



# Complementing consumer input through text mining social media tools- An exploratory case study

Muhammad Tariq Yousafzai<sup>1,\*</sup> and Syed Gohar Abbas<sup>2</sup>

<sup>1</sup>FAST-National University of Computer and Emerging Sciences, Peshawar Campus, K.P.K, Pakistan (25000).

<sup>2</sup>Magellan Research Centre, IAE Lyon, Université Jean Moulin, France.

## ARTICLE INFO

### Article history:

Received: 18 July 2013;

Received in revised form:

29 September 2013;

Accepted: 3 October 2013;

### Keywords

Social Media Platforms,  
Text Mining,  
New Product Development Process,  
Consumer Insights,  
Discovery in a text.

## ABSTRACT

The paper studies in tandem, the application of Text-Mining to Social Media platforms such as Social Networking Sites and Blogs. The research design incorporates triangulated use of a single case study approach to figure out through cluster analysis whether consumer insights can be captured for an organization through the use of Text-Mining software. The research argues favorably in theory and reveals numerically that through Text-Mining Software we can acquire novel consumer insights for the new Product Development (NPD) Process. The research tries to open a new avenue for market researchers to collect consumer insights for forming and revising their social media campaigns. The research leaves issues such as complete applicability and appropriate choice of software for further studies.

© 2013 Elixir All rights reserved

## Introduction

"Markets as conversations" is concept which has its roots in historical time when people exchanged goods and services in a mutually satisfying manner. As the world economy becomes more sophisticated year by year, the marketing concepts are likewise, becoming more sophisticated. With the advent of social media the traditional viewpoint of consumer voice has been made more relevant due to its inherent ability of reaching millions of people in a relatively short span of time.

The rise of social media can be seen from the fact that during the last few years, words from social media landscape have entered into the lexicon as 'Word of the Year' such as Blog in (2004), Podcast in (2005) and Facebook in (2007).

According to Dave Evans (2008) in social media platform conversations between two persons can reach a significant number of (n=2)90th persons. This is indeed a great improvement when compared with traditional media where the conversations can reach a square number of participants (n<sup>2</sup>). Social media is interactive in nature and relatively a new concept. Likewise, the concept of text mining is also a new concept which warrants an in-depth study.

The technology used to elicit information over the World Wide Web is known as "text mining". According to Henning et al, (2004) electronic Word-of-Mouth (eWOM) refers to any statement by a customer regarding a firm or product which is spread across a number of users through the internet. The concept of text mining is distinct from Netnography which is a qualitative method that make use of optimized ethnographic research techniques to analyze the social media in online communities such as Google/yahoo groups etc. In crux, the scope of topic coupled with the text mining tool software available at a Micro Firm in Pakistan inspire us to choose a Text mining technique to explore social media for the end-user inputs.

## Purpose of the Study

The researcher has adopted an ambitious approach to find out the exploratory purpose of the paper to collect consumer input from an organizational view point. To achieve the purpose of the paper the following research question has been formulated:

**RQ:** Can we extract consumer input for an organization through text mining?

## Significance of the Study

Gathering relevant information from social media online communities in its various forms can be intriguing to diverse spectra of audience ranging from marketing researchers, New product designers and innovators of new products and services in the domain of technological landscape. By capturing multifaceted and continuous feedback through text-mining these can significantly raise the product quality and hence the consumer satisfaction. This undertaking is targeted for organizations interested in adopting social media marketing bandwagon and social media scanning services can also benefit from this research.

## Literature Review

According to Dave Evans (2008) social media is, "The democratization of information, transforming people from content readers to content publishers. It is the shift from the broadcast mechanism to a many-to-many model, rooted in conversations between authors, people and peers"

Due to the inherent changes brought in by the spread of internet, marketing has undergone several changes particularly in the new social media landscape. As a consequence today, marketing practices have become more consumer dominant as evident from the prophetic definition of Dave Evans (2008) cited above for public consumption. The authors have drawn upon their earlier research on social media practices for their classification of social media in the figure (2.1) below.

Classification of Social Media Types by Kaplan and Haenlein (2010)		Social Presence / Media Richness		
		Low	Medium	High
Self-Presentation/ Self Disclosure	High	Blogs	Social Network Services	Virtual Social Worlds
	Low	Collaborative Projects	Content Communities	Virtual Game Worlds

Source: Adopted from Kaplan and Haenlein (2010)

### Blogs

In the year (1997) Peter Merholz first introduced the word Blog and soon after in (2004) it became Word of the year by the Merriam Webster Dictionary owing to its increased popularity. Blogs represents a type of website which is constantly updated and entries are date stamped accordingly. In the domain of blogosphere a typical Blog may allow the opportunity to comment for its viewers which may take the form of *praise*, *critique* or *suggestions* (Magnuld & Faults, 2008).

### Social Networking Sites

Social Networking Sites (SNS) help in maintaining and facilitating relationships through effective use of technologies such as a system (SaaS) which include but is not limited to Facebook, You tube, Flickr, Linked in, Twitter and so on (Bernoff, 2008).

### Content Communities

These communities facilitates dispense of media content among users. The media served on the communities include text, videos, photos and PowerPoint presentations. However, Kaplan and Haenlien (2010) proclaim that these communities are prone to infringement of copyrights.

### Virtual Game World

According to popular belief Virtual Game worlds represents a true manifestation of social media marketing. These also represent a virtual three dimensional replica landscape which is visible in the form of avatars which interact with others as would human do in a real life scenario.

### Virtual Social Worlds

These social worlds allow the users to change their behaviors more frequently and are virtually similar to their daily life conduct. This kind of platform allows more interactions such a self representation such as in the Second Life Application (Kaplan & Haenlien 2010).

### Collaborative Projects

This type of user-generated content allows for concurrent creation of content by a multiplicity of authors. These types of collaborative projects include Wikis, bookmarking application, Encyclopedia and so on.

### Text Mining

In order to form a foundation for the subsequent investigation it would logical to touch upon slightly over the concept of text mining and its relevance in the current scenario. Text Mining as the name depicts is a particular area of data mining which can be defined as the process of eliciting relevant information from a platform such as social media. According to Barbier and Liu, (2011) the underlying idea behind text mining is to figure out new relevant information which is latent in the clutter of wide spread information available over World Wide Web.

### Social Media Marketing motives from text mining perspective

According to Ulrich and Eppinger (2003) New Product development (NPD) process generally consist of a sequence of idea conception, development, design, testing and

commercialization. However, firms may use a variety of approaches to NPD like concurrent or overlapping stages of NPD. Customer insights are taken to consideration in a multiplicity of ways to shrink the NPD process so as to offer disclosure of products and services ahead of competitors. Urban and Hauser (1993) contends that firms also make effective use of quality deployment function (QFD), product testing, prototype development and test launching to arrive at improved marketing campaigns for new product introductions. On the other hand, some authors such as Ciccantelli and Magidson (1993) contends that consumers have little or no idea about the future offering of firms therefore, the firm cannot truly incorporate consumer insights when developing new products and services. Some critics, on the other hand, believe that in consumers' goods manufacturing the end-user information is critical for success of new products.

Caimcross, (1997) contends that from a perspective of consumer led-new product innovations the internet provides the ideal platform for collecting new product relevant information as it is ubiquitous and cost effective without geographical restrictions. Likewise, Sawhney et al, (2005) proclaim that larger strata of customers can contribute in NPD when compared to conventional marketing research practices.

### Research Design/Methodology

Unlike major researches which adopts either of the dichotomous approaches ie. Qualitative and Quantitative, this research is based on mixed approach. As a longitudinal study, the authors have tried to triangulate their observations, interviews with cluster analysis of a single case study at a Micro firm in Pakistan. The choice of case study make sense as we are trying to explore a new phenomena of text-mining in real life social media environment, where the concepts are boundary spanning. This approach is in accordance with Yin (2003) who favors case study research strategy when there is both qualitative and quantitative data in the research.

### Empirical Data Analysis

At the Micro firm Pakistan we have a planned strategy to find out insights from social media through text mining using the Generic text mining software. In an effort to confirm or reject the research questions the researchers have garnered the data from the three text mining studies carried out in the following Figure below (4.1)

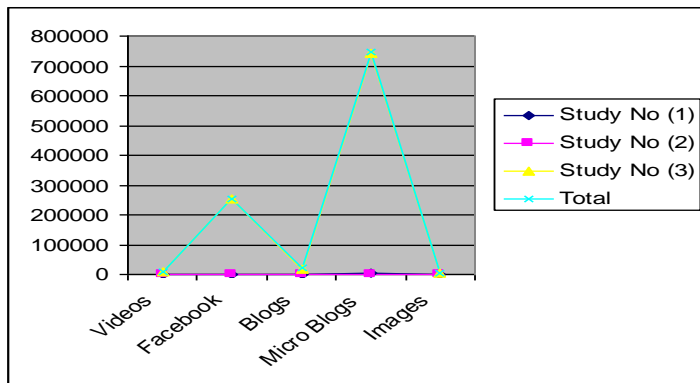
Social Media	Study No (1)	Study No (2)	Study No (3)	Total
Videos	374	114	9847	10335
Facebook	1827	393	252960	255180
Blogs	2056	118	19031	21205
Micro Blogs	4114	1140	741644	746898
Images	112	100	4435	4647
			<b>Total Posts</b>	<b>1038265</b>

The above figure shows that the maximum number of posts was generated Micro Blogs and Facebook and the least number were through images and videos respectively. When the data set were subjected to further analysis the following clustrous representation were generated as shown in Figure (4.2) below. The data gathered were interpreted through a sequence of coding, noting, abstracting, refining, and generalizing of the data.

### Consumer Insights and Product Development Process

The consumer insights attained through text mining which are also known as "*knowledge discovery in a text*" refers to

extracting useful information to be incorporated in the New Product Development at companies.



**Figure (4.2) Social Media Platform Posts of Various Search Terms**

The product development flow at the chosen Micro firm has consumer insights as its initial point. According to the company website all the products are a function of consumer driven product development process at the said firm. Inwardly, this means that consumer insights are incorporated at ideation stage of the new product development stage. Further the consumer insights collected from Social media can also be shared with research and development (R&D) so that the consumer perspective is available for a team driven by technological perspectives.

#### Conclusions

This paper proposes a simple method through which businesses can incorporate ongoing consumer discussion over the World Wide Web in their new product development stages.

This paper through its novel idea of using social media for capturing consumer insights to be made part of new product development process in an organization has further enhanced the business application of this novel approach at least on academic level. By using a mixed method the research renders a clear view of identifying suitable social media platforms and then collecting the desired consumer insights through text-mining

practices. Finally, this research has opened a new avenue for market research firms to collect consumer insights for forming and revising their social media campaigns.

#### Reference List

- W.Glynn Managold,D.j.F. (2009). "Social Media: The New Hybrid Element of The Promotion Mix" *Business Horizons* 53(4): 357-365.
- Thorsten Henning-Thurau, K.P.G Gianfranco walsh, Dwayne D. grenker (2004). "Electronic Word Of Mouth Via Consumer Opinion Platforms: What Motivates Consumers To Articulate Themselves On The Internet?" *Journal Of Interactive Marketing* 18(1):38-52
- Mohanbir Sawahney, G.V, Emanuela Prandelli (2005) "Collaborating To Create: The Internet As A Platform For Customer Engagement In Product Innovation" *Journal Of Interactive Marketing* 19(4):4-17
- Karl Ulrich, S.E (2003) *Product Design and Development*. Network, Mcgraw-Hill/Irwin
- Kaplan, M. h. A. M (2009). "Flagship Brands Stores Within Virtual Worlds: The Impact of Virtual Worlds: The Impact of Virtual Store Exposure on Real Life Attitude Towards The Brand and Purchase Intent". *Recherche et Applications en marketing* 24 93)
- Josh Bernoff, C. L. (2008-2011). *Groundswell, Expanded and revised Edition: Winning in The World Transformed by Social Technologies*. Harvard Business Review Press: Expanded and Revised Edition.
- Glen I. Urban, J. r. H. (1993) *design and Marketing of New Product*
- Geoffery Barber, H. L. (2011). *Data Mining in Social Media Chapter 12. Social Network Data Analytics*. Newyork, Springer-Science+ Business Media, LLC.
- Evans, D. (2008). *Social Media Marketing: An Hour a Day*, Sybex
- Urban, Hauser (1993). "Consumer Idealised Design: Evolving Consumers in The Product Development Process. "The Journal of Product Innovation Management 10 (4): 341-347