

Public opinion on OER and MOOC: a sentiment analysis of twitter data

URL	http://weko.wou.edu.my/?action=repository_uri&item_id=52
-----	---





Abeywardena, I.S. (2014). Public Opinion on OER and MOOC: A Sentiment Analysis of Twitter Data. *Proceedings of the International Conference on Open and Flexible Education (ICOFE 2014), Hong Kong SAR, China.*

Public Opinion on OER and MOOC: A Sentiment Analysis of Twitter Data

Ishan Sudeera Abeywardena

School of Science and Technology, Wawasan Open University, Penang, Malaysia
ishansa@wou.edu.my

Abstract

The Open Educational Resources (OER) movement has gained significant momentum recently as a global effort culminating in the 2012 Paris OER declaration. However, the purist definition of OER has blurred since then morphing into Massive Open Online Courses (MOOC). Even though OER are a significant part of the MOOC movement, it might not be a defining one. However, this has not yet been fully verified with respect to the opinion of the general public who are the main stakeholders of both the movements. To answer this question, this paper attempts to explore the public opinion and perceptions regarding OER, MOOC and their complementary roles. A text mining approach is used to analyse raw Twitter data in the domains of OER and MOOC within a timespan of 12 months. Sentiment analysis is applied to the data to understand how public perceptions have changed during this time period. The major contribution of my paper is a chronological view of public opinion on OER and MOOC post Paris OER declaration.

Keywords: open educational resources, OER, MOOC, text mining, opinion mining, sentiment analysis

1. Introduction

Arguably, the months since the Paris Open Educational Resources (OER) Declaration (UNESCO, 2012) are the most crucial in terms of the future direction of the whole OER movement. Many new OER initiatives have blossomed since addressing the 10 key recommendations made to policy makers. On the other end of the spectrum, Massive Open Online Courses (MOOC) have gained momentum as an innovative way of increasing access and equity in education. McAuley, Stewart, Siemens, & Cormier (2010) claim that

“The large scale of the community, from several hundred to several thousand participants, maximizes the possibility that the “long tail” effect will enable someone with even the most esoteric interests within the overall focus of the MOOC to find people with whom to share and collaborate”.

Thus, MOOC have wide implications on how education will be perceived in the future. Although OER play a major role in the delivery of MOOC, it might not be a defining one. As such, the two camps are partially divided when it comes to which is the best way forward. Affirming this, Daniel (2012) argue that

“...what MOOCs will not do is address the challenge of expanding higher education in the developing world. It may encourage universities there, both public and private, to develop online learning more deliberately, and OER from MOOC courses may find their way, alongside OER from other sources, into the teaching of local institutions”.

This division results in confusion among the stakeholders as they are now faced with a difficult choice in terms of organisational, institutional or national policy. In an attempt to identify the extent of this dilemma, this paper looks at the sentiment of the public (stakeholders) with respect to OER and MOOC. To achieve this objective, a text mining approach is used to extract data from social media for sentiment analysis.

The major contribution of this paper is a chronological view of public opinion on OER and MOOC for a timespan of 12 months post Paris OER declaration. Through this view, a roadmap can be identified for future research and development based on public demand. This is the major advantage of the preliminary findings presented.

The remainder of the paper is divided into methodology, results, discussion and conclusion.

2. Methodology

At present, *Twitter*¹ is one of the largest social networking platforms in existence used by over 230 million monthly active users (Twitter, 2013). The platform allows users to share concise (140 characters) posts in the form of microblogs called *tweets*. These tweets are seen by selected individuals, groups or publically by the whole Twitter network depending on the social status set. Following a tweet, the social interaction takes place in the form of response tweets or as *retweets* where the initial tweet is shared virally. This rich interaction results in a large number of tweets being generated on a particular topic.

It is observed that the OER and MOOC communities actively use social media such as Twitter for teaching and learning purposes (McAuley, Stewart, Siemens, & Cormier, 2010). Furthermore, due to the unpoliced and liberal nature of this social networking platform, frank opinions are shared by users on the benefits as well as the shortcomings of both the ideologies. As such, Twitter was found to be the ideal data source to analyse the public opinion regarding OER and MOOC.

The extraction of tweets was done manually using the native search mechanism of Twitter². The search terms ‘OER’ and ‘MOOC’ were used in the search. Only the abbreviated forms were used in the search as I had identified empirically that the terms ‘Open Educational Resources’ and ‘Massive Open Online Courses’ were seldom used in tweets. This is mainly due to the 140 character restriction of a tweet. Only the ‘Top’ tweets were extracted as these are the best matches for a given search query. Furthermore, the use of the ‘Top’ tweets acted as a first level cleansing of the dataset. The raw data was organised using the MS Excel 2010 spreadsheet application. A second level cleansing of the dataset was conducted manually to remove partial, irrelevant and repeated tweets. Only distinct tweets were used. All user details were removed to ensure anonymity and unbiased analysis. The details of the final dataset are provided in Table 1.

Table 1 The Twitter dataset used in the study consisting of tweets which include the terms ‘OER’ and ‘MOOC’.

Search Query	Tweet History	Timespan	No. of Distinct Tweets
OER	1/11/2012 – 31/10/2013	12 months	1209
MOOC	1/05/2013 – 31/10/2013	6 months	2823

The sentiment analysis was done using the *Semantria*³ software application which comes in the form of a plugin for the MS Excel spreadsheet application. In preparation for analysis, an identity column was added to the dataset to enable the analysis of individual tweets with respect to sentiment. A basic sentiment analysis was conducted on the dataset using the Semantria plugin. The plugin uses a cloud based corpus of words tagged with sentiments to analyse the dataset. Through statistical inferences, each tweet is tagged with a numerical sentiment value ranging from -1.5 to +1.5 and a polarity of (i) negative; (ii) neutral; or (iii) positive. The positivity increases with the sentiment value.

The sentiment data is then reorganised according to individual months of the tweet history to identify the public opinion on a particular topic over a given timespan.

3. Results

The analysis of the data looks at two major aspects which are (i) the number of tweets over a given timespan; and (ii) the public sentiment over a given timespan. Figure 1 and Figure 2 show the number of tweets on ‘OER’ and ‘MOOC’ respectively. Figure 3 shows the public opinion on ‘OER’ over a 12 month timespan whereas Figure 4 shows the public opinion on ‘MOOC’ over a 6 month timespan.

¹ <http://www.twitter.com>

² <https://twitter.com/search-home>

³ <https://semantria.com/>

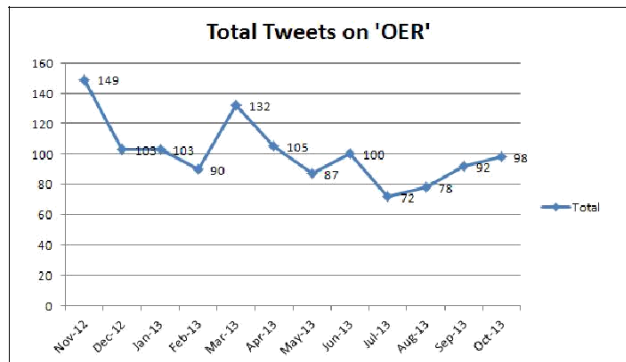


Figure 1 Total tweets on ‘OER’ for a timespan of 12 months from November 2012 to October 2013.

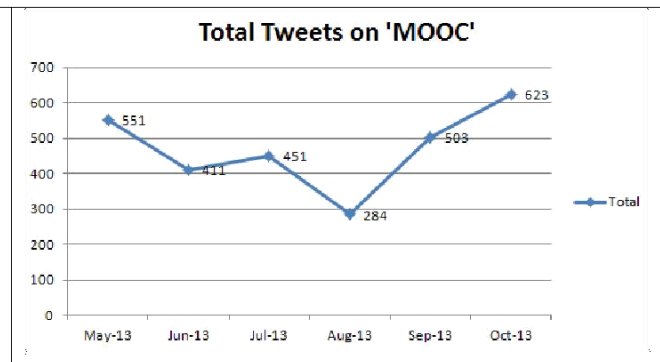


Figure 2 Total tweets on ‘MOOC’ for a timespan of 6 months from May 2013 to October 2013.

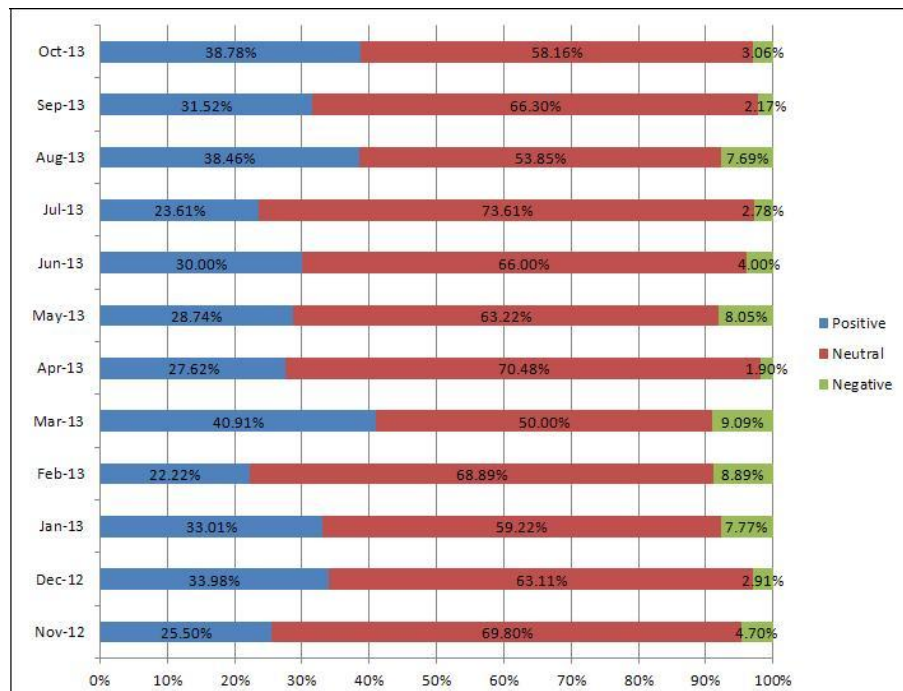


Figure 3 Public opinion on ‘OER’ over a 12 month timespan.

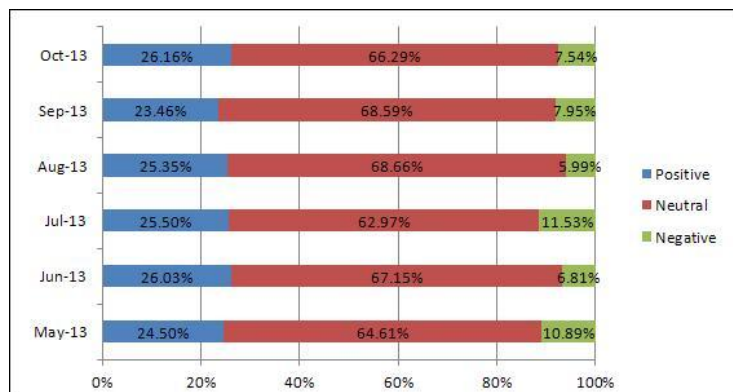


Figure 4 Public opinion on ‘MOOC’ over a 6 month timespan.

Figure 5 provides a comparison between the numbers of tweets on 'OER' vs. 'MOOC' for a 6 month timespan between May to October 2013. Figure 6 depicts the change in public opinion on 'OER' vs. 'MOOC' for the same timespan.

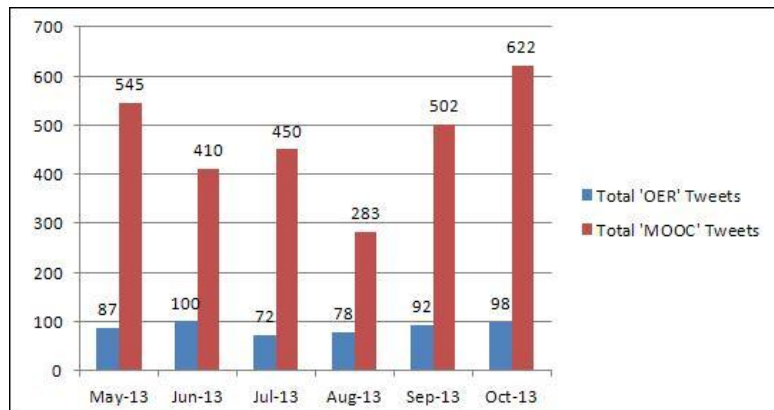


Figure 5 Comparison between the numbers of tweets on 'OER' vs. 'MOOC' for a 6 month timespan between May to October 2013.

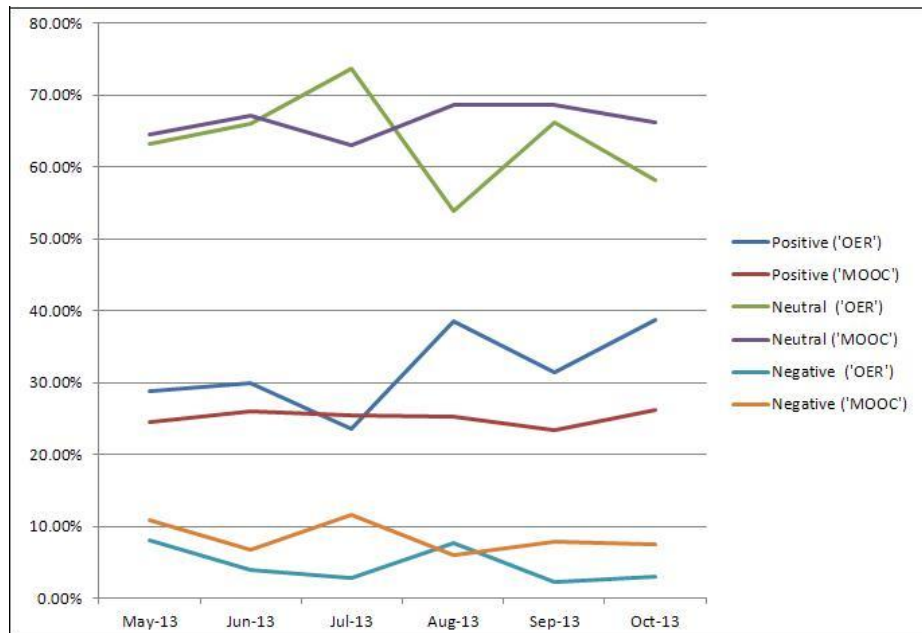


Figure 6 Change in public opinion on 'OER' vs. 'MOOC' for a 6 month timespan between May to October 2013.

4. Discussion

With reference to Figure 5, it can be seen that there is relatively more discussion taking place on the topic of MOOC in comparison to OER. As shown in Table 1, the total number of distinct tweets on OER for a timespan of 12 months is only 42.8% of the total number of distinct tweets on MOOC within a time span of six months. Furthermore, Figure 1 and Figure 2 suggest that the interest in OER is on a downward trend whereas the interest in MOOC is on the rise. There may be several factors contributing to this trend. Among them, the novelty of MOOC, the involvement of the private sector, the brand names associated with the recent MOOCs delivered, the keen interest of highly reputed conventional institutions in the concept and the large marketing budgets could be key influences.

Although the interest in OER seems to be declining, Figure 3 suggests that the public opinion on OER still remains largely positive. By comparing the averages of the positive values shown in Figure 3 and Figure 4, it can be seen that the average positive public opinion on OER is 31.19% in comparison to 25.16% on MOOC.

Considering the timespan of 6 months from May to October 2013, the change in public opinion on OER and MOOC are highlighted in Figure 6. Even though the numbers of tweets on MOOC are considerably larger than on OER suggesting more interest in the former, public opinion on MOOC seem to remain unchanged with respect to the positive impact it has. In contrast, the stakeholders seem to feel more positive about OER and the benefits it has. Furthermore, this positivity towards OER seems to be on an upward trend.

When considering the negative public opinion on OER and MOOC, Figure 6 suggests that the negativity towards both the concepts are reducing. However, the negativity towards OER is slightly less than towards MOOC. The more interesting trend is the change in public neutrality towards OER and MOOC. Figure 6 shows only a slight change in neutrality towards MOOC whereas there is a decline in the public neutrality towards OER. This suggests that the public are now forming informed opinions regarding OER whereas caution and skepticism still surrounds the relatively new concept of MOOC.

The strength of this study is that it takes an objective look at both MOOC and OER from a public's perspective. In contrast to an instrument driven survey which would restrict the expression of opinion by the public to a certain reference frame, the use of sentiment analysis takes into consideration all that the public has said and their relative connotations. Although a considerably large number of tweets have been analysed, it can still be argued that the sample size is not sufficient to generalise any trends. This is one of the weaknesses of the study. However, the advantage is that this model can be replicated using a larger sample size when the data is available in the future.

5. Conclusion

The debate on Open Educational Resources (OER) vs. Massive Open Online Courses (MOOC) has been raging for the past couple of years. Advocates of the two camps have been bombarding the public with the benefits of these ideologies. Given that both OER and MOOC are pervasive forces which can change the education landscape of the future, the general public (stakeholders) is in a state of confusion with respect to adopting a particular ideology for their institution, organisation or nation. This study looks at OER and MOOC from the public's perspective. It uses sentiment analysis of Twitter data to approximate the public opinion on OER and MOOC in terms of the benefits they promise.

The analysis of the twitter data suggests that there is growing interest on MOOC in comparison to OER. However, when considering the public opinion on MOOC, it is apparent that the stakeholders haven't yet formed strong opinions on MOOC due to it being a relatively recent phenomenon. In contrast, the positivity towards OER is on an upward trend. The neutrality towards OER has also decreased over a time span of 12 months post Paris OER declaration suggesting that the public is slowly but surely forming informed opinions about the use of OER.

The contribution of this paper is a chronological view on how the interest and public opinion has changed with respect to OER and MOOC following the 2012 Paris OER declaration. Three main trends can be identified which need to be taken into consideration by the stakeholders when making a policy decision: (i) there is an increasing amount of interest on MOOC; (ii) the public still hasn't formed strong opinions regarding MOOC due to the novelty of the ideology; and (iii) the positivity towards OER is growing.

This paper discusses the preliminary analysis of the data. It is my intention to further probe the data and identify the reasons behind the trends identified in this paper. The next step is to potentially pinpoint certain phenomena which would have a positive or negative impact on the public opinion with respect to MOOC and OER.

6. Acknowledgements

I acknowledge the support provided by Dr K S Yuen, Chair of the conference organising committee; Dr K C Li, Vice-chair of the conference organising committee; and the Open University of Hong Kong (OUHK) by extending full sponsorship for my participation at the Inaugural International Conference on Open and Flexible Education (ICOFE 2014).

7. References

- Daniel, J. (2012). Making sense of MOOCs: Musings in a maze of myth, paradox and possibility. *Journal of Interactive Media in Education*, 3.
- McAuley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). *The MOOC Model for Digital Practice*. Retrieved 12 15, 2013, from http://www.elearnspace.org/Articles/MOOC_Final.pdf
- Twitter. (2013). *About Twitter, Inc.* Retrieved 12 15, 2013, from Twitter.com: <https://about.twitter.com/company>
- UNESCO. (2012, June 22). *2012 PARIS OER DECLARATION*. Retrieved June 13, 2013, from [unesco.org: http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/WPFD2009/English_Declaration.html](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/WPFD2009/English_Declaration.html)