusion of social media tools. These tools can be used among others for event detection, crisis management as well as for trend and campaign monitoring. The majority of them perform some kind of filtering to a vast amount of social media data in order to present only relevant information to its user promising increased efficiency. According to [1], more than 200 social media tool are available in the market today that vary considerably in terms of general scope, functionality, application areas as well as on the matter of price; ranging from zero to several thousand Euros per month [12]. The main problem with the commercials tools is that although they typically provide a wide range of functionalities their price often exceeds the budgets of most SMEs. On the other hand low-budget or free tools usually provide only basic functionality, they focus on single Social Networking Sites and don't offer adequate support to their users [11]. Social media tools are also developed for academic research but these tools are usually created as a proof of concept and are not available to the general public to be used in an applied setting.

Considering the multitude of social media tools available today it is surprising that none has been created specifically to address the needs of SMEs, which represent a large percentage of the economic power of most countries [13]. The main barriers that hinder SMEs to adopt social media are scarcity in time, financial resources and technology skills [14] as well as negative views about their usefulness, or unfamiliarity with a particular technology. In this paper we present the development process of a social media tool called SONETA (Social Networks Analysis Tool) that aims to help SME's lift such barriers. We describe the initial assumptions that have been made in order to create a first version of the tool as well as the results of a first evaluation iteration that was conducted involving three typical regional SMEs.

The rest of the paper is structured as follows. In section 2 of the paper we discuss how SONETA addresses the specific needs of SMEs and the conceptual design of the tool. In section 3 of the paper we discuss the main components of the system from an architectural point of view, and in section 4 we describe a typical usage scenario of the tool. Next, we describe a study in which usability experts and end-users (three SMEs) evaluated SONETA in terms of its usability and usefulness in actual practice. The paper concludes with a discussion on the implications of this research and directions for future work.

2   Conceptual Design

Time and money constraints as well as lack of IT expertise [11] are special characteristics of SMEs that could hinder them in adopting and using social media as effectively as large corporations. These characteristics have been taken into consideration during the creation process of SONETA.

Browsing social networking sites in order to monitor customer opinions and conversations relevant to a specific business domain can be a tedious and time-consuming task. Outsourcing the task, hiring new or training existing personnel to be responsible for social media management is typically not an option for SMEs. Thus, appropriate social media tools that can support the task of data filtering would be of great value for SMEs. For this reason, the highest priority in terms of functionality that has been given to SONETA was to enable users to navigate the unstructured stream of social media and to help them extract useful and relevant to their business information. Various data visualizations techniques have been considered that could help the user quickly recognize relevant information as well as to facilitate further exploration of the data. The lack of IT expertise, which is another characteristic of SMEs, has been addressed by giving special emphasis on creating a usable and intuitive user interface that would not require time consuming training sessions and lengthy instructional material.

Geographic location of customers is considered as an important factor for businesses [15]. Geographic segmentation helps marketers identify the specific needs of their customers. This is true for multi-national businesses as well as for SMEs. Even though SMEs may not use geographic segmentation marketing, geo-location still remains an important factor for SMEs since they often have local clientele and are not

attempting to broaden their business beyond their own region [14]. SMEs are proba-
bly interested in monitoring only social media data that originate in a specific and
confined geographic location and not in a specific country or even the entire world.
However, since not all SMEs have their customers confined in specific geographic
regions, restriction of social media results should be an optional feature that the user
could choose to use or not.

In order to meet the above requirements, we created an experimental prototype that
is able to detect social media trends for specific keywords in specific geographic are-
as, entitled SONETA. These trends are defined as the most important related terms
around a user-defined topic, keyword, phrase or geographic region. In this initial de-
velopment phase SONETA extracts information only from the Twitter stream.

3   Architectural Design

The SONETA system architecture is consisted of four main layers (Fig. 1): the com-
munication layer that is responsible for querying Twitter, the storage layer where all
the data (processed and raw) are stored, the preprocessing layer that transforms the
unstructured data into top-level keywords, and the presentation layer that visualizes
the top keywords. In the next paragraphs we provide a brief technical description of
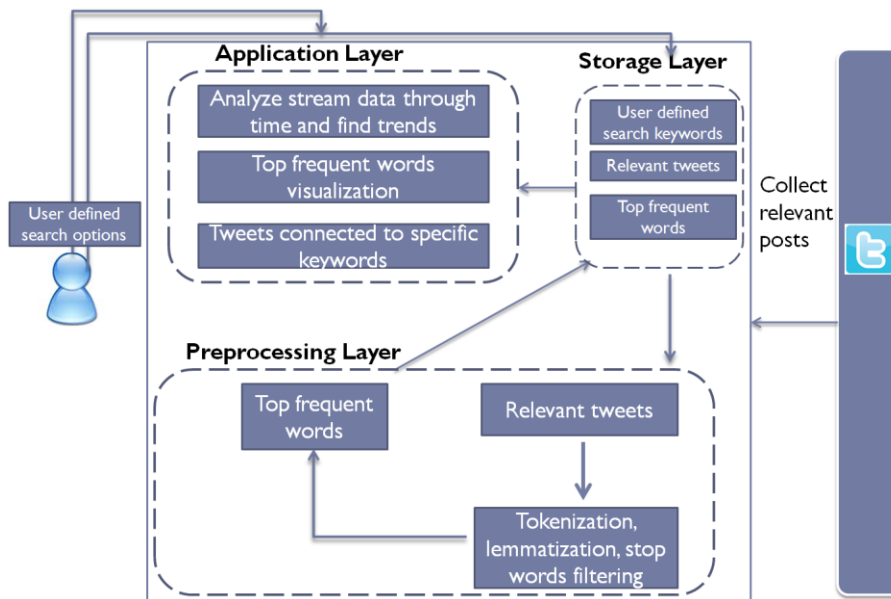each layer.

Fig. 1. Architectural design

Communication Layer: SONETA communicates with the Twitter Search API 1.1 in
order to retrieve tweets related to user-defined keywords. To this end, we used a PHP

library as a client for accessing Twitter resources. SONETA asks for tweets containing specific keywords posted from a specific location.

Storage Layer: Results from the communication layer are stored in the storage layer. The storage layer consists of two main components: a MySQL Database and a data directory. The database stores the user-defined keywords, the results from the Twitter stream and the user credentials. It has been designed in such a way that new components and services can be easily added later. The data directory stores the results from the preprocessing layer.

Preprocessing Layer: Stored tweets are the input of the preprocessing layer. The preprocessing layer is implemented in PYTHON with the help of the NTLK package. For each tweet, external links, mentions and stop words of different languages are removed since they don't provide any useful information for the purposes of our system. Retrieved results are tokenized and the most frequently appearing ones are stored in the data directory component of the Storage layer using a JSON format.

Presentation Layer: Processed results are visualized at the presentation layer. Visualizations are implemented with the D3.JS, a javascript library for data visualizations. Using this library, we developed a timeline graph, a word cloud, a bubble chart and a treemap visualization. In addition, Google maps was used for building the geo-location feature.

4 Typical Usage Scenario

In this section a typical usage scenario of SONETA is presented. First, the SONETA user starts by creating a new stream for searching across Twitter (Fig. 2). A stream can be viewed as a subset of the Twitter stream.

The user can restrict this subset by applying appropriate search filters. Currently, the tool offers two options: keyword filtering and geo-location filtering. When keyword filtering is active, the steam includes Twitter messages containing the specified keywords. This can be useful for monitoring a brand name, a product or a domain that is of interest for an SME. In addition, SONETA provides the option to filter Twitter messages according to a specific geographic region. Other tools also provide such an option, but require setting a longitude-latitude pair and the length of the desired radius. SONETA simplifies this process, so that even less tech-savvy users can easily perform it. In specific, a map is provided to the SONETA user, in which a circle represents the region of interest. Initially, this region is placed around the user's current location, but it can be resized (using the handlers; Fig. 2) and moved (by drag-and-drop). Using this filter the user retrieves all the twitter messages originating from the specified geographic location. Examples of how geo-location filtering can be useful for an SME are: understanding trend topics in customer conversations located in specific regions, estimating the impact of targeted advertising campaigns, and collecting ideas for new user products or services customized to a specific region. Finally, both filters can be combined in order to receive messages that contain a user-specified keyword in a specific region.
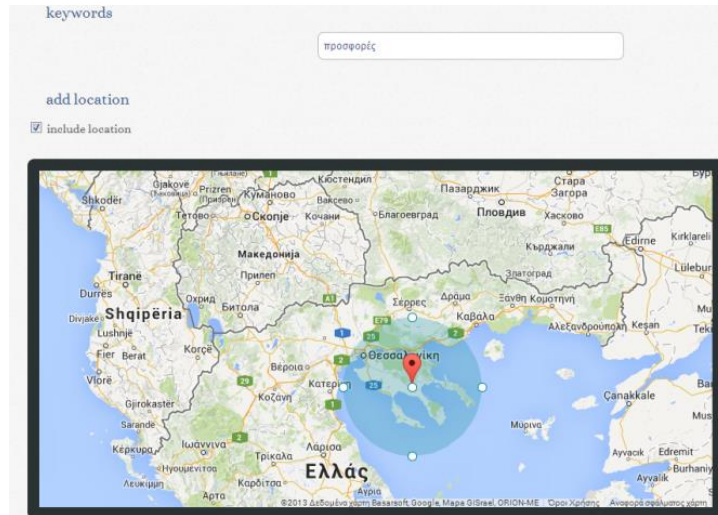
Fig. 2. New stream creation interface

Figure 3 presents an example of the output produced by SONETA. The timeline presents two bar charts of the total number of tweets per day that are contained in the user-defined stream. The bottom bar chart provides a bird's-eye-view and helps the user to easily identify areas of interest, whereas the top bar chart displays an adjustable zoomed portion of the bottom bar chart. Timeline representation was used in order to provide a spatial perspective of the trends [16].
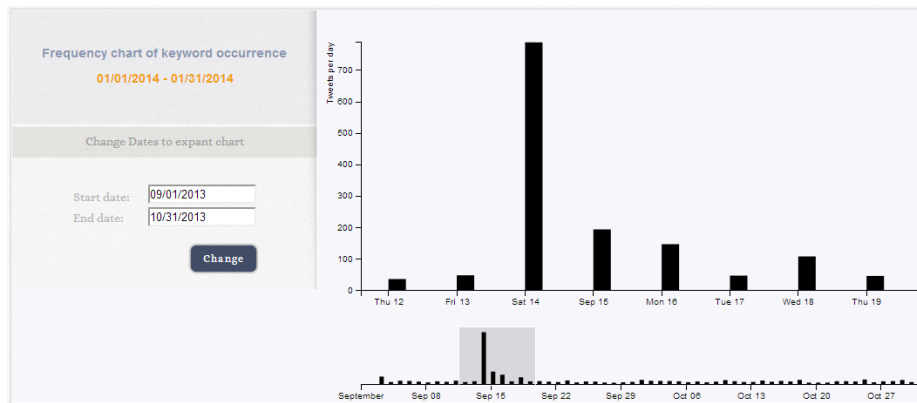

Fig. 3. An example of the SONETA output: total numbers of tweets per day

For each day, the SONETA user can see the most frequent words for the selected stream by clicking on the respective bar of the timeline. The initial representation of the most frequent words is a word cloud (Fig. 4, left), which is considered a powerful representation of textual data [7]. Monochromatic color combination and size are used to denote the frequency of the words appearing in the selected stream. SONETA pro-

vides two additional data visualizations: a) a bubblecloud in which the size and the color represent word frequency (Fig. 4, mid), and b) a treemap that uses the size and the color of the rectangles to represent word frequency (Fig. 4, right).
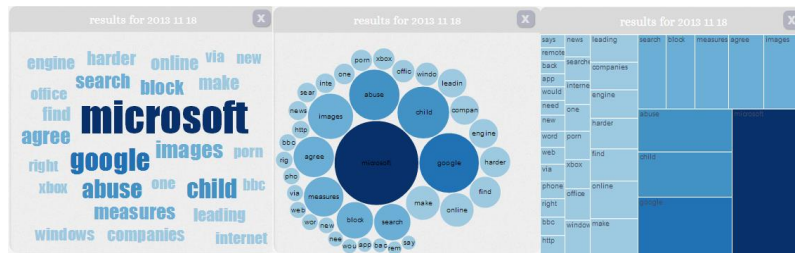

Fig. 4. An example of the SONETA output: top keywords and the available visualizations

5 Evaluation Study

This section presents an evaluation study of SONETA in terms of its usefulness and usability.

5.1 Method

The evaluation methodology we followed had two phases. During the first phase, a formative usability evaluation of SONETA was conducted by three usability experts. The experts performed heuristic evaluations to review the tool's interaction and interface design. Feedback from this phase was used to identify important usability issues and to improve the tool before it could be given to representative end-users. In the second phase, SONETA was evaluated by three typical regional SMEs after they had used it in their business activities for about a month. The SMEs that were selected to participate in the evaluation study were typical Greek companies with diverse profiles, business goals, operational and organizational structures and initial attitude towards social media. SME1 employs 16 persons and is a Business-to-Business company active in undertaking research and development projects. SME2 is a six-person software design and development company, mainly involved in Business-to-Business activities. SME3 employs eight persons and is a Business-to-Consumer company that acts as an online meta-retailer of shopping deals on a variety of products and services.

Since we were interested to examine whether SONETA could aid SMEs in achieving their objectives and help them engage with social media more effectively, we chose an evaluation scenario in which the companies would use the tool for an extended period with as little intervention from us as possible. First, a pre-study interview was conducted with the SMEs in order to profile them and identify their needs and problems in regard to social media usage. After the initial interview, SONETA was introduced to the participants followed by a training session in which instructional material was given to each SME. Subsequently, the participating SMEs were left to use SONETA for a period of one month for whatever reason they deemed appropri-

ate. During this period and following a ten-day interval, three semi-structured interviews were conducted with each SME in order to collect qualitative feedback about whether and how they used SONETA, as well as to collect suggestions for improvement. Each interview placed emphasis on a different aspect. The first interview focused on the deployment process (e.g. level of effort required, training quality), the second interview on the user experience (e.g. usability, usefulness, learnability) and the final interview on overall impact (e.g. business goals support, change in policies). Apart from evaluating the features and functionalities that SONETA currently provides, a secondary objective was to identify future directions for improvement

5.2   Results and Findings

The pre-study interview showed that all three SMEs were primarily active in Facebook, which is typical for Greek companies. All of them had Twitter accounts but they were considerably less active in Twitter compared to Facebook. Two of the participating SMEs just mirrored Facebook posts on Twitter, whenever possible, while the third one used Twitter mostly to monitor content posted by others. SME3 reported being particularly interested to learn how to use Twitter effectively for the purpose of promoting their business. In the pre-study interview, no SME mentioned the importance or the need of geo-location as a feature of a social networking tool. Finally, even though all SMEs mentioned that they had problems to find interesting content to post they did not mention the need for tools that could help them find relevant to their business information in the social media stream.

Even though in the pre-study interview the SMEs did not mention the need for tools that filter data from the Twitter stream and identify information relevant to their businesses trends, their engagement with SONETA made them realize the potential of this feature. The participating SMEs reported that the possibility to restrict results in a specific geographic location was very interesting conceptually. All SMEs experimented with this SONETA feature, but their perceptions about the usefulness of the results varied considerably among them. SME1 and SME2 struggled to find appropriate ways to use it in their specific business domain. They reported having difficulties in formulating strategies on how to act upon information they extracted through SONETA and difficulties in finding appropriate keywords in order to create relevant streams. They asked about examples or good practices of using SONETA in similar to their own domains. On the other hand, SME3 found SONETA extremely useful. In contrast to the other participating companies, SME3 was active in a more consumer oriented sector. In addition, they were more motivated compared to the other two SMEs, since they explicitly stated in the pre-study interview that they wanted to learn to use Twitter more effectively and that they strongly believed in the power of social media. They used various keywords as well as the geo-location feature to find trending conversations that were particularly useful to them. The following excerpt from the transcribed interviews indicates that they did find SONETA useful:

"It [SONETA] helped me increase my followers on Twitter by looking for trending conversations in specific areas as well as in Greece in general. Taking part in these

conversations allowed me to create 115 followers in a matter of two weeks. The tool allowed me to find what the competitions is tweeting about and to identify what topics are gaining attraction. This helped me make decisions about what to tweet myself."

In regard to the tool's usability the users reported at the beginning of the evaluation that they faced difficulties in learning how to use it. They reported that they would most likely not be able to use SONETA without the introductory material that was given to them at the beginning of the study. User comments helped us realize some moderate user interface problems that were resolved as soon as possible. In addition, we included supplementary instructions for first-time users, describing the main components of a stream creation process as well as how results can be further explored. We also added functionalities for the management of the streams as suggested by the participating SMEs. They confirmed that the word cloud visualization was particularly useful because it helped them to explore specific domains and identify important keywords and trends in a straightforward and quick manner. However, SMEs asked for the additional possibility to see specific Tweets, apart from the word cloud. This request was also implemented at a later time, allowing users to see all tweets that contain a certain keyword by simply clicking on it on the word cloud.

6   Conclusions and Future Work

In this paper we described the need for social media tools that specifically address SMEs because of their increased constraints, such as lack of resources and IT expertise, compared to larger corporations. To that end, we developed SONETA, a tool that filters information from the Twitter stream.

SONETA was introduced to three typical SMEs who used it for a period of one month in the context of an evaluation study that was designed to maximize ecological validity. The main goal of the evaluation was to understand if such a tool can support SMEs in engaging effectively with social media data. Evaluation showed that in general, SONETA was perceived by the participating SMEs as a reasonably useful tool. More specifically it was found that:

(a) SONETA can be particularly useful for SMEs that are mostly engaged in Business-to-Customer activities.
(b) There is a need for exemplary scenarios and case studies of using SONETA in specific domains.
(c) Attitude towards social media in general, and Twitter in particular, conditioned the perceived quality of the tool.

The feedback that was received from the evaluation study was used to further improve the tool, and will guide the development of future versions. For instance, the next version of SONETA will provide the option to retrieve data from multiple social networking sites. In addition, we plan to examine how such data can be integrated and summarized in order to help users gain insights from different perspectives. Future work also involves investigating how various types of visualizations may help users gain insight from social media data that originate from different domains. Finally,

considering the real time nature of social media, the visualization layer in the SONETA system architecture could also be modified to support real time updates.

References

1. Brett, M., Kevin B., Yamara, S., Beomjin, K..: Twitterreporter: Breaking News Detection and Visualization through the Geo-tagged Twitter Network. In Wei Li, editor, CATA, pp. 84-89. ISCA (2011)
2. Diakopoulos, N., De Choudhury, M., Naaman, M.: Finding and Assessing Social Media Information Sources in the Context of Journalism. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2451–2460. ACM, New York, NY, USA (2012)
3. Best, D., Bruce, J., Dowson, S., Love, O., McGrath, L.: Web-based Visual Analytics for Social Media. In AAAI ICWSM SocMedVis: Workshop on Social Media Visualization, pp. 2–5, AAAI Technical Report WS-12-03 (2012)
4. Takeshi, S., Makoto, O., Yutaka, M.: Earthquake Shakes Twitter Users: Real-time Event Detection by Social Sensors. In: Proceedings of the 19th international conference on World wide web (2010)
5. Mangold, W.G., Faulds, D.J.: Social Media: The New Hybrid Element of the Promotion Mix. Business Horizons 52 (4), pp. 357–365 (2009)
6. Bulearca, M.: Twitter: A Viable Marketing Tool for SMEs? Global Business Management Research 2(4), 296-309 (2010)
7. Bughin, J., Manyika, J.: How Businesses Are Using Web 2.0: A McKinsey Global Survey. McKinsey Quarterly Web Exclusive, McKinsey and Company (2010)
8. Hawn, C.: Take Two Aspirin And Tweet Me In The Morning: How Twitter, Facebook, And Other Social Media Are Reshaping Health Care. Health Affairs 28(2), 361 (2009)
9. Kietzman, J.H., Hermkens, K., McCarthy, I.P., Silvestre, B.S.: Social media? Get serious! Understanding the functional building blocks of social media Business Horizons 54(3), pp. 241–251 (2011)
10. Chamlertwat, W., Bhattarakosol, P., Rungkasiri, T., Haruechaiyasak, C.: Discovering Consumer Insight from Twitter via Sentiment Analysis. Journal of Universal Computer Science 18(8), 973-992 (2012)
11. Stavrakantonakis, I., Gagiu, A.-E., Kasper, H., Toma, I., Thalhammer, A.: An approach for evaluation of social media monitoring tools. Common Value Management, 52-64 (2012).
12. Kasper, H., Dausinger, M., Kett, H., Renner, T., Finzen, J., Kintz, M., Stephan, A.: Marktstudie Social Media Monitoring Tools, Fraunhofer-IRB, Stuttgart (2012).
13. Michaelidou, N., Siamagka, N.T., Christodoulides, G.: Usage, Barriers and Measurement of Social Media Marketing: An Exploratory Investigation of Small and Medium B2B Brands. Industrial Marketing Management. 40, 1153–1159 (2011).
14. Stockdale, R., Ahmed, A., Scheepers, H.: Identifying Business Value from the Use of Social Media: An SME Perspective. In: PACIS 2012 Proceedings, Paper 169 (2012)
15. Broderick, A, Pickton, D. : Integrated Marketing Communications, Edinburgh Gate Pearson Education Ltd. Second Edition (2005)

16. Heer, J., Bostock M., Ogievetsky, V.: A Tour through the Visualization Zoo. Communications of the ACM 53(6), 59 (2010)