

# Using Infographics to Represent Meaning on Social Media

Erick López-Ornelas<sup>(✉)</sup> and Saúl Hermilio Sánchez Hernández

Information Technology Department and Theory and Design Processes Department,  
Universidad Autónoma Metropolitana, Cuajimalpa, Mexico  
elopez@correo.cua.uam.mx, saulsan7@hotmail.com

**Abstract.** In this paper we focus on the automatic creation of infographics based on automatic extraction of tweets. To achieve this goal, we define a model where five types of information can be represented: spatial, chronological, quantitative, hierarchical, and contextual or, as is usually the case, a combination of all five. To demonstrate the suitability of this approach a prototype was generated in order to create automatic infographics without any user assistance.

**Keywords:** Twitter · Social Media · Infographics · Visualization

## 1 Introduction: Social Media and Infographics

Social Media has become an essential component in the navigation of everyday life [1]. Social Media influences some aspects of human behavior from the way in which organizations operate to the way people shop and spend their time. Using Social Media vast amounts of information can be disseminated to worldwide audiences in an instant, while the web simultaneously offers an arena for public and private social interaction.

Social network sites (SNS) can be defined as virtual collections of user profiles which can be shared with others. Despite the prominence of the internet and social networking in modern life, research concerning information representation and visualization has been limited. We focus on the use of Twitter and the large amount of information that is generated daily. Twitter is a popular social media service that allows people to share updates, news, and information (known as “tweets”) with people in their Twitter network and beyond. In our approach we used tweets extracted from Twitter. A tweet is a little message of no more than 140 characters that users create in order to communicate thoughts, feelings, or even participate in conversations.

In short, there is a large amount of information generated in social networks which is not properly explored or exploited by users. Even if the tweets have a lot of information, there exists much more information that can be extracted and analyzed automatically.

In the other hand, the human brain is more able to identify and understand relationships and patterns if data is encoded into visual forms [2]. One form that has been used frequently are infographics whose definition are: “The use of computer-supported, interactive visual representations of data to amplify cognition [3]”; infographic is graphic visual representation of information, data or knowledge intended to clarify and integrate difficult information quickly and clearly [4]. For education (definition) for

infographic: a collection of graphic organizers integrates different media in simple diagrams: text, images, symbols and schemas [5]. In Human and Computer Interaction, infographics can improve user cognition by utilizing graphics to enhance the human visual system's ability to see patterns and trends. In other words, Infographic is a new way to visualize data. Another concept frequently used is information visualization (InfoVis) or data visualization [6]. Visualization is defined as [7] "mechanisms by which humans perceive, interpret, use and communicate visual information". The main aim of visualization is to communicate information more clearly and effectively by using graphical means [8].

Although infographics have been used for information visualization, they rarely have been used in Social Media and in any case created automatically. Usually it required user support for its coherent creation. In this article we explore the idea of using the large amount of information generated by Twitter and we propose the automatic construction of infographics coherently identifying the type of information we need to represent.

## 2 Types of Infographics

There are many graphic types for visualizing data, from bar graphics to pie charts, from tables to diagrams. Actually, most of the graphics used in visualization applications are a part of our lives since many years. Graphs allow us to explore data and observe patterns that no other approach can achieve.

Arabic numerals are preferable in infographics, the heading of table should put underlined and centered above them. Human mind can recognize visual information with more successful and lasting way when compared to with written or verbal information transfer [8]. Therefore, infographics design should be experienced to carry transmission of data visualization.

In informatics research [9], it is found that a rich interactive infographic capable of showing far more digestible information at a glance than conventional, tabular representations. The essential text content has been explained with well designed infographics. Just by reviewing the graphics, we can understand the whole idea of the reports. Moreover, with today's technology, infographics can also be transformed to animated images for the website version [9]. A graphical symbol or icon is defined as the smallest graphical unit that carries meaningful information.

Some major types of infographics base on its usability [10] are as follows:

- **Statistical Based.** This type of infographic includes diagrams, charts, graphs, tables, and lists. Among the most common devices are horizontal bar charts, vertical column charts, and round or oval pie charts, that can review statistical information. It can be made in interactive manner as well.
- **TimeLine Based.** Timeline show the sequence of events according to the time each event had happened. A timeline enables an audience to realize chronological relationships very quickly. Sometimes it shows in tabular, year-by-year paragraphs, etc.
- **Process Based.** These process based usually can be found in cooking magazines or explain about recipe using infographic. Also this type of infographic can be used to

clarify in workspaces of factory or offices. It can make readers to understand about its practices in limited space.

- Location or Geography Based. With widely use of GIS, maps can also consider as the best way to show geography based infographics. They include symbols, icons, diagrams, graphs, tables, arrows and bullets. There are many well known GIS notation that used in maps to identify highways, streets, subways, and facilities. Many familiar icons and symbols designed for places like tourist spots, hospitals, airports etc. Scale is the imperative consideration additionally because all places and landmarks are marked according to the exact scale or ratio.

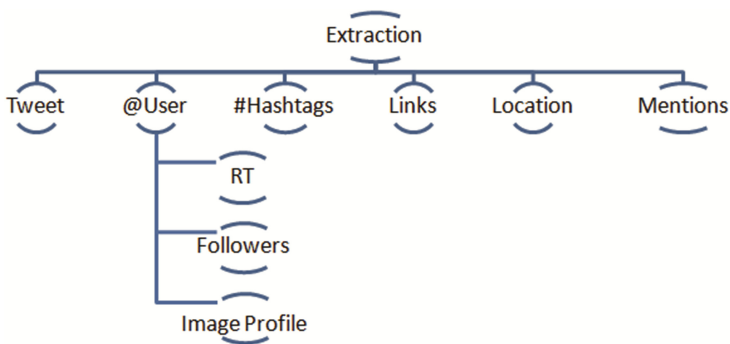
### 3 Mining Twitter Data

Twitter has its own convention that renders it distinct from other textual data. Consider the following Twitter example message (“tweet”): RT @john has a cool #car. It shows that users may reply to other users by indicating user names using the character @, as in, for example, @john. Hashtags (#) are used to denote subjects or categories, as in, for example #car. RT is used at the beginning of the tweet to indicate that the message is a so-called “retweet”, a repetition or reposting of a previous tweet.

The Twitter Application Programming Interface (API) [11] currently provides a Streaming API and. Through the Streaming API [12] users can obtain real-time access to tweets in sampled and filtered form. The API is HTTP based, and GET, POST, and DELETE requests can be used to access the data.

In Twitter terminology, individual messages describe the “status” of a user. Based on the Streaming API users can access subsets of public status descriptions in almost real time, including replies and mentions created by public accounts. Status descriptions created by protected accounts and all direct messages cannot be accessed. An interesting property of the streaming API is that it can filter status descriptions using quality metrics, which are influenced by frequent and repetitious status updates, etc.

Among all these elements extracted by the API, we are interested in analyze those listed in Fig. 1.



**Fig. 1.** Data extraction using the API of Twitter

The API uses basic HTTP authentication and requires a valid Twitter account. Data can be retrieved as XML or the more succinct JSON format. The format of the JSON data is very simple and it can be parsed very easily because every line, terminated by a carriage return, contains one object. Using this API, we can extract large amounts of data, however we need to find a better way to display and visualize this data.

### 4 Infographics: Modeling and Creating

Despite the difficulty in creating a design model, it would be useful to have one, in order to understand the overall picture of the infographic design process and especially the type of information that will be used.

The major challenge in order to design and create successful infographic is to understand what type of information it is trying to communicate. We have defined five different approaches – whether spatial, chronological, quantitative, hierarchical, contextual or, as is usually the case, a combination of all five.

The first three approaches are has been explored widely [13], but the hierarchical and contextual approaches are the most important contribution made in this work, together with the automatic creation process.

In the first step the user has to choose a topic of interest. This is defined using the #Hashtag that has to be entered to the system. The system has to extract, using the Twitter API, all relevant information related to this specific #Hashtag. Then the system asks what kind of representation has to be represented whether spatial, chronological, quantitative, hierarchical, and contextual. The system, process the data and define automatically the design of infographics and the most important the type of information that should be used. Then this infographic is presented to the user for validation.

In the Fig. 2 we shows the model used to create infographics focusing in the type of information that we need to visualize.

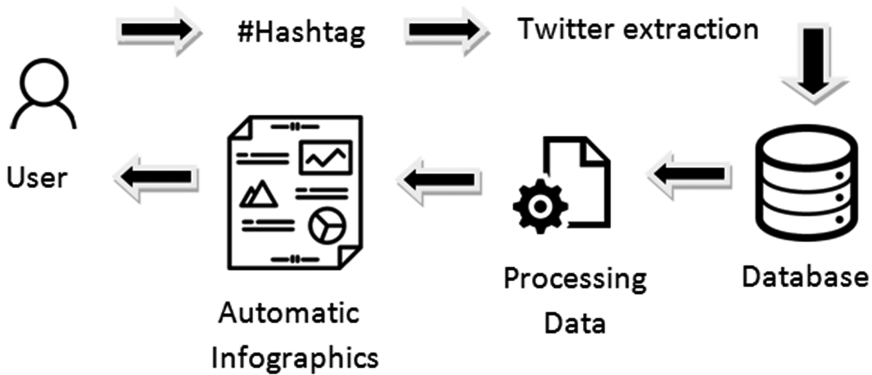


Fig. 2. Model of an automatic infographic creation

The different types of information that can be represented are described below.

## 4.1 Spatial

This information describes relative positions and the spatial relationships in a physical or conceptual location. Using this approach, it is possible to identify the spatial information and to determine where each Tweet has been written and published. The information is displayed using a map and we have the option to display the different geographical places. The essential elements for build this type of spatial infographics includes: the user profile, the user name, the tweet and the user location. In Fig. 3, we show an example of this approach.



Fig. 3. Spatial infographics using Twitter

## 4.2 Chronological

This information describes sequential positions and the causal relationships in a physical or conceptual timeline. In this type of infographics the result is displayed as a timeline, some chronological aspects can be discovered using this approach. In these representations we have to order all Twitter elements chronologically in order to be presented on the infographics, using this approach the date and time of tweet publication is determinant. In the Fig. 4, we show an example of automatic construction of a chronological infographic.

## 4.3 Quantitative

This Information describes scale, proportion, change, and organization of quantities in space, time or both. This infographic shows different data organized by different trends or details about the search in the form of numbers, graphics etc. So, the user can draw some conclusions and logic of their topics of interest.

Twitter elements required to build this type of infographics are the user names, the account names, the tweet, the platform where the tweet was published, the user profile



Fig. 4. Chronological infographics using the Hashtag #epn

(picture age, sex, etc.), and the numbers of times that the tweets is shared and marked as favorite.

With all this information we need to select and perform some operations in order to display the suitable infographics, using graphics or manipulating data. All this manipulation is performed based on style sheets and JavaScript. In the Fig. 5 we show the result of a quantitative infographics created automatically.



Fig. 5. Quantitative infographic using the Hashtag #Mexico

#### 4.4 Hierarchical

Hierarchies are some structures based on a criterion of subordination, i.e., we can define different levels taking into account some factors such as scales, the influences degree, periods of time or importance of a subject. This category was a contribution in this article, which emerged thanks to the previous analysis to represent information.

Using this design we can simplify the infographic and obtain visually different trends, patterns, measuring the followers or shared publications. That is the main reason we include this type of representation in our design model.

For the automatic construction of this infographics, it was necessary to extract the following items from the tweet: the user names, the profile picture, the user account information, the number of times a tweet has been shared and bookmarked. Having this information we can identify the influential users, or some communities in social network. In Fig. 6, we show an infographic designed identifying the most popular tweets using the Hashtag #donaldtrump.



Fig. 6. Hierarchical infographics using the Hashtag #donaldtrump

#### 4.5 Contextual

Contextual information is also used to design infographics. Using this information, it is possible to represent symbolic data or graphics, and to explore a set of circumstances surrounding specific issues or facts for some behavior or pattern from the original topic.

We realized that extracting Twitter information, it could exist the interest of the user to know significant topics linked to an specific Hashtag, for example what are different Hashtag related to the original one or who is the most influential user that is not mention directly in the Hashtag but is related to the original keyword.

All these variables can also be extracted and represented with data in order to identify a context. To construct this infographics, we used Cloud Words in order to visualize all surrounding topics. Normally, the elements extracted: the tweet, the user profile, the original Hashtag and the related Hashtag. The visual design of the infographics can change according the context of each Hashtag search. In the Fig. 7 we show an example.



Fig. 7. Contextual infographic using the Hashtag #futbol

## 5 Conclusions

The Twitter messaging service is wildly popular, with millions of users posting more than 200 million tweets per day. This stream of messages from a variety of users contains information on an array of topics, including conventional news stories, events of local interest (e.g., social movements), opinions, real-time events (e.g., earthquakes or traffic jam), and many others. Unfortunately, all this information is not well exploited by users.

Automatic extraction using Twitter's APIs provide access to tweets from a particular time range, from a particular user, with a particular keyword, or from a particular geographic region. We think that this form of automatic extraction must be used regularly in order to have a better understanding of what happens on Twitter.

In other hand, infographics are traditionally viewed as visual elements such as charts, maps, or diagrams that aid comprehension of a given text-based content. We have shown in this paper how to represent Twitter data in order to get new meaning and discover new knowledge.



In this article, we have described an approach for efficient extraction of information from Twitter searching a particular subject (the use of Hashtag #). This information is properly reorganized and presented through infographics automatically. These infographics can be designed to represent Twitter information on qualitative, temporal, geographic, hierarchical or contextual. A prototype tool was created to determine the suitability of this kind of infographics.

In conclusion, we need to find new ways to represent all this large amount of data that is generated by Social Media. We argue that visualization and specifically Infographics has a great potential. Some aspects should be improved, like the visual design and the interactive interaction, but we think is a good beginning and research must continue in this direction.

## References

1. Amichai-Hamburger, Y., Vinitzky, G.: Social network use and personality. *Comput. Hum. Behav.* **26**, 1289–1295 (2010)
2. Cleveland, W.S.: *The Elements of Graphing Data*, Revised edn. Hobart Press, New Jersey (1994)
3. Card, S., Mackinlay, J., Shneiderman, B.: *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers, Los Altos (1999)
4. Smiciklas, M.: *The Power of Infographics: Using Pictures to Communicate and Connect with Your Audience*. Que Biz-Tech Series. Que Publishing, Indianapolis (2012)
5. Serenelli, F., Ruggeri, E., Mangiatordi, A., Ferri, P.: Applying the multimedia learning theory in the primary, school: an experimental study about learning settings, using digital science contents. In: *Proceedings of the European Conference on e-Learning is the Property of Academic Conferences* (2011)
6. Card, S.: Information visualization. In: Sears, A., Jacko, J.A. (eds.) *Human-Computer Interaction: Design Issues, Solutions, and Applications*, pp. 510–543. CRC Press, Boca Raton (2009)
7. Scaife, M., Rogers, Y.: External cognition: how do graphical representations work? *Int. J. Hum. Comput. Stud.* **45**, 185–213 (1996)
8. İnan, B., Dur, U.: Analysis of data visualizations in daily newspapers in terms of graphic design. *Proc. Soc. Behav. Sci.* **51**, 278–283 (2012)
9. Spry, K.C.: An infographical approach to designing the problem list. In: *IHI 2012: Proceedings of the 2nd ACM SIGHT International Health Informatics Symposium* (2012)
10. Artacho-Ram, M.A., Diego-Mas, J.A., Alcaide-Marzal, J.: Influence of the mode of graphical representation on the perception of product aesthetic and emotional features: an exploratory study. *Int. J. Ind. Ergon.* **38**, 942–952 (2008)
11. Twitter API: (2010). <http://apiwiki.twitter.com/>
12. Kalucki, J.: Twitter streaming API (2010). <http://apiwiki.twitter.com/Streaming-API-Documentation>
13. Rajamanickam, V.: *Infographics seminar handout* (2005, 2011)