Information balance between newspapers and social networks

Francesco Mazzeo Rinaldi[1], Andrea Russo[2], Giovanni Giuffrida[1]

[1]Department of Political and Social Science, University of Catania, Italy, [2]Department of Physics and Astronomy, University of Catania, Italy.

Abstract

Competing newspapers, tend to publish the same information in a given time frame. However, each editor tends to aggregate and present the news according to certain criteria such as editorial policies, filtering strategies, readers base, etc. Thus, the proper choice and filtering of information makes one newspaper different from the other and, the proper management of such criteria, may deem the success or failure of a newspaper.

From the editor's perspective, the news selection process is a trade-off between informativeness and attractiveness, as determined by the readership. Moreover, is it possible that cultural and political inputs from social media may impact the news selection process?

Political news on social networks represent nowadays a valuable informative asset that gives the possibility to correlate newspaper information with public request expressed on social networks. We believe that it is possible to develop a theory to mitigate the newspaper's cultural identity with the public information needs collected on social media.

In our work, we show how to measure the society's request for information through the analysis of public reaction to certain articles on social networks, in particular we present how studying the hashtags and articles shared can be conveyed to understand social dynamics in nowadays discussion.

Keywords: Information; Newspaper; Social media, Social need; Big Data.

1. Introduction

The common goal of all newspapers is to make money by selling news. Thus, their news, and their presentation, must be attractive for the readership. However, considering the amount of new information constantly produced, a proper continuous smart selection of those is necessary and represent a crucial factor for the newspaper survival.

In this context, information extracted in real time from social media provides a valid input for the selection process. From a research perspective, data becomes crucial to understand how newspapers gain information to balance their cultural basis with the population's informative needs.

The purpose of this paper is to demonstrate that social network data could be successfully used as a new resource for newspapers. Many newspapers use big data to improve their articles, but in this case, we propose a new way to analyze big data to improve topic selection. It would be feasible that, with a more peculiar use of big data there will be the possibility to respond to unconscious requests for information (on a specific topic) by the social network's users. This unconscious need could be recognized and used in order to take advantage of that, in order to create highly informative content, so as to achieve the end of people's interest in the topic, even though the newspaper is specialized in a different sector.

From the methodologic point of view, hashtags can be used to estimate if there is apprehension or curiosity about a particular topic. For instance, we could study whether the word "Bojo", an acronym for Boris Johnson communicating closeness to the persons, is used more frequently among readers talking positively, about Brexit, compared to the ones not approving the Brexit plan. Thanks to this logic it is plausible to relate the use of the name to the newspaper's position on Brexit.

Today, every newspaper has its public digital space where readers discuss and share information. Thus, it becomes possible to enhance such a network by improving it or to attract new clusters of people that have a similar culture to the newspapers' one.

In particular, we can understand which topics are more suitable for the newspaper as recognized by the public. For example, Justice and Economics news are the most prominent themes for newspapers analyzing the political system, as well as wars and scientific discoveries.

By measuring real-time hashtag trends on social networks for particular topics we could promptly inform the newspaper about groundbreaking news.

For instance, if a newspaper selects the news about Brexit highlighting the economic effects rather than the causes of Brexit itself, assuming that to be the informational need of people if favor of Brexit, there is the possibility that the reader's request appears on the social networks.

It would be possible to analyze the selection process of the newspaper by studying the hashtags inside the articles published, in relation to the sharing network generated by the user activity.

Thus, social media data analysis becomes a key factor to understand readers' news needs to prepare a more attractive and accurate article set on a newspaper.

A case study for this subject could be the commentaries' data mining on articles that discuss a new law proposal of the Italian government discussed in late 2019. The analysis shows whether readers commenting the article are in favour or against it. (Francesco Mazzeo Rinaldi, Giovanni Giuffrida, and Tom Negrete).

2. Scientific results & methodology

Big Data and algorithms are nowadays a very common topic in social and political sciences. The initial debate about technology and social sciences have produced many publications with many interesting results. In many cases technology has produced some relevant suggestions to advance scientific research in those disciplines. Furthermore, through its continuous improvement in recent years, technology has helped to better understand our own society. *Mutatis mutandis.*

Thanks to social network data and its related algorithms, today there is the possibility to measure the unbalance between newspapers and social needs for information. There are some relatively new researches that using big data have brought innovative results. The "traditional" sentiment-analysis on a topic or an article, can be used by a newspaper to intensify the research on data-driven strategies to be used commercially. The benefits of sentiment-analysis have long been recognized, but with this research we want to show how big data, if used for a precise topic, could answer the need for information of the social network users, widening the research domain and its methodological settlement. A popular topic on social network could create a contrast between different points of view or show disinformation. In both cases, but more specifically when there is little or no information, through the use of big data it is plausible to detect the peoples' need for information (among supply and demand). The existence of the Points of consideration (which will be later examined in depth) is valuable to highlight the scarcity of information (or high number of discussion on a topic) on social media, and the absence of articles (supply) from newspapers. An accurate data analysis could show a demand for information, which will later be acquired from the newspapers and as a result, suitable articles be produced. For instance, it has been used to study trend topics in social networks, and the corresponding offer of information, thus showing whether there is an offer from newspapers.

Such a correlation provides the opportunity to estimate the public reaction to certain articles. For instance, in the political context, the algorithm can analyse keywords and hashtag used as a thematic aggregator, so to measure the level of attractiveness of a specific news.

2.1. Data source

We propose to use Twitter as the data source because of three crucial aspects: first of all, Twitter makes it easier to acquire data more than the other competitors. The text format is easily analysed because of its shortness (compared to other social networks such as Facebook or TikTok): you have just 280 characters to express your idea, so the data mining process is simpler. Furthermore, Twitter user base is very interesting: there are public figures and various experts, and all their discussions are represented by an hashtag.

Linking user ids on social media and on newspapers, in some cases, may happen because the id or the email is the same on both platforms. This could be a very useful piece of information for our analysis. It is possible that a newspaper can ask the id users on social media of their subscribers to understand what they think about some kind of hashtag or a particular information.

Moreover, users need to give permission to link their social media account with their newspaper one. GDPR legislation allow this typology of data gathering, and the data that we use for our research are the ones that were approved by users.

Thanks to the data analysis, we can measure whether the newspaper is unbalanced in the direction of the public information needs, or if it tries to push a specific group of people more sensitive to a certain argument towards a determined discussion. The data could confirm or deny if the newspaper has succeeded in its purpose.

The dissimilarity between newspaper and social network can be assessed by defining a score measure, named "Points of consideration". For example, if the newspaper push (score 8/10 given by the frequency and the visualization of the article on a definite topic on a precise date) on economic articles and (score 3/10) on justice articles, but the social network asks for (score 8/10 given by the frequency of people discussing on a particular hashtag or a content in a day) Justice articles and (score 3/10) economics articles, this idiosyncrasy can be detected within data correlating newspaper and social network. From the evaluation of people debating over a particular topic in a determinate moment on Twitter, it is possible to calculate the importance of it and its hashtag. This is what Twitter call "Trendtopic. From this principle it is possible to understand the importance of that subject. For newspapers side, it is possible to understand if an article catches the readers' attention and give them the information by, the number of articles present in that newspaper, with how many people have viewed the articles on the same date about that topic. The equivalence of these two results gives the balance between the request for information in social media and the offer of

information from the newspapers, which will provoke the result of the ending the discussion on that trend topic.

Figure 1. An example of how it's possible to figure the "Points of consideration" between article and social media.

Figure 1 shows an example of connection between newspaper and social media data with some "points of consideration".

With a correlation approach, by studying the hashtags and articles shared, it is possible to understand who is acting as a micro-influencer on the social network and to connect it with those who want to augment/increase the popularity of the articles related to some specific argument.

Additionally, it is worth noting that the catalogue of hashtags is constantly growing. So, by moving back in the influence network, it is possible to understand who has shared the first hashtag (correlated to some news) and for which purpose. Also, this catalog will be helpful to understand the starting point of the news sharing process, for instance if some sudden or unexpected event has generated the evolution of a certain hashtag in time.

Moreover, it is important to know the entire set of those who "consume" the article as information on social media, in order to understand when this information has started to spread. Many newspapers can benefit from this service for th**e** decision-making processes, allowing to grow some specific argument-sectors in the newspaper.

2.2. Hashtag and sharing information before Social media

Sometimes if the article is very old and keywords are not created as a hashtag, it could be difficult to understand when the topic started and when those articles became relevant on social media. Consequently, it is necessary to understand which article the editor promoted the most in the past, and how readers had used those articles on social media.

A data mining program can be used to understand article keywords and extract the possible hashtag associated to it. However, to choose the correct hashtag it is necessary a human-supervised action that include an examination of the cultural environment. For instance, the

political culture in the particular time frame when the articles were published could turn to be a crucial variable that can significantly modify the information in time.

As a consequence, a fundamental result of this project could be to understand the evolution of the newspaper by articles publication history. This could be done analyzing user's accounts that had shared the article in the past. Thus, it is possible to analyze the entire history section of the micro-influencers, i.e. the accounts that sponsored the post on Social media, but with some low kind of appearance in social media itself, and to see which person had previously shared the post on his profile.

The micro-influencer and the information analysis will be useful to understand why people on social networks had used some specific information to share ideas or to spread propaganda for his/her favorite candidate/party. By enlarging the perspective of the present work, the results of this analysis can show the evolution of the information over time. This could help other researchers to better understand the reasons for our current social and political condition, consequently the motivation behind certain political actions occurring in such a context.

3. Conclusion

Trend topics on social networks represent nowadays a valuable informative aid to newspaper, that gives the possibility to respond to unconscious requests for information (on a specific topic) on social network, with newspaper information supply.

Thanks to the use of big data, it is possible to evidence the unbalance of information request by people on social network and information supply by newspaper. This dissimilarity can be assessed by defining a score measure, named "Points of consideration", given by the elaboration of inner data from newspaper and patterns model from public data on social network. The equivalence of these two results gives the score balance between the request for information in social media and the offer of information from the newspapers, if in the end, the newspaper will supply correctly the information request with a specific high-quality article that will provoke the full information-satisfaction for people, as consequence there will be the end of the discussion on that trend topic.

References

Gallino Luciano, (2007) Tecnologia e Democrazia – *Biblioteca Einaudi*

Gustav Jakob Petersson & Jonathan D. Breul, (2019) Cyber society, big data, and evaluation. *Routledge editori*.

Lupton Deborah (2018). How do data come to matter? living and becoming with personal data. *Big Data & Society*.

Mazzeo Rinaldi Francesco & Giuffrida Giovanni, (2017) Big data e valutazione: una relazione ancora da costruire. *Rassegna italiana di valutazione* (aiv) n68 anno xxi issn 1826-0713, issne 1972-5027.

Resnyansky Lucy (2019). Conceptual frameworks for social and cultural big data analytics: answering the epistemological challenge. *Big Data & Society*.

Sadowsky Jathan (2019). When data is capital: datafication, accumulation, and extraction. *Big Data & Society*.