Indian Journal of Science and Technology, *Vol 9(43), DOI: 10.17485/ijst/2016/v9i43/104754, November 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Analysing the Role of User Generated Content on Consumer Purchase Intention in the New Era of Social Media and Big Data

Prabha Kiran* and S. Vasantha

School of Management Studies, Vels University, Pallavaram, Chennai - 600117, Tamil Nadu, India;
prabha.ram22@gmail.com, vasantha.sms@velsuniv.ac.in

Abstract

Objective: Big Data refers to the overwhelming amount of data that is being captured today by society, computers, cell phones and the internet. These data sets are so large and are of varied in nature, type and format that it becomes difficult to actually capture, manage, analyze, transform, model and organize this unstructured data for realizing company's goal of discovering information and gain insights into consumer purchasing behavior. The paper attempts to offer this understanding of insights into consumer's requirements through studying this social media big data. **Methods/Statistical Analysis:** The paper proposes that Social Media and Big Data are related to development of consumer purchase behavior. The unstructured data that is generated also known as User Generated Data (UGC) plays a very important role in forming consumer purchase intention. **Findings:** Through this study it was found that the new paradigm shift in the consumer's purchase intention is driven by Social Media and Big Data. The researcher has found a perfect model fit using Structural Equation Modeling and proven through hypothesis that Social Media and Big data combined together are responsible to generation of UGC's which impact purchase intention of consumers. **Application/Improvement:** the paper proposes that social media and big data are intersecting each other in a novel way and new methods and techniques need to be developed in order to get better insights into the unstructured data so that consumer requirements are better understood by marketers.

Keywords: Big Data, Consumer Behavior, Purchase Intension, Social Media

1. Introduction

The data world has exploded like a nuclear bomb reaction and is getting stronger day by day. For the matter of fact, the world would be generating 1ZB of data that's equal to 1024 exabytes by the year 2020 as suggested by a report by EMC. India alone will be contributing 2.9 zettabyte that's almost 7 percent of the global digital data universe. Simplifying it further, Cisco explains one Exabyte as equivalent to 36,000 years of HD-TV video or same as streaming entire Netflix Catalogue 3177 times.

Similarly, the big data market has been studied by Nasscom report in 2015, and it is estimated to be at $25 billion and growing at a pace of 45 percent Compound Annual Growth Rate (CAGR). Indian Big data market is not far behind and it is estimated at $1 billion and growing at a staggering pace of 83 percent CAGR. The social media big data is not far, as over 7000 photos are uploaded on Flickr per minute, over 600 videos are uploaded on YouTube per minute and Twitter alone generates 12 terabytes of data every day (Nasscom- Crisil report 2015). The International Data Corp. (IDC) study has found that globally there are 15 billion connected devices as per report by the year 2015. The Cisco report has highlighted that by 2017 the global internet traffic will reach 1.4 Zettabytes per year and 120.6 exabytes per month. The report has

**Author for correspondence*

stressed on the fact that by 2017 nearly half of all the traffic will originate from non-PC devices and mobile data traffic will reach 11.2 exobytes per month by 2017.

The EMC report also states that India is not utilizing the digital information available and only half a percent is actually used for analysis. The study has revealed that if the information is systematically analyzed 36 percent of it will eventually reveal valuable insights about the users. The report has also raised question marks on security aspect and has highlighted that almost 7.61 percent of India's digital needs protection of which only 56 percent is protected and the remaining is highly vulnerable to security threats. Gartner report of 2015 study has put forward that by year 2015, 85 percent of fortune 500 organizations will be able to exploit the big data for competitive advantage but it has also revealed that the global IT job requirement will be scarcely filled as out of the 4.4 million expected jobs will find only one third of the skilled professionals.

This revolution of humongous data has transformed each and every level of the business and has changed the way organizations operate. Mckinsey has rightly said about the Big Data as the fifth wave in the technological revolution. The personal data has not become an asset to the organizations and it at par with precious metals like gold, silver and oil. These data are considered as treasure and the organizations see opportunities waiting to be tapped. Social media data, website data, mobile data, recent data associated with cloud technologies and the data through connected devices are all changing the competitive landscape of the business and paving way for predictive analytics.

The Internet of Things implies to all the equipment connected to internet and interacts with the virtual and real world and Big Data have made the situation even more interesting. The wearable's and sensors offer a very high level of connectivity to the users.

2. Objectives

- To examine the relevance of big data in social media environment.
- To study the Social Media big data and its impact on Consumers purchase intention.

2.1 From Big Data to Social Big Data

The study report by McKinsey explains the big data phenomena as those data sets that are very large in size and way beyond the capacity of available database software tools that assist in the process of capturing and analyzing the data[1]. The big data can be generated from everywhere: pictures, videos, online purchases, the GPS locations, information through mobile devices so on and so forth. But the fact is that big data is more than just being called as the data being measured in huge volumes as terabytes or xetabytes[2]. Another very important facet of this big data is described by four V's: Volume, Velocity, Variety and Value[3]. Social media data are hence contributing in a huge proportion to fulfill the four V's and by the amount of user base and connectivity that has reached people one can only say that Big Data has now become the Social Big Data. The unstructured data like the texts, audio, video, click streams and the log in data of users are all laden with sentiments and connection among the users as well as with the brand, product, services and relationship with the organizations. The online reviews, blogs, comments and communities are all part of the big social media data. It is believed that more 80 percent of the available global digital data is unstructured and needs the marketers to understand and arrange it into a well-structured format for it to have meaningful interpretations that can be measured and analyzed. An important study carried out by IBM has stated that most of the companies are currently focused on capturing an analyzing the internal data sources that more structured to gain insights into consumer's intentions and behavior. Very few organizations look to capture the data outside their firewalls such as social media data just because they are unstructured and need more proficiency in handling, analyzing and interpreting. The study highlights the fact that external source contributes heavily to Big Data, 43 percent contribution is from social media, 38 percent from audio and 34 percent from videos and photos.

2.2Amalgamation of Big Data and Social media

The importance of social media cannot be relegated as it is one of the important waves of IT as described by McKinsey. Studies have revealed that 90 percent of all purchased are hugely impacted by the influence from one of the social group the consumer is subscribed to. Over 90 percent of consumers trust recommendations from people they know, more than 67 percent of the shoppers are willing to spend more for the recommended product from their friends and family members. Consumers want to share their experience more and hence there is a change in the consumer's behavior too, according to the study[4] percent of the Facebook users have "liked" a brand in past and 53 percent of twitter users have recommended companies or products to others based on their experience. The Loyal customers or the Fans of a particular brand are more likely to buy the same brand in their next purchase. Social media has the sharing feature which helps in increasing the awareness of a brand by 246 percent just by clicking a "like" button and by 98 percent if "send to a friend" is done. The information available to users is huge across each of their networks and the most important aspect is that companies have no control over the conversations users go through about a particular product or a brand. Relevant and engaging contents often allows users to interact more deeply about a product as these conversations provide value addition to the user, allows them to find answers to their questions and helps them gain better and novel solutions for their search issues. In fact, according to BrightEdge survey conducted in 2014 to study the Content and Search Marketing they have stated that 84 percent of marketers are assigning larger budget for developing a specified content strategy. The study reveals that it is very important engage consumers if the company wants to have competent and cohesive marketing operations. Figure 1 describes the amalgamation of big data and social media.

Social media serves as a guide for driving innovations in product design too. Studies have shown that 26 percent marketers utilize inputs from social media sources for R&D and product development and 46 percent use it to forecast future requirements also called predictive analysis.

2.3 The Big Social Data through Social Media Impacting Purchase Intentions

There has been an increase in the research of the ever dynamic social media and various networks in last few years as it has a very strong impact on the inter-relationships among the consumers and organizations. The interdisciplinary approach is the most important feature of social media[5]. In past when the social media was not present and internet based communication was fast becoming quite popular, measuring the affinity between the two interacting parties was very tough. Traditionally this was done using questionnaire method or interview method[6]. Recent studies have stated that observation method is best suited to analyze the strength of the bond between the two individuals in social media environment[7]. The intensions to purchase a product can be studied merely by observing the interaction in Facebook without actually interviewing. The new methods of analysis allow companies to define segmented correlations and measure the impact of the interactions on intension to buy. Studies have been had conducted traditional hypothesis testing on the data collected from 118 participants who are part of a social network and the study was conducted in a random field experiment that constituted 9167 users and close to 1.4 million friends on Facebook. The study was also expanded to analyze the effect of viral stories and innovative marketing campaigns on social media. It was found that these contents have a very strong impact on consumers and their decisions to purchase. The greater influence of peers in social media is highlighted and impact of user generated content is found to have direct influence on purchase decisions[8]. These new found changes need a completely novel approach towards research design so that all aspects of the interaction can be covered. Modern approaches like the graph theory models or neural networks and analyzing the current events are required to get an insight into consumer behavior[9].

Figure 1. Intersection of Big Data and Social Media.
Source: http//blogs.znet.com/Hinchcliffe

3. Research Methods

Qualitative analysis was conducted that resulted in finding the sources of User generated content that lead to development of consumer behavior. The paper attempts to study the impact of the UGC's on consumers purchase intention. The major source of UGC's as identified were Social Media and Website Big Data that is as a result of consumer interaction. The content analysis is stated to be a very useful method in identifying the major influencers[10]. In another study it has been stated that qualitative analysis can be further scrutinize to verify its validity as it has been congregated from public sources[11]. In order to increase the validity and improve the results survey research method has also been carried out[13]. Hence quantitative research was carried out and primary data was collected and statistical tests were conducted to validate the constructs.

3.1 Theoretical Framework

Figure 2 explains the Structure Equation Modeling diagram that encompasses the social media and big data for formation of purchase intention among consumers.

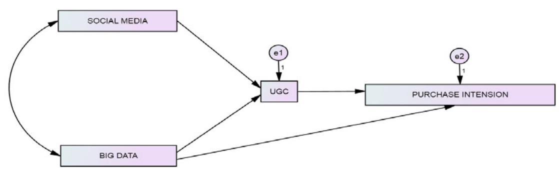

Figure 2. The proposed model.

3.1.1 Social Media

Social media incorporates a wide range of online avenues that enable consumers to interact, collaborate and share information related product and services[13]. Social media has enabled consumers to actively participate in various online communities and gather information and insight about the product and hence they are no longer dormant recipients of product related information. It helps in locating right information at right time and allows users to access same interest groups[14]. They have been promoted to content creators and distributers from merely being consumers in the form of videos, text messages, and audio. It is hence proven that consumers have the power to influence other users purchasing behavior and intentions on a very high level as which was not seen before[15].

3.1.2 Big Data

The advancements in technology have steered the way the information is accessed and due to this the importance of Big Data has also been increased[16]. In the forthcoming years the volume of data being generated and gathered is expected to explode[17]. Although the context of big data is enormous there is still a lot of non-uniformity and the data generated are mostly unstructured. Due this scenario there is a need for more advanced technology that can collect and analyse these unstructured data and infer meaningful inferences[18]. The communications done through electronic mode that speaks about the brand, product or services, delivered through the existing, past or future users is considered as electronic word of mouth[19]. The main challenge with big data analysis still remains about extracting meaningful inferences from large scale data which is available freely on internet[20]. Various tools and web services are being developed by scientist that will enable smooth interface between the user and website and will make data collection more easy and structured. Tools like Milne and Witten designed for Wikis, Reips and Garaizer designed especially for twitter[21]. Data generated are all various forms of User generated content that needs more structuring for in depth data analysis.

3.1.3 User Generated Content- UGC

The content that is generated though the interactions of members within a media and that involves the information to be created, influence, disseminated and consumed by others is referred to as User Generated Content[22]. The information generally includes product, brand or specification related subject matter that propels the consumer to gain more insights about the product and drive their purchase intentions[23]. Recent Studies states that electronic word of mouth and user generated content are although used in tandem but are very different. Only in case of brand specific products UGC and e-WOM can be used interchangeably[24].The study hence investigates the impact of these UGCs on consumer purchase intention.

The study is very important as UGC has a significant influence on consumers especially when marketplace is taken into consideration and retailers have limited control over these UGC and this is also a very big challenge for them[25].

3.1.4 Purchase Intention

Users of social media are exposed large amount of information either intentionally or unintentionally. In past studies have shown that the information available online influences consumers purchasing intentions[26], the level of influence alone may vary depending on consumer to consumer and the kind of product or services being searched[27].

4. Data Collection

In order to test the proposed framework, the data was collected through a structured questionnaire which was circulated to social media users and consumers who intend to search online before purchasing a product. Total of 250 questionnaires was distributed randomly for the period of two months (Feb 2016-March 2016) and we were able to collect 202 usable responses. The questionnaire consisted of demographic questions and Likert scale questions each containing five items were used (1= Strongly disagree, 2= Disagree, 3= Neutral, 4= Agree, 5= Strongly agree). Table 1 explains demographic of the respondents.

Table 1. Demographic summary

Gender	Frequency	Percent
FEMALE	94	46.5
MALE	108	53.5

Age	Frequency	Percent
26-33	52	25.7
34-41	98	48.5
ABOVE 41	52	25.7

Income	Frequency	Percent
BELOW 40000	56	27.7
40000-60000	66	32.7
60001-80000	36	17.8
ABOVE 80000	44	21.8
Education	Frequency	Percent
UG	12	5.9
PG	118	58.4
Professional Courses	72	35.6

5. Data Analysis and Results

Hypothesis:

Table 2 shows the relations between the constructs

H1: Social Media is positively associated with UGC- Accepted

H2: Big Data is positively associated with UGC- Accepted

H3: UGC is positively associated with Purchase Intention- Accepted

H4: Big Data is positively associated with Purchase Intention- Accepted

The items were tested so that they are uni-dimentional, as discussed by[28], they have stated that the items should be associated significantly with the indicators of the constructs and the association should be with only one construct. Bentler and Hair et al. have also studied the unidimentionality in their studies[29]. Model Fit indices were studied and based on these indices, ($\chi 2$/df=1.854; goodness of fit [GFI]=0.93; adjusted goodness of fit [AGFI]=0.91; Bentler comparative fit index[CFI]=0.972; root mean square residua [RMSR]l=0.06; and root mean square error of approximation [RMSEA]=0.047) the model fit was achieved, and it can be concluded that unidimentionality was achieved. Convergent and Discriminant validity was tested by using Confirmatory Factor Analysis (CFA). Table 2 shows the result of CFA. As discussed by Fornell and Larker in 1981 and Chen and Pulraj in 2004, standardized factor loadings of every indicator (>=0.5), composite reliability CR (>=0.7) and average variance extracted AVE (>=0.5)[30]. The results support the test. Discriminant validity was further analysed by calculating squared root of AVE measured and the result showed that that it was greater than the correlation coefficient between the constructs of same column. Common method bias was also studied as chances potential biases due to multiple responses are high. Table 3 explains the Factor Loading, CR and AVE values

Table 2. Hypothesis and results of goodness of fit indices

				Std RW	C.R.	P
H1	UGC	<---	Social media	0.39	5.305	***
H2	UGC	<---	Big Data	0.26	3.716	***
H3	Purchase intension	<---	UGC	0.88	16.571	***
H4	Purchase intension	<---	Big Data	0.41	5.381	***

Goodness-of-fit indices	
X2/d.f.	1.854
Goodness-of-fit index (GFI)	0.930
Adjusted GFI (AGFI)	0.906
Comparative fit index (CFI)	0.972
RMSEA	0.047

Note: *p < 0.05, **p < 0.01, ***p < 0.001.

Std R.W = Standardized Regression Weights, C.R = Critical Ratio.

Table 3. Factor loading, CR and AVE values

Variable	Item	Factor loading	CR*	AVE**
Social Media	SM1 SM2 SM3	0.74 0.86 0.82	0.851	1.95
Big Data	BD1 BD2 BD3 BD4	0.79 0.86 0.76 0.86	0.83	2.67
UGC	U1 U2 U3	0.85 0.80 0.88	0.837	2.13
Purchase Intention	PI1 PI2 PI3	0.73 0.74 0.84	0.815	0.77

*CR= Composite Reliability, **AVE= Average Variance Extracted

6. Discussion

The paper emphasizes that companies should explore the available knowledge and gain awareness regarding consumer's preferences. These can help companies in gaining insights into consumer's requirements and develop strategies accordingly. The traditional concept of seasonal demand may no longer exist and the in depth study might reveal more location based demands. Another aspect is regarding customized service and making the consumer a part of branding. By analyzing the digital footprints of consumers companies can deliver bundled services focused towards the consumer choice and gain an edge in the competitive market[31]. One of the best measures to analyze the consumer's likings the companies should observe the search carried out by them. Companies can create awareness in order to increase the initial interest in the product that eventually will lead to increasing search volumes. Big data has opened up prospects in many industries. Merely by following the consumers digital foot prints companies can strategies their marketing campaigns and maximize company's performance. Big social data allow the freedom to capture and observe the data and the reactions from the consumers towards the product. The purchase intentions of users can be impacted heavily by the interactions users have with each other in the social network communities. The focus of big social data analysis remains on the pattern used by the consumers for searching and interacting with other like-minded people. It can allow businesses to penetrate into consumer's minds and discover their need. The content creator can be identified and its influence on other readers can be ascertained.

7. Conclusion

Big data has not been completely understood by all and hence it still has a lot of potential to understand it and implement it to the benefit of organizations. In spite of few companies having understood the underlying potential, it still needs an expert to actually research the available to data to get insights out of it. Social media has been used by many companies to market their products and is in a very nascent stage or maturity. Many organizations are utilizing the existing platforms and channels for gathering consumer related information for developing a perfect strategy to implement perfect technology and get accurate results. Hence big social data is very useful and since social media has penetrated so deep in our lives we cannot ignore the impact it can have on the consumer decisions. Due to these changes in the social environment has led to deeper and wider studies on the impact these platforms have on consumer purchase intentions. The challenges still persist as identifying the real content creator and the one who can influence is tricky. Observation methods are best suited for conducting research in these contexts. This in-depth research will help in enhancing the decision making capabilities of the managers as the decisions will be more data driven than being just hypothetical. Purchase intensions are driven by social media communications and big data is also generated by social media platforms that imply there needs a proper system to understand these unstructured data derive inferences from them for analyzing the consumer preferences. Finally, the onus is on humans for delivering the best possible process that will help in discovering the best possible amalgamation of finding insights using machine data and sentiments of people.

8. Reference

1. Dellarocas XM, Zhang NF, Awad A. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. Journal of Interactive Marketing. 2007; 21(4):23–45.
2. Reiss PC, Wolak FA. Structural econometric modeling: rationales and examples from industrial organization. Chapter 64 Handbook of Econometrics; 2007. p. 6.

3. Predictive-analytics-for-the-small-business [Internet]. [cited 2016 Apr 12]. Available from: http://blog.monitor.us/2014/01/predictive-analytics-for-the-small-business-part-1/.
4. Garg R, Smith MD, Telang R. Measuring information diffusion in an online community. Journal of Management Information Systems. 2011; 28(2):11–38.
5. Russom P. Big data analytics, Best Practices Report, Fourth Quarter, The Data Warehouse Institute, Renton, WA; 2011 Sep.
6. Bhattacharjee S, Gopal RD, Lertwachara K, Marsden JR, Telang R. The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts. Management Science. 2007; 53(9):1359–74.
7. Kauffman RJ, Wood CA. Revolutionary research strategies for e-business: A philosophy of science view in the age of the Internet, economics. Chapter 2. In: Kauffman RJ, Tallon PP, editors. Economics, Information Systems, and Electronic Commerce: Empirical Research, M. E. Sharpe: Armonk, NY; 2009. p. 31–62.
8. Aral S, Walker D. Identifying social influence in networks using randomized experiments. IEEE Intelligent Systems. 2011; 26(5):91–6.
9. Davenport TH, Harris JG. Competing on analytics: The new science of winning, Harvard Business Press: Boston, MA; 2007.
10. Boyer KK, Swink M L. Empirical elephants—why multiple methods are essential to quality research in operations and supply chain management. Journal of Operations Management. 2008; 26(3):338–44.
11. Tangpong C. Content analytic approach to measuring constructs in operations and supply chain management. Journal of Operations Management. 2011; 29(6):627–38.
12. Jick T D. Mixing qualitative and quantitative methods: Triangulation in action. Administrative Science Quarterly. 1979; 24(4):602–11.
13. Tuten TL. Advertising 2.0: Social media marketing in a web 2.0 world. Westport, CT: Greenwood Publishing Group; 2008.
14. Mittal P, Garg S, Yadav S. Social network analysis using interest mining: A critical review. Indian Journal of Science and Technology. 2016; 9(16):1–8.
15. Accenture. Retail technology vision [Internet]. 2013. [cited 2015 May 27]. Available from: https://www.accenture.com/t20150527T211700__w__/cn-zh/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Local/zh-cn/PDF_2/Accenture-Retail-Technology-Vision-2013.PDF.
16. Anderson C. The end of the theory: the data deluge makes the scientific method obsolete [Internet]. 2008 [cited 2008 Jun 23]. Available from: http://www.edge.org/3rd_culture/anderson08/anderson08_index. html.
17. Kaisler S, Armour F, Espinosa JA, Money W. Big data: Issues and challenges moving forward, 46th Hawaii International Conference System Sciences (HICSS); 2013. p. 995–1004.
18. Eynon R. The rise of Big Data: What does it mean for education, technology, and media research? Learn. Media Technology. 2013; 38(3):237–40.
19. Liu X, Wang Y, Zhao D, Zhang W, Shi L. Patching by automatically tending to hub nodes based on social trust. Computer Standards & Interfaces. 2015; 44:94–101.
20. Cuzzocrea A, Song IY, Davis KC. Analytics over large-scale multidimensional data: The big data revolution! Proceedings of the ACM 14th International Workshop on Data Warehousing and OLAP; 2011. p. 101–4.
21. Milne D, Witten IH. An open-source toolkit for mining Wikipedia. Artificial Intelligence. 2013; 194:222–39.
22. Reips UD, Garaizar P. Mining twitter: A source for psychological wisdom of the crowds. Behavior Research Methods. 2011; 43(3):635–42.
23. Daugherty T, Eastin MS, Bright L. Exploring consumer motivations for creating user-generated content. Journal of Interactive Advertising. 2008; 8(2):16–25.
24. Smith AN, Fischer E, Yongjian C. How does brand-related user generated content differ across YouTube, Facebook, and Twitter? Journal of Interactive Marketing. 2015; 26(2):102–13.
25. Riegner C. Word of mouth on the web: The impact of web 2.0 on consumer purchase decisions. Journal of Advertising. 2007; 47(4):436–47.
26. Wang YJ, Minor MS, Wei J. Aesthetics and the online shopping environment: Understanding consumer responses. Journal of Retailing. 2011; 87(1):46–58.
27. Yang Z, Cai S, Zhou Z, Zhou N. Development and validation of an instrument to measure user perceived service quality of information presenting web portals. Information and Management. 2005; 42(4):575–89.
28. Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate data analysis. Upper Saddle River, NJ: Prentice Hall; 1998.
29. Bentler PM, Chou CP. Practical issues in structural modeling. Sociological Methods and Research. 1987; 16(1):78–117.
30. Fornell C, Larcker DF. Structural equation models with unobservable variables and measurement error: Algebra and statistics. Journal of Marketing Research. 1981; 18(3):382–8.
31. Brown JJ, Reingen PH. Social ties and word-of-mouth referral behavior. Journal Consumer Research. 1987; 14(3):350–62.