Measuring the return on communication investments on social media:

The case of the higher education sector

Luciana Oliveira
CRACS / INESC TEC & CEOS.PP Polytechnic of Porto
Rua do Campo Alegre 1021/1055, Porto, Portugal
lgo@eu.ipp.pt

Álvaro Figueira
CRACS / INESC TEC & University of Porto
Rua do Campo Alegre 1021/1055, Porto, Portugal
arf@dcc.fc.up.pt

Abstract—**Measuring the return on communication investments on social media has become one of the top key issues for organizations joining social networks. However, this field has been lacking articulation between what is conveyed as social media key performance indicators and the alignment of strategic organizational goals. Therefore, we propose a methodology to measure the performance of each organization on social media, to determine their positioning in the sector and to evaluate which are the content strategies used to boost the highest performing organizations. Thus, we identify how to determine which organizations should be closely monitored within the sector and which type content strategies can foster higher organizational performance on social media.**

Keywords—social media; content strategy, analytics; organizational performance; benchmarking; positioning; higher education; semi-supervised learning; text mining; categorization

I. INTRODUCTION

Social media performance measurement and benchmarking has been on the top key concerns of all organizations adopting social networks as communication channels and interaction environments. However, most organizations have been adhering to the trend of creating a social presence in multiple networks [1-5], without previously defining a clear strategic approach in line with the overall business objectives. This has been posing real challenges for organizations when trying to determine what and how to measure their performance and return on investment (ROI) on social media.

Regardless of the popularity that social media has been gaining in academic research, there are very few contributions related to social media content analysis in the Higher Education Sector (HES) [6-8] and, to the best of our knowledge, none related to a combined benchmarking analysis of performance and strategic editorial content.

According to Porter [9] a successful firm is one with an attractive relative positioning, which should translate as an outcome and not as a cause. The author also states that, competitive advantage is attained within some scope, and the choice of scope is a central one in strategy. The main goal of benchmarking the organization's performance is to seek to improve internal practices on a regular basis, by continuously developing inter-organizational learning processes. Therefore, to gather meaningful learning from other organizations, whose performance is superior, it is fundamental to investigate and understand 'how' superiority was attained.

In this paper, we present a methodology for retrieving and categorizing Facebook posts in strategic editorial areas as well as three main groups of procedures and sector's results aimed at benchmarking organizational positioning, performance and content strategies.

II. SCOPE AND METHODOLOGY

The research scope comprises one full academic year (twelve months) of social media communication messages of 43 Portuguese Higher Education Institutions, in the most used social network in this sector, which is Facebook.

Using automatic data retrieval, 15,444 Facebook posts were collected using the Facebook's API, during 1st September N and 31st August N+1. The posts were then automatically categorized into one of the seven editorial areas of the editorial model proposed in [10], in a semi-supervised learning process supported by the use of five text mining classifiers, as proposed in [11]: Support Vector Machines [12], Random Forests [13], LogiBoost [14], K-Nearest Neighbours [15], and MultiLayer Perceptrons [16].

We decided to adopt this procedure because automatic categorization provides a highly scalable solution for real-time monitoring and it has been proved to be at least as good as the resulting agreement of a set of five human communication specialists. It is also important to note that the editorial model is comprised of seven strategic editorial areas (education, research, society, identity, administration, relationship and information), aimed at fostering organizational goals and, therefore, provide the main domains in where social media communication can boost tangible and non-tangible organizational performance.

After the retrieval, cleaning and classification of social media messages, a set of analytics and diagnose tools were developed to provide knowledge on each of the agents positioning in the sector, its performance throughout the schoolyear and the adopted content strategy.

III. RESULTS

In this section, we present both the tools that we developed to provide measurement of the communication investments on social media, as well as the core knowledge outputs that we were able to generate.

A. Relative positioning analysis

To determine the positioning of each agent in the sector we used two main variables: x "effort' (total number of messages posted during the analyzed period) and y "response" (an weighted sum of total amount of interactions that each agent has gathered during the same period and used as a score function [10], in which each type of interaction has a distinct weight: 'likes' weigh 0.5 points, 'comments' weigh 0.8 points and, 'shares' weigh 1 point. It is important to highlight that each agent's relative positioning corresponds to 'where' he is in relation to the other agents at 31st August N+1.

A perceptual map, based on the (x,y) variables, composed by four quadrants was also built and used to identify efficient, advanced, incipient and inefficient agents.

Using these methods and tools we identified one advanced agent (highest "effort" and highest "response") in the sector and a remaining group of 42 incipient agents (very low "effort" and "response"), with ten times higher values on both variables in relation to the first. This scenario led us to conduct additional analysis on the relative positioning of agents in the sector, where we considered that one single advanced agent as an outlier. Therefore, we deepened the analysis by recomputing only the positioning of the agents in the incipient quadrant, to identify, among these, which could, additionally, be considered as having a good relative performance positioning. As a result, we identified four incipient-efficient and two incipient-advanced. The great majority of agents remained in the incipient quadrant, leading us to believe that the sector is mainly composed by early social media adopters, to whom interorganizational learning might be crucial to improve performance.

B. Performance analysis

After identifying the relative positioning of agents at the end of the period of analysis, we computed their initial positioning (using the data from the first month of the period of analysis: 1st to 31st September N) and we also tracked which was the performance trajectory on that (x,y) plane that each agent has undergone, by computing the trajectory's angle. Thus, the developed performance trajectories were composed of a departure point and of an arrival point for each agent (two-point segment angle).

In this performance analysis, we were able to identify positive and negative trajectories' angles. This means that all throughout the entire school year some agents have positively evolved, either more on "effort" than on "response" (or the opposite) or have regressed, also either more on "effort" than on "response" (or the opposite).

We were also able to observe that the positive and negative variations of the trajectories (two-point segment) occurred in very distinct ratios for each agent. For this reason, the point of

departure of the trajectories was normalized, i.e., the vectors were brought down to (0,0).

This method provided us with a clearer visualization of the angle of the performance trajectories (two-point segment), as well as their length, both in a positive and negative axis. At this point, we were able to conclude that around 70% of the trajectories revealed some form of positive performance and only around 30% of the trajectories segments were effectively expressing a regression from the origin to the 3rd quadrant.

Still, at this point, it was not possible to accurately report on how positive was a specific positive performance trajectory, nor on how negative was a negative regression trajectory. This raised the research and organizational problem of not being able to establish the needed comparisons among organizations, sustained on solid statistical and analytical arguments.

Therefore, to overcome this limitation, we designed a diagnose tool that consists of a circumference, where 8 zones of 45° intervals between 0° and 360° are identified as independent diagnosable areas. Using this tool to overlap the great diversity of performance trajectories' angles we could create a quick visual aid for diagnosing each agent's performance.

To each one of the 8 zones of the diagnose circumference a small description of the past, current and future outcome / return in terms of social media performance was elaborated. For instance, in the zone from angles 0° to 45° performance trajectories are both growing and developing within a healthy ratio of "effort" and "response". However, in terms of future projections, the organization should be advised to seek to improve their content strategy in the short term, in order to move the angle of its trajectory to values above 45°, where, definitely with less "effort" the amount of "response" continues to increase. In this way, sustainability and return would be higher.

Within the diagnose circumference we have also identified the two best opposite positive and negative performance angles: angle 135° as being the most favorable for any organization, and angle 315° as being the worst angle possible. By doing so, organizations are provided with a sense of a to-reach top performing trajectory, as well as to-avoid direction in terms of diagnose.

From our analysis, and considering the initial set of 42 agents, we were able to identify that 17 organizations had a good positive performance trajectory, 4 had a very good performance trajectory, 18 agents presented acceptable trajectories that need intervention on the short term, i.e., need to improve their content strategy; and finally, we detected 3 very bad performance trajectories belonging to organizations that need immediate intervention on their social media content strategy.

C. Editorial benchmarking

The final step of the methodological procedures that we considered as necessary to provide a sustainable way to measure the return on communication investments on social media, consisted of conducting editorial benchmarking. Editorial benchmarking consists of identifying the best

practices in the sector, in this case, in terms of content strategies, and to gather knowledge that other organizations can use to improve their performance, according to their own set of internal present and future organizational goals.

Since we conducted positioning analysis and performance analysis, we believe that the best way to gather the best practices within the sector would be to join together the two sets of the top leading agents that we have separately identified. By doing so, we are not only reuniting the capital from those who, at the end of the period of analysis, had the best relative positioning, but also those who had the best performance thorough the entire year. Additionally, since for benchmarking purposes an outlier might be a highly valuable resource to learn from, and not just an atypical value or experimental failure, we decided to include the agent that we have previously mentioned and who outperformed every other agent in the sector, by far.

Naturally, by bringing these two sets together, we ended up finding commonalities, i.e., agents that belong to both sets, which provides us with confidence on the analyses that we conducted: it would only be natural for some of the best well positioned agents to have had some of the best performance trajectory angles. However, we found 3 exceptions: on the one hand, two of the best well positioned agents did not have one of the best trajectory angles, on the other hand, we discovered that one of the agents with one of the best performance trajectory angles ended up not being among the top best positioned agents at the end of the period of analysis.

However, for the purpose we have set out to accomplish, we decided to consider only the agents in common on both of the sets of top agents together with the former outlier, ending up with a total of unique 7 agents, which we consider to be the "cream of the crop" to learn from.

IV. CONCLUSIONS AND FUTURE WORK

We have presented a set of methodologic procedures and three main groups of analyses aimed at benchmarking organizational positioning, performance and content strategies. Following the proposed benchmarking methodology and comparative views, organizations can identify its sector positioning, assess their performance on Facebook and identify and evaluate the type of content areas that might be worth pursuing to better attain both social media and organizational goals.

Recapturing the seven editorial areas that we previously mentioned, and that are composed by seven strategic areas for Higher Education Institutions (education, research, society, identity, administration, relationship and information), we noticed that the main content areas pursued by the set of 7 agents consists of the areas 'Identity", "Education", "Research" and "Society", in this order of importance.

The editorial area "Identity" is, by far, the most prominent area in of the content strategies, which comprises institutional promotion, advertising (identity, image, reputation), or any other form of organizational self-promotion. The editorial area "Education" is the second content area most prone to generate "response" from audiences, and it consists on promoting / advertising higher education courses and / or additional / complimentary training provided by the organization. The third most relevant area, "Research", includes informing on and / or posting calls for participation in: congresses, seminars and other scientific meetings; and referring to internal and external research results and awards (papers, articles, books, proceedings, etc.). Ultimately, indirectly, advertising the organization's scientific capital contributes to build and enhance its reputation. The last most relevant content area, "Society", includes a variety of items that are aimed at reinforcing the bonds between the organization and its surrounding environment, namely through establishment and promotion of partnerships, contracts, knowledge and technology transfer, and promotion of student's employability. In the HES the input provided by this content area is essentially pursued by posting about job offers (national and international) in the organizations' core training areas, thus gathering enormous amounts of response among students and alumni, either through Facebook's 'likes', 'shares' and 'comments'.

As an output of the benchmarking analyses conducted in previous subsections, we have revealed how to identify top performing agents' worth monitoring and learning from and the editorial areas that stand out among the combined set of top performance agents: "Identity", "Education", "Research" and "Society". We believe that the proposed methodology and diagnoses are transferable to other organizational / business / industry sectors, considering the corresponding adaptation of the editorial model to the strategic areas of communication / intervention with the most potential to improve the overall organizational performance.

REFERENCES

[1] Fournier, S., et al., The uninvited brand. Business Horizons, 2011. 54(3): p. 193-207.

[2] Gensler, S., et al., Managing Brands in the Social Media Environment. Journal of Interactive Marketing, 2013. 27(4): p. 242-256.

[3] Hanna, R., et al., We're all connected: The power of the social media ecosystem. Business Horizons, 2011. 54(3): p. 265-273.

[4] Kaplan, A.M. and M. Haenlein, Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, 2010. 53(1): p. 59-68.

[5] Kirtiş, A.K. and F. Karahan, To be or not to be in Social Media arena as the most cost-efficient marketing strategy after the global recession. Procedia-Social and Behavioral Sciences, 2011. 24: p. 260-268.

[6] Chauhan, K. and A. Pillai, Role of content strategy in social media brand communities: a case of higher education institutes in India. Journal of Product & Brand Management, 2013. 22(1): p. 40-51.

[7] Lai, L.S. and W. To, Content Analysis of Social Media: A grounded Theory Approach. Journal of Electronic Commerce Research, 2015. 16(2): p. 138.

[8] Qi, B. and L. Mackie. Utilising Social Media Technology to Raise Brand Awareness in Higher Education. in WEBIST 10th International Conference on Web Information Systens and Technologies. 2014.

[9] Porter, M., Towards a dynamic theory of strategy. Strategic management journal, 1991. 12(S2): p. 95-117.

[10] Oliveira, L. and Á. Figueira, Social Media Content Analysis in the Higher Education Sector: From Content to Strategy. International Journal of Web Portals (IJWP), 2015. 7(2): p. 16-32.

[11] Figueira, Á. and L. Oliveira. Analyzing Social Media Discourse - An Approach using Semi-supervised Learning. in Proceedings of the 12th International Conference on Web Information Systems and Technologies. 2016. Rome, Italy: ScitePress.

[12] Crammer, K. and Y. Singer, On the algorithmic implementation of multiclass kernel-based vector machines. The Journal of Machine Learning Research, 2002. 2: p. 265-292.

[13] Breiman, L., Random forests. Machine learning, 2001. 45(1): p. 5-32.

[14] Friedman, J., T. Hastie, and R. Tibshirani, Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). The annals of statistics, 2000. 28(2): p. 337-407.

[15] Altman, N.S., An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician, 1992. 46(3): p. 175-185.

[16] Freund, Y. and R.E. Schapire, Large margin classification using the perceptron algorithm. Machine learning, 1999. 37(3): p. 277-296.