On the social perception of science in Twitter: a proposal for a real-time indicator

Elena Denia*

INGENIO [CSIC - Universitat Politècnica de València] / Institute of Innovation and Knowledge Management

Abstract

This work aims a study on the social diffusion of knowledge and citizen participation in the social evaluation of science in the specific area of cosmology. We propose a practical study on Twitter –by means of tools of computational sociology– on public attention (or interest) to a scientific discipline to see if a positive perception is linked to the flow of information of high scientific rigor and /or vice versa.

*elenadenia@ingenio.upv.es

1 Social perception of cosmology

The present study represents an original contribution to the branches of knowledge of *Public Understanding of Science*, *Science Communication* and *Physics and Society*.

On the one hand, it offers the first quantitative analysis in Twitter of the social attention cycles and the dynamics of information propagation of media, institutions and general public restricted to the specific area of cosmology. The distinct kind of temporal variations will be appreciable in terms of the events or discoveries that take place over time in such a field, as well as in terms of those different actors in the system.

On the second hand, this work offers the first quantitative analysis which measures the interaction of social perception with high level scientific information, i.e a study about the impact of science in society in Twitter. Once again, it has been particularized to the case of the astrophysics & cosmology field. Nevertheless, the method has been designed as an original proposal for a new tool for measuring the social perception of science and technology, a social aspect studied extensively over the years by european Eurobarometer surveys, by Fecyt reports in Spain and others.

To do so, in this chapter we will focus on an empirical research based on real-time data extracted from Twitter about the social attention to astronomy topics. Attitudes of communities and individuals toward cosmology achievments, sometimes meaning an acquisition of new scientific information.

The preliminar experimental proposal comes inspired by the works performed in [2] and [4].

Our goal is to answer the following two research questions.

1.1 Research questions

- A favourable perception of achievements in cosmology (in Twitter) is linked to the flow of information of high scientific rigor?
- And vice versa?

1.2 Methodology

Presumably, computational sociology, agent-based modeling and social network analysis may offer proper tools to our research. Our next step will be determine the most suitable computational techniques, for now, outside from the scope of this preliminar proposal.

One of the strengths of this work is that intends to claim the studies on social attitudes through a real time method, not through the answers based on traditional surveys. For instance, the last reveal that 25% of Spaniards believe that the Sun rotates around the Earth, that is to say, one in four is geocentrist[1]. Is that a reliable result? Some scholars attribute it to that it was a question

[1] http://elpais.com/elpais/2015/04/23/ciencia/1429792444_486485.html

Attention	$f_c(\%)$	V	V_{128}	V_p	$\frac{V_P}{V}(\%)$
Media					
Institutions & associations					
Public: scientists					
Public: general					

Table 1: Statistics of the clusters (community subsets).

located at the end of the survey so that respondents could answer while they were tired and not thinking clearly.

1.3 Public attention cycles to cosmology

Our first aim is to study the structure of the network and its evolution, specifically the dynamics of cosmology information in Twitter.

Orientative statistics of the clusters (i.e. community subsets) are shown in the next table, where f_c is the fraction of phrases in the cluster (hashtags and keywords?), V the total volume (over 1 year), V_{128} the volume around the public attention peak (128 hours), V_P : Volume at the peak (1 hour), $\frac{V_P}{V}$ is the peak to total volume.[2]

Let Figure 1 as an example of the temporal variation of hashtag mentions (even if it is outdated or old, there must be noisy background), computed in [4]. In our case, clusters are the four community subsets, providing us information about how is behaving the attention pattern depending on the cluster nature and the cosmological discovery or new, allowing us comparisons between subcommunities and also between events.

For instance, in our representation media attention is expected to present a keen peak associated to an astro new, while general public probably would show a progressive loss of attention over time, perhaps maintaining a 'background noise'. It could be of interest to see how thread volume increase and decay over time. Thread volume in news sources increases slowly but decrease quickly, while for public attention the increase could be rapid and decrease much slower.

Additionally, to study such the background noise mentioned previously could be also of interest for future comparations with other areas of science.

Other expected result is a time lag (hours, days...) for the institutions & organizations attention (for instance, linking to press realease about a discovery in the field), media news and in the last place public attention.

Presumably, could be found potential analogies of natural systems similar dynamics to what one sees in the attention cycles.

It is also of interest to point out the comments in [4] for a model for the news cycle:

> We argue that in formulating a model for the news cycle, there

[2]Time periods are inspired by the work of Yang et al. (2011), but should be adapted to our case once we get the real data.

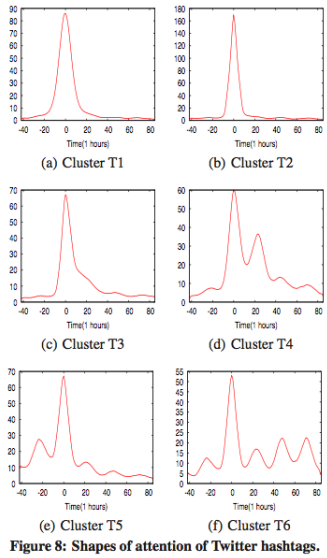

Figure 1: Shapes of attention of Twitter hashtags [4]

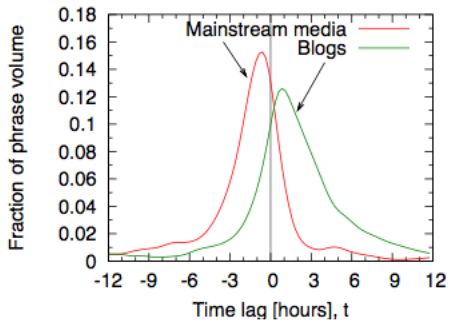

Figure 2: Example: Time lag for blogs and news media. Thread volume in blogs reaches its peak typically 2.5 hours after the peak thread volume in the news sources. Thread volume in news sources in- creases slowly but decrease quickly, while in blogs the increase is rapid and decrease much slower. [3]

> are two minimal ingredients that should be taken into account. The first is that different sources *imitate* one another, so that once a thread experiences significant volume, it is likely to persist and grow through adoption by others. The second, counteracting the first, is that threads are governed by strong *recency* effects, in which new threads are favored to older ones. (There are other effects that can be included as well, including the fact that threads differ in their initial *attractiveness* to media sources, with some threads having inherently more likelihood to succeed. However, we omit this and other features from the present discussion, which focuses on identifying a minimal set of ingredients necessary for producing the patterns we observe.) (Yang & Leskovec, 2011).

Moreover, we may check, for example, whether the theory of the american sociologist Anthony Downs on Issue-Attention Cycle, based on specific issues, can be reflected in our system. In a paper published in 1972 *Up and Down with Ecology: the Issue-Attention Cycle,* Downs states the question why some issues came to merit the interest of policy makers and other do not. He ventured the hypothesis that a particular question, important as it was, attracted the attention of public opinion only for a short period of time. Problems, assures Downs, acquire prominence suddenly and, after a period of time, resolved or not, "[...] gradually fade from the center of public attention."[1]. [3] Problably, such an analysis will be open to interpretation due to the large number of contextual variables and it shall not be fitted to a time-dependent theoretical model of attention.

Until now, we do not know if the comments are favorable or unfavorable, but that scientific information raises an activity or measurable flow. As an analogy, we know that the Voynich manuscript contains a message, we do not know what that message. In the next section we will proceed to decrypt the message.

1.4 Interaction of public perception of cosmology and scientific rigor

This is a study about attitudes and knowledge dissemination.

First of all it is necessary to classify the agents defining the community subsets.

The main parameter 'atenttion' is defined by the community (or sample) identified by keywords and #hashtags. Such a community is formed by different agents as media, institutions & associations, scientists and general public. We must look at the mean and standard deviation of each subset. These are the **control variables**.

Let C be the set generated by Twitter users talking about astronomy in a period of time t.

[3] Please be aware that data miners did not prove that issue-attention cycle exist. http://ae-portal.blogspot.com.es/2009/06/issue-attention-cycle.html

$$C = \{media, institutions, scientists, public\} = \sum_{i}^{4} s_i$$

Once we have defined our control variables (Media; Institutions & associations; Public: scientists; Public: general), we must think about the **dependent variables**. To do so, we propose the following:

- *Attention*: defines the community talking about astrophysics through keywords, #...
- *Perception* (image of cosmology and its progress): identified by opinion mining or manually (the more viable)
  - Favorable (+)
  - Hostil/resistances (-)
  - Neutral (*)
- *Scientific rigor*: identified by computational approaches and based on sharings, likes and content as follows:
  - Mild: based on sharings and likes
  - Soft: sharings with scientific content behind (general media)
  - Strong: sharing with hard scientific content behind (high level outreach and links to scientific articles)
  - Lack: pseudosience content behind

We must pay special attention to the pseudoscience contribution, whether it is significant or not poses an appreciable contribution to the perception (+,-,*) measurements.

The following step is more tricky, and the proposed approach comes inspired by the structure of the work 'Public attention to science and political news and support for climate change mitigation', published in the journal *Nature Climate Change* [2].

Table 2 shows public attention to cosmology issues of different community subsets and their perception computed by opinion mining techniques differenciated as positive, negative or neutral. We must pay particular attention to the uncertainty provide by the chosen technique (to be determined, dependent on the technique).

Table 3 shows again the public attention to cosmology issues of the different community subsets and their grade of interest in terms of the scientific rigor behind the shared information. Scientific rigor is considered *mild* when users retweet othe users talking about astropysical issues or like tweets about this

	Dependent Variables			
	Attention	Perception (image)		
		Positive	Negative	Neutral
Media				
Institutions				
Public: scientists				
Public: general				

Table 2: Subsets and perception of cosmology issues

	Dependent Variables				
	Atenttion	Rigor			
		Mild	Soft	Strong	Lack
Media					
Institutions					
Public: scientists					
Public: general					

Table 3: Community attention subsets and information in cosmology issues

topic; *soft* when users actively tweet links to scientific content as blogs, newspapers, press releases, etc.; *strong* when users share hard scientific content like scientific articles; and *lack* when users include links to pseudoscience content.

After this results, we must estimate if a favourable perception of achievements in cosmology is linked to the flow of information of high scientific rigor and/or vice versa by crossing *Percepcion* with *Scientific rigor*.

1.5 Conclussions

The most interesting contribution relies on the interaction of social perception of a science discipline with high level scientific information shared on Twitter. So that, this work may be considered as a innovative proposal for a real-time indicator of the social perception of science in Twitter.

The study is opened to a deeper analysis concerning to the content of the tweet, since the linked content may have different formats and sources (or channels) like papers, blogs, media news, videos, etc.

It would allow us to make considerations about the social image of science and technology, considering different factors, agents, impact and other variables.

References

[1] A Downs. Up and down with ecology-the issue-attention cycle. *The public interest*, 1972.

[2] P. Sol Hart, Erik C. Nisbet, and Teresa a. Myers. Public attention to science and political news and support for climate change mitigation. *Nature Climate Change*, 5(JUNE):541–545, 2015.

[3] Jure Leskovec, Lars Backstrom, and Jon Kleinberg. Meme-tracking and the Dynamics of the News Cycle. *Proceedings of the 15th {ACM} {SIGKDD} International Conference on Knowledge Discovery and Data Mining*, pages 497–506, 2009.

[4] Jaewon Yang and Jure Leskovec. Patterns of Temporal Variation in Online Media. *Time*, 468:177–186, 2011.