BIG DATA AND ONLINE SOCIAL NETWORKS:
TOOLS FOR BETTER USE OF INFORMATION ON BUSINESS

Edwin E. González-Carril, Universidad del Turabo (edwin.egonzalez@upr.edu)
Angel Ojeda-Castro, Universidad del Turabo (ut_aojeda@suagm.edu)

ABSTRACT

The business strategies of companies are being increasingly impacted by social media (Jobvite, 2014). The increase of hand-held smart devices for communication allows the users to create a great quantity of data in the social media that businesses can analyze (Rossouw, 2012). This paper presenting the importance of the use of social networks and the implementation of the techniques and tools of big data in the process management business services. With this article we present results and evidence of the impact of information on social networks and the role of big data for analysis and interpretation. Based on the comprehensive review of the literature, 48 articles from 2010 to 2015 were analyzed. The literature review leads us to conclude that the phenomenon of big data is still in its early beginnings and is still much to discover and research in this field (Martínez-Martínez, & Lara-Navarra, 2014). The application of big social data provides businesses improve the customer experience, achieve more targeted marketing, cost reduction and reinvent processes to make more precise and accurate decisions. These findings demonstrate that the use of big data tools and social networks can help businesses achieve higher quality and better management in their services.

Keywords: Big Data, Social Media, Big Social Data, and Social Networks.

INTRODUCTION

Internet Users

Nowadays, 70% of the world population of users of the internet have daily access and it is projected that this number will increase exponentially (Kemp, 2016). According to the Internet World Stats (2015), the use of the Internet has increased on the international level of 753.0% during the period from 2000 to 2015. Accordingly, in 2011, the data generated on the web augmented to approximately 1.8 Zettabyte that represents 1.8 trillion of GB (Bakshi, 2012; Suh, Barac, Bogdanovic & Radenkovic, 2015).

People expect all the processes and services to be connected to the Internet in real time (Kemp, 2016). The proliferation of technological equipment and the exponential growth of data that the users generate through the social media, demands that this topic become a target of study for investigators (Martínez-Martínez, & Lara-Navarra, 2014). The 93% of the participants use the Internet daily to access different technologies, such as: social network platforms, wikis, blogs, websites (Rainie, Smith & Duggan, 2013). Meanwhile, one fourth of the Internet users use social media (Zerofox, 2015). This signifies that the social media play a fundamental role in establishing communication between people and their companies.

The data produced in the social media is not defined and is composed of different formats, including text, audio, video, log files, among others (Suh et al., 2015). The analysis is creating a great challenge for businesses because, by not being well-defined data, the tools of Business Intelligence (BI) do not have the ability to effectively process them (Tascón, 2013).

The application of analyzing big social data is becoming a focus of study for researchers. Social media is increasingly impacting business strategies of companies (Jobvite, 2014). According to Luis Kreither, Executive Director of Kr Consulting, big data "has become a big challenge due to its ability to store and analyze a lot of data. And this, is also due to the emergence of social networks, whose information cannot be managed with the traditional sources" (Gerencia, 2014).

The objective of this paper is to present results and evidence of the impact of information flow in social networks and describe the role of big data for analysis and interpretation of the data generated by these means.

RESEARCH METHODOLOGY

The development of this research was to conduct an exhaustive search of the literature and analyze 48 professional references dating from the period 2010 to 2016. This work consists of the following parts: the first generation of data they produce will be described by the social networks; secondly, the concept of big data and its application in social networks will be defined; in the third, we will reflect on the benefits that companies can obtain using data to monitor their name brand. Finally, the results of this work will be presented.

Social Media

At the present, social media plays an important role in the daily and social activities of people (Del-Fresno-Garcia, 2014). Social networks have become an indispensable part of modern society for the development and dissemination of information (Wang, Huang, Rong, Wang, & Lai, 2014). People, who are social beings who need to communicate, can now accomplish this through the Internet, thanks to social media (Martínez-Martínez & Lara-Navarra, 2014). This new means of virtual communication increases the participation of users significantly each year at an accelerated rate (McGrath, 2011).

On July 2015, the data showed that 73% of Americans had at least one profile on a social network; representing an increase of 6% compared to 2014 (Statista.com, 2015). Kemp (2016) stated that there was an increase of 8.7% in the use of social networks worldwide. In its most recent study, he reported that there are 2.31 billion users in social media, of which 1.97 billion media access these via their mobile phones (Kemp, 2016). Also, Statisticbrain.com reported that by March 2015, Facebook surpassed 1,374 billion monthly connected users around the world. A more recent report showed that in January 2016, the most used social network in the world was Facebook, with 1.550 million active users monthly; followed by WhatsApp with 900 million and in third place QQ, with 860 million (Kemp, 2016; Statista.com, 2015). According to Duggan, Ellison, Lampe, Lenhart, & Madden (2015), 52% of adult users use several of these platforms.

Users share a large volume of data through social media. According to an infographic company Zerofox.com (2015), there are about 4.5 billion daily likes and five profiles are made every second on Facebook. While on Twitter, they are an average of 500 million daily tweets tweeted and 6,000 every second. Also, Google+, nearly 800,000 new users are created and more than 1.5 billion photos are uploaded. Social Times, presents an infographic that summarizes what millions of users do through social media every minute. The data shows that users upload about 250 thousand photos, making more than three million updates to their status, sending more than 150 thousand messages showing more than 15 million pages visited and shared with likes, among other details (Lafferty, 2014).

Moreover, the spread of technological equipment of communication encourages users to increasingly provide more data on social media that the companies can analyze and benefit from (Rossouw, 2012). In 2014, about 60% of the world population had a mobile phone; while by 2019 this figure is expected to increase to 67% (Statista.com, 2015). By January 2016, the use of mobile phone to access social networks, increased by 17% compared to 2015; equivalent to 283 million new users (Kemp, 2016). Similarly, it is expected that the number of smartphone users worldwide will reach 2.6 billion by the year 2019 (Kemp, 2016).

According to a study by Cisco, it is projected that between 2011 and 2016, the amount of mobile data traffic will grow at an annual rate of 78% (Cisco.com, 2012). While the number of mobile devices connected to the Internet will exceed the number of people on the planet (Cisco.com, 2012). In addition, it is estimated that traffic data through the cloud of electronic devices (mobile cloud), will represent 71% (7.6 Exabyte) per month of total mobile data traffic by 2016; compared with 45% (269 petabytes) per month reached in 2011 (Cisco.com, 2012).

Devices like smart phones, web cameras, radio frequency identification readers (RFID) and sensor networks are sources continuously generating data without necessarily using human intervention; causing a significant increase in

the volume of data and speed from which it is being generated (Wang & Alexander, 2015). These information sources collect data on the geographical location of a person, genetic information, blog entries, preferences for certain products and interactions that users perform through social media; providing scientists access to large amounts of information previously unavailable (Titiunik, 2015). The concept of big data promises to be the solution for this situation, since it refers to the processes by which it collects and analyzes large amounts of data, structured and unstructured, from various sources and taking into account volume, speed, variety and value (Martínez-Martínez & Lara-Navarra, 2014; Zaslavsky, Perera, Georgakopoulos, 2013).

Moreover, companies can obtain the following benefits through social media in the marketing process: increase exposure, increase traffic, build loyal fans, provide market information, potential customers target, improve search rankings, reduce marketing costs and increase sales (Stelzner, 2014). In 2014, the top five social media platforms most commonly used to carry out marketing were: Facebook (94%), Twitter (83%), LinkedIn (71%), YouTube (57%) and Bloggin (55%). In addition, 54% of sellers chose Facebook to advertise, followed by LinkedIn and Twitter (Stelzner, 2014).

Big Data

Big data has the potential to revolutionize the world market, because when analyzing large amounts of customer information, provides more accurate, precise and better quality business information, in order to offer more and better services to its market (Wang & Alexander, 2015). It also creates the skills and other challenges in the process of analysis to support a level of more accurate and timely decision-making (Monroy, 2013). According to Kim, Pelaez, & Winston (2013), big data stands out today as the new capability to drive business rates. In a survey by Gartner, more than 75% of companies invested or plan to invest in big data in the next two years (Van der Meulen & Woods, 2015). Some goals that companies may have to implement a project of big data are: improve customer service, streamlining existing systems, achieve more targeted marketing and cost reduction (Van der Meulen & Woods, 2015).

Lewis, Zamith, & Hermida (2013), describes big data as a high volume of information produced for and by human activity, due to the growing ubiquity of mobile devices, tracking tools, a variety of sensors, computers and low-cost storage. That is, big data refers to a large amount of data that cannot be processed and analyzed in traditional manner, due to its complexity (Liu, 2013; Wang & Alexander, 2015). In big data, most data are unstructured and contributes to a much wider variety of data types (Jeske, Grüner, & Weiß, 2013). In addition, the data is too large, they are generated quickly and do not conform to the restrictions of the designs of conventional databases (Dumbill, 2012).

According to several researchers, big data refers to large volumes of data that are characterized by their volume, velocity, variety, value, variability and validity also expressed as 6V (Russom, 2011; Dumbill, 2012; Zikopoulos & Eaton, 2011; Bellini, Di Claudio, Nesi, & Rauch, 2013; Demchenko, Grosso, De Laat, & Membrey, 2013; Megahed & Jones-Farmer, 2013; Rajpathak & Narsingpurkar, 2013). It is based on the analysis that can be performed using different data sources, which in the past were not of great importance to businesses, but now can be used to gain a new perspective and offer business advantages (Fernández, 2014). It also implies the relationship between sets of data, the complexity of these relationships, the speed at which data changes, and technological capacity to process in timely manner (Huang & Huang, 2015).

In recent years, big data has been of great interest to companies because of its ability to store and analyze a lot of data; which are due in large part to the emergence of social networks, whose information cannot be managed with traditional databases (Gerencia, 2014). Similarly, big data helps to improve the process in the supply chain and provides real-time estimates that help make critical decisions within companies (Wang & Alexander, 2015). The analysis of data collected with big data are of great importance for decision-making, in order to create market offerings (Wang & Alexander, 2015).

Big Social Data

Social media has been incorporated into the study of big data, thanks to its rapid growth and the variety of interactions and content that users generate (Martínez-Martínez & Lara-Navarra, 2014; West, 2013). Users of social media react to events, issues and institutions in order to express their personal opinions and emotions (Bravo-Marquez, Mendoza, & Poblete, 2014). Big social data is a phenomenon proposed by Manovich, which analyzes the data generated and the feelings of users through social media, using as a tool big data (Manovich, 2011). The analysis of feelings tries to get the real opinion of people about a specific product, services, organizations, events and values (Ghosal, Bhattacharjee, Das, & Ningombam, 2014). Access to all this information through big data allows better management in the company, since by getting the most accurate information, they are able to improve their products, services, and have greater operational efficiency in real-time (Wang & Alexander, 2015). When data is analyzed in real time, companies have the opportunity to follow the thoughts, opinions, ideas and feelings of people (Manovich, 2011); allowing them to reach customers more directly and meet their needs which might otherwise go unnoticed (Martínez-Martínez & Lara-Navarra, 2014).

Recent researches have shown that the average social big can be used to predict outcomes, such as income from films, winners of the elections, moods localized outbreaks (Asur & Huberman, 2010; Bollen & Mao, 2011; Lassen, Madsen & Vatrapu, 2015; Vatrapu, Mukkamala, & Hussain, 2014). Asur and Huberman (2010), revealed that the content of media can be used to predict real-world results. In their study, they analyzed nearly three million tweets on Twitter.com users to predict box office revenues of movies before its release in theaters. The authors demonstrated that their model predictions exceeded the Hollywood Stock Exchange and there was a strong correlation between the amounts of attention was on a given topic (in case of an upcoming film) and its ranking in the future.

Vatrapu et al. (2014) in his study analyzed data generated by Facebook and its relation to sales of the company H&M; using Data Analytic Tool Social tool. The authors found that there were statistically significant correlations between actual results of the companies in their quarterly sales activities of social media (messages, likes, comments), as well as positive, negative expressions or neutral sentiment of users toward the company. An extension of this study was performed by Lassen et al. (2015), which showed that social media Twitter and Facebook were used to predict quarterly iPhone sales and revenues of the company H&M; using social big data.

The high content of value generated by this type of information is enabling organizations improved their decisions, leading to obtaining competitive advantages in different fields (Tabares, & Hernández, 2014). Big social data analysis is of great value for companies and researchers because they can understand the thinking and actions of the person; also improve the processes that lead to decision-making and efficiently manage products and services (Hobbs, 2014).

SUMMARY

The great quantity of data that is being generated via the social media on the Internet requires tools and analytic solutions that allow deleting information that is not important and that can be useful. Big data promises to be a solution to this problem, now that the knowledge of the information that is generated by the social networks can be used to improve businesses. The data that is provided by the social media and the communication devices represent new and abundant sources of information that businesses can use to revise and create competitive advantages in their organization.

The high demand in the use of social media allows companies to obtain information on the location of users, schedules and specific uses of the different alternatives in services that access or share with their friends via portable devices such as mobile smartphones and tablets. The use of this information will be useful to the organization to establish a specific market and customized for each potential customer. Also, the social big data, provides ample opportunity for companies to collect information that can instantly be used as a tool for building competitive advantages in their organization.

However, social media has created major challenges for big data, such as processing large volumes of data streaming and data of human expression that are made by the publications produced by these media (Bravo-Marquez et al., 2014). Lisa Kart, research director at Gartner, indicated that the challenges companies face when implementing a project of big data are the interpreting data, human resources skills, organizational governance, financing and return on investment (Van der Meulen & Woods, 2015).

In this paper, we proposed the importance of implementing big social data tools and techniques for managing products and business services. We have shown that use big social data allows companies to obtain benefits for customers and suppliers. It also allows companies ample opportunity to gather information that can instantly use it to build competitive advantage.

However, the social big data is a recent phenomenon that requires longer be studied. Analyzing the feelings of users in social media, it is an example of the challenge that companies have to interpret this unstructured information. Big social data presents a particular complexity that comes through user's interactions made between them using different types of devices and also connecting to each other devices themselves (Martínez-Martínez, & Lara-Navarra, 2014). Also, the literature demonstrates with empirical studies the effectiveness of big social data for predicting outcome (Asur & Huberman, 2010; Bollen & Mao, 2011, Lassen et al., 2015; Vatrapu et al., 2014). Similarly, there are still questions about the tools or analytic solutions for the management and generation of knowledge (Martínez-Martínez, & Lara-Navarra, 2014). The challenge facing companies will continue to consist in understanding and extracting useful information from complex data sources and unstructured (Kim, et al. 2013).

REFERENCES

Asur, S., & Huberman, B. A. (2010, August). Predicting the future with social media. In Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/*ACM International Conference* on (Vol. 1, pp. 492-499). IEEE.

Bakshi, K. (2012, March). Considerations for big data: Architecture and approach. In *Aerospace Conference, 2012 IEEE* (pp. 1-7). IEEE.

Bellini, P., Di Claudio, Nesi, P., & Rauch, N. (2013). Tassonomy and review of Big data solutions navigation. *Big Data Computing,* Akerkar, R. (Ed.), CRC Press, pp: 57-101

Bollen, J., & Mao, H. (2011). Twitter mood as a stock market predictor. *Computer*, 44(10), 0091-94.

Bravo-Marquez, F., Mendoza, M., & Poblete, B. (2014). Meta-level sentiment models for big social data analysis. *Knowledge-Based Systems*, *69*, 86-99.

Cisco.com. (2012). *Índice de pronóstico de red visual de Cisco proyecta que el tráfico de datos de internet móvil global crecerá 18 veces entre los años 2011 a 2016*. Retrieved from http://globalnewsroom.cisco.com/es/la/press-releases/indice-de-pronostico-de-red-visual-de-cisco-proyec-nasdaq-csco-0851825

Del-Fresno-García, M. (2014). Haciendo visible lo invisible: Visualización de la estructura de las relaciones en red en Twitter por medio del Análisis de Redes Sociales. *El profesional de la información*, *23*(3), 1386-6710.

Demchenko, Y., Grosso, P., De Laat, C., & Membrey, P. (2013, May). Addressing big data issues in scientific data infrastructure. In *Collaboration Technologies and Systems (CTS), 2013 International Conference on* (pp. 48-55). IEEE.

Duggan, M., Ellison, N. B., Lampe, C., Lenhart, A., & Madden, M. (2015). *PewResearchCenter Internet, Science & Tech*. Retrieved from Social Media Update 2014: http://www.pewinternet.org/2015/01/09/social-media-update-2014/

Dumbill, E. (2012). *Planning for big data*. " O'Reilly Media, Inc

Fernández, A. (2014). Los datos de las personas a los pies de la empresa. Retrieved from *ABC Tecnología*:
http://www.abc.es/tecnologia/redes/20141102/abci-datos-personas-pies-empresas-201410312203.html

Gerencia (2014). Big Data: Impactando la toma de decisiones de las empresas. Retrieved from
http://www.microbyte.cl/gere/flipbook/201404/#/27

Ghosal, T., Bhattacharjee, S., Das, S. K., & Ningombam, D. (2014). A Comparative Study on Sentiment Analysis of
Big Data. *Communication, Cloud and Big Data: Proceedings of CCB 2014*.

Hobbs, Abbi (2014). "Social media and big data". POSTnote, n. 460, March. *Houses of Parliament. Parliamentary
Office of Science & Technology*. http://www.parliament.uk/briefing-papers/POST-PN-460/social-media-
and-bg-data

Huang, P., & Huang, P. (2015). When Big Data Gets Small. *International Journal of Organizational Innovation
(Online), 8*(2), 100-117. Retrieved from http://search.proquest.com/docview/1721368656?accountid=44848

Internet World Stats. (2015). *World Internet Users and 2015 Population Stats*. Retrieved from
http://www.internetworldstats.com/stats.htm

Jeske, M., Grüner, M., & Weiß, F. (2013). Big Data in Logistics: A DHL perspective on how to move beyond the
hype. *DHL Customer Solutions & Innovation*, *12*.

Jobvite. (2014). *Social Recruiting Survey Results 2014*. Retrieved from Jobvite.com: https://www.jobvite.com/wp-
content/uploads/2014/10/Jobvite_SocialRecruiting_Survey2014.pdf

Kemp, S. (2016). *Digital, Social & Mobile Worldwide in 2015*. *Wearesocial.net*. Retrieved from
http://wearesocial.com/uk/special-reports/digital-in-2016

Kim, H. J., Pelaez, A., & Winston, E. R. (2013). Experiencing Big Data Analytics: Analyzing Social Media Data in
Financial Sector as a Case Study. *Proceedings for the Northeast Region Decision Sciences Institute*
(NEDSI), 62-69.

Lafferty, J. (2014). *Infographic*: What happens in a Facebook minute? Retrieved from Adweek.com:
http://www.adweek.com/socialtimes/infographic-what-happens-in-a-facebook-minute-2/299587?red=if

Lassen, N. B., Madsen, R., & Vatrapu, R. (2015). Predictive Analytics with Big Social Data. In *International
Conference on Computational Social Science*.

Lewis, S. C., Zamith, R., & Hermida, A. (2013). Content Analysis in an Era of Big Data: A Hybrid Approach to
Computational and Manual Methods. *Journal of Broadcasting & Electronic Media*, *57*(1), 34-52.
doi:10.1080/08838151.2012.761702

Liu, L. (2013). Computing infrastructure for big data processing. *Frontiers of Computer Science*, *7*(2), 165-170.

Manovich, L. (2011). Trending: The promises and the challenges of big social data. *Debates in the digital
humanities*, *2*, 460-475.

Martínez-Martínez, S., & Lara-Navarra, P. (2014). El big data transforma la interpretación de los medios sociales. *El
Profesional de la Información*, *23*(6), 575-581.

McGrath, L. C. (2011). Social networking privacy: Important or not? *Interdisciplinary Journal of Contemporary
Research in Business*, *3*(3), 22-28

Megahed, F. M., & Jones-Farmer, L. A. (2013). A statistical process monitoring perspective on big data. *Frontiers
in Statistical Quality Control, 11th ed. Springer, New York*.

Monroy, J. M. V. (2013). Big Data: las oportunidades profesionales que vienen. *Bit,* (193), 53-55.

Rainie, L., Smith, A., & Duggan, M. (2013). Coming and going on Facebook. *Pew Research Center's Internet and American Life Project*.

Rajpathak, T., & Narsingpurkar, A. (2013). Managing knowledge from Big Data analytics in product development. *White Paper, Tata Consultancy Services*

Rossouw, L. (2012). Big Data–Grandes Oportunidades. *Gen Re–Risk Insights*, *16*(2).

Russom, P. (2011). Big data analytics. *TDWI Best Practices Report, Fourth Quarter*, 1-35

Statista.com (2015). *Statistics and facts about Social Networks*. Retrieved from http://www.statista.com/topics/1164/social-networks/

Statisticbrain.com. (2015). *Social Networking Statistics | Statistic Brain*. Retrieved from http://www.statisticbrain.com/social-networking-statistics

Stelzner, M. (2014). 2014 *Social Media Marketing Industry Report*. Retrieved from Social Media Examiner: http://www.socialmediaexaminer.com/SocialMediaMarketingIndustryReport2014.pdf

Suh, J., Vujin, V., Barac, D., Bogdanovic, Z., & Radenkovic, B. (2015). Designing Cloud Infrastructure for Big Data in E-Government. *RUO. Revija za Univerzalno Odlicnost*, *4*(1), A26.

Tabares, L. F., & Hernández, J. F. (2014). Big Data Analytics: Oportunidades, Retos y Tendencias. *Cali, Colombia*.

Tascón, M. (2013). Introducción: Big Data. Pasado, presente y futuro. Telos: *Cuadernos de comunicación e innovación*, (95), 47-50.

Titiunik, R. (2015). Can big data solve the fundamental problem of causal inference? *PS: Political Science & Politics*, *48*(01), 75-79

Van der Meulen, R. & Woods, V. (2015). *Gartner Survey Shows More Than 75 Percent of Companies Are Investing or Planning to Invest in Big Data in the Next Two Years*. Retrieved from http://www.gartner.com/newsroom/id/3130817

Vatrapu, R., Mukkamala, R. R., & Hussain, A. (2014). Towards a Set Theoretical Approach to Big Social Data Analytics: Concepts, Methods, Tools, and Empirical Findings. In *2014 International Conference on Social Media and Society Conference*.

Wang, L. Z., Huang, Z. G., Rong, Z. H., Wang, X. F., & Lai, Y. C. (2014). Emergence, evolution and scaling of online social networks. *PloS one*, *9*(11), e111013.

Wang, L., & Alexander, C. A. (2015). Big data driven supply chain management and business administration. *American Journal of Economics and Business Administration*, *7*(2), 60.

West, O. (2013). "Big data insights". *Webit Congress 2013*. ComScore. Retrieved from https://www.comscore.com/esl/Insights/Presentationsand-Whitepapers/2013/Big-Data-Insights

Zaslavsky, A., Perera, C., & Georgakopoulos, D. (2013). Sensing as a service and big data. *arXiv preprint arXiv:1301.0159*

Zikopoulos, P., & Eaton, C. (2011). *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media

Zerofox.com. (2015, March 6). Big Data Just Got Bigger [*Infographic*]. Retrieved from Zerofox.com: https://www.zerofox.com/whatthefoxsays/social-media-big-data-just-got-bigger/