Big Data Usage in the Marketing Information System

Alexandre Borba Salvador, Ana Akemi Ikeda
Faculdade de Administração, Economia e Ciências Contábeis, Universidade de São Paulo, São Paulo, Brazil
Email: alexandre.salvador@usp.br

Received 18 June 2014; revised 20 July 2014; accepted 11 August 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY).
http://creativecommons.org/licenses/by/4.0/

Abstract

Data generation, storage capacity, processing power and analytical capacity increase had created a technological phenomenon named big data that could create big impact in research and development. In the marketing field, the use of big data in research can represent a deep dive in consumer understanding. This essay discusses the big data uses in the marketing information system and its contribution for decision-making. It presents a revision of main concepts, the new possibilities of use and a reflection about its limitations.

Keywords

Big Data, Marketing Research, Marketing Information System

1. Introduction

A solid information system is essential to obtain relevant data for the decision-making process in marketing. The more correct and relevant the information is, the greater the probability of success is. The 1990s was known as the decade of the network society and the transactional data analysis [1]. However, in addition to this critical data, there is a great volume of less structured information that can be analyzed in order to find useful information [2]. The growth of generation, storage capacity, processing power and data analysis provided a technological phenomenon called big data. This phenomenon would cause great impacts on studies and lead to the development of solutions in different areas. In marketing, big data research can represent the possibility of a deep understanding of the consumer behavior, through their profile monitoring (geo-demographic, attitudinal, behavioral), the statement of their areas of interest and preferences, and monitoring of their purchase behavior [3] [4]. The triangulation of the available data in real time with information previously stored and analyzed would enable the generation of insights that would not be possible through other techniques [5].

However, in order to have big data information correctly used by companies, some measures are necessary,
such as investment on people qualification and equipment. More than that, the increase of information access may generate ethic-related problems, such as invasion of privacy and redlining. It may affect research as well, as in cases where information could be used without consent of the surveyed.

Predictive analytics are models that seek to predict the consumer behavior through data generated by their purchase and/or consumption activities and with the advent of big data, predictive analytics grow in importance to understand this behavior from the data generated in on-line interactions among these people. The use of predictive systems can also be controversial as exemplified by the case of American chain Target, which identified the purchase behavior of women at the early stage of pregnancy and sent a congratulation letter to a teenage girl who had not yet informed her parents about the pregnancy. The case generated considerable negative repercussions and the chain suspended the action [4].

The objective of this essay is to discuss the use of big data in the context of marketing information systems, present new possibilities resulting from its use, and reflect on its limitations. For that, the point of view of researchers and experts will be explored based on academic publications, which will be analyzed and confronted so we may, therefore, infer conclusions on the subject.

2. The Use of Information on the Decision-Making Process in Marketing

The marketing information system (MIS) was defined by Cox and Good (1967, p. 145) [6] as a series of procedures and methods for the regular, planned collection, analysis and presentation of information for use in making marketing decisions. For Berenson (1969, p. 16) [7], the MIS would be an interactive structure of people, equipment, methods and controls, designed to create a flow of information able to provide an acceptable base for the decision-making process in marketing. The need for its implementation would derive from points that have not changed yet: 1) the increase in business complexity would demand more information and better performance; 2) the life cycle of products would be shortened, requiring more assertiveness from marketing managers to collect profits in shorter times; 3) companies would become so large that the lack of effort to create a structured information system would make its management impractical; 4) business would demand rapid decisions and therefore, in order to support decision making, an information system would be essential for marketing areas; 5) although an MIS is not dependent on computers, the advances in hardware and software technologies would have spread its use in companies, and not using its best resources would represent a competitive penalty [7].

The data supplying an MIS can be structured or non-structured regarding its search mechanisms and internal (company) or external (micro and macro environment) regarding its origin. The classic and most popular way of organizing it is through sub-systems [8]-[10]. The input and processing sub-systems of an MIS are the internal registration sub-system (structured and internal information), marketing intelligence sub-system (information from secondary sources, non-structured and from external origins), and the marketing research sub-system (information from primary sources, structured, from internal or external origins, generated from a research question).

3. Big Data

The term big data applies to information that could not be processed using traditional tools or processes. According to an IBM [11] report, the three characteristics that would define big data are volume, speed and variety, as together they would have created the need for new skills and knowledge in order to improve the ability to handle the information (Figure 1).

The Internet and the use of social media have transferred the power of creating content to users, greatly increasing the generation of information on the Internet. However, this represents a small part of the generated information. Automated sensors, such as RFID (radio-frequency identification), multiplied the volume of collected data, and the volume of stored data in the world is expected to jump from 800,000 petabytes (PB) in 2000 to 35 zettabytes (ZB) in 2020. According to IBM, Twitter would generate by itself over 7 terabytes (TB) of data a day, while some companies would generate terabytes of data in an hour, due to its sensors and controls. With the growth of sensors and technologies that encourage social collaboration through portable devices, such as smartphones, the data became more complex, due to its volume and different origins and formats, such as files originating from automatic control, pictures, books, reviews in communities, purchase data, electronic messages and browsing data. The traditional idea of data speed would consider its retrieval, however, due to the great number of sensors capturing information in real time, the concern with the capture and information analysis speed emerges,
leading, therefore, to the concept of flow. The capture in batches is replaced by the streaming capture. Big data, therefore, regards to a massive volume of zettabytes information rather than terabytes, captured from different sources, in several formats, and in real time [11].

A work plan with big data should take three main elements into consideration: 1) collection and integration of a great volume of new data for fresh insights; 2) selection of advanced analytical models in order to automate operations and predict results of business decisions; and 3) creation of tools to translate model outputs into tangible actions and train key employees to use these tools. Internally, the benefits of this work plan would be a greater efficiency of the corporation since it would be driven by more relevant, accurate, timely information, more transparency of the operation running, better prediction and greater speed in simulations and tests [12].

Another change presented by big data is in the ownership of information. The great information storages were owned only by governmental organizations and major traditional corporations. Nowadays, new corporations connected to technology (such as Facebook, Google, LinkedIn) hold a great part of the information on people, and the volume is rapidly increasing. Altogether, this information creates a digital trail for each person and its study can lead to the identification of their profile, preferences and even prediction of their behavior [5].

Within business administration, new uses for the information are identified every day, with promises of benefits for operations (productivity gains), finance (control and scenario predictions), human resources (recruitment and selection, salary, identification of retention factors) and research and development (virtual prototyping and simulations). In marketing, the information on big data can help to both improve information quality for strategic planning in marketing and predict the definition of action programs.

4. Use of Big Data in the Marketing Information System

Marketing can benefit from the use of big data information and many companies and institutes are already being structured to offer digital research and monitoring services. The use of this information will be presented following the classical model of marketing information system proposed by Kotler and Keller (2012) [10].

4.1. Input-Sub-Systems

4.1.1. Internal Reports

Internal reports became more complete and complex, involving information and metrics generated by the company’s digital proprieties (including websites and fanpages), which would also increase the amount of information on consumers, reaching beyond the data on customer profile. With the increase of information from different origins and in different formats, a richer internal database becomes the research source for business, markets, clients and consumers insights, in addition to internal analysis.

4.1.2. Marketing Intelligence

If in one hand the volume of information originated from marketing intelligence increases, on the other hand, it
is concentrated on an area with more structured search and monitoring tools, with easier storage and integration. Reading newspapers, magazines and sector reports gains a new dimension with the access to global information in real time, changing the challenge of accessing information to selection of valuable information, increasing, therefore, the value of digital clipping services. The monitoring of competitors gains a new dimension since brand changes, whether local or global, can be easily followed up. The services of brand monitoring increase, with products such as GNPD by Mintel [13] and the Buzzz Monitor by e. Life [14] or SCUP and Bluefin.

4.1.3. Marketing Research
Since the Internet growth and virtual communities increase, studying online behavior became, at the same time, an opportunity and a necessity. Netnography makes use of ethnography sources when proposing to study group behavior through observation of their behavior in their natural environment. In this regard, ethnography (and netnography) has the characteristic of minimizing the behavior changes setbacks by not moving the object of study from its habitat, as many other study groups do. However, academic publications have not reached an agreement on technique application and analysis depth [15]-[17]. Kozinets (2002, 2006) [16] [17] proposes a deep study, in which the researcher needs to acquire great knowledge over the object group and monitor it for long periods, while Gerbera (2008) [15] is not clear about such need of deep knowledge of the technique, enabling the understanding of that which could be similar to a content analysis based on digital data. For the former, just as ethnography, the ethical issues become more important as the researcher should ask for permission to monitor the group and make their presence known; and, for the latter, netnography would not require such observer presentation from public data collected. The great volume of data captured by social networks could be analyzed using netnography.

One of the research techniques that have been gaining ground in the digital environment is the content analysis due to, on one hand, the great amount of data available for analysis on several subjects, and, on the other hand, the spread of free automated analysis tools, such as Many Eyes by IBM [18], which offers cloud resources on terms, term correlation, scores and charts, among others. The massive volume of information of big data provides a great increase in the sample, and, in some cases, enables the population research, with “n = all” [4].

4.2. Storage, Retrieval and Analysis
With the massive increase of the information volume and complexity, the storage, retrieval and analysis activities are even more important with big data. Companies that are not prepared to deal with the challenge find support in outsourcing the process [11]. According to Soat (2013) [19], the attribution of scores for information digitally available (e-scores) would be one of the ways of working with information from different origins, including personal data (data collected from fidelity programs or e-mail messages), browsing data collected through cookies, and outsourced data, collected from financing institutes, censuses, credit cards. The information analysis would enable the company to develop the client’s profile and present predictive analyses that would guide marketing decisions, such as identification of clients with greater lifetime value.

4.3. Information for the Decision-Making Process in Marketing
The marketing information system provides information for strategic (structure, segmentation and positioning) and operational (marketing mix) decision making. The use of big data in marketing will analyzed below under those perspectives.

4.3.1. Segmentation and Positioning
For Cravens and Piercy (2008) [20], a segmentation strategy includes market analysis, identification of the market to be segmented, evaluation on how to segment it, definition of strategies of micro segmentation. A market analysis can identify segments that are unacknowledged or underserved by the competitors. To be successful, a segmentation strategy needs to seek identifiable and measurable, substantial, accessible, responsive and viable groups.

Positioning can be understood as the key characteristic, benefit or image that a brand represents for the collective mind of the general public [21]. It is the action of projecting the company’s offer or image so that it occupies a distinctive place in the mind of the target public [10]. Cravens and Piercy (2008, p. 100) [20] connect the segmentation activity to the positioning through identification of valuable opportunities within the segment. Segmenting means identifying the segment that is strategically important to the company, whereas positioning
means occupying the desired place within the segment.

Digital research and monitoring tools enable studies on the consumer behavior to be used in behavioral segmentation. The assignment of scores and the use of advanced analyses help to identify and correlate variables, define predictive algorithmics to be used in market dimensioning and lifetime value calculations [19] [22]. The netnographical studies are also important sources to understand the consumer behavior and their beliefs and attitudes, providing relevant information to generate insights and define brand and product positioning.

4.3.2. Product
From the positioning, the available information should be used to define the product attributes, considering the value created for the consumer. Information on consumer preferences and manifestations in communities and forums are inputs for the development and adjustment of products, as well as for the definition of complementary services. The consumer could also participate in the product development process by offering ideas and evaluations in real time.

The development of innovation could also benefit from big data, both by surveying insights with the consumers and by using the information to develop the product, or even to improve the innovation process through the use of information, benefitting from the history of successful products, analyses of the process stages or queries to an idea archive [23]. As an improvement to the innovation process, the studies through big data would enable the replication of Cooper’s studies in order to define a more efficient innovation process, by exploring the boundary between the marketing research and the research in marketing [24].

4.3.3. Distribution
Internal reports became more complete and complex, involving information and metrics generated by the company’s digital proprieties (including websites and fanpages), which would also increase the amount of information on consumers, reaching beyond the data on customer profile. With the increase of information from different origins and in different formats, a richer internal database becomes the research source for business, markets, clients and consumers insights, in addition to internal analysis.

In addition to the browsing location in the digital environment and the monitoring of visitor indicators, exit rate, bounce rate and time per page, the geolocation tools enable the monitoring of the consumers’ physical location and how they commute. More than that, the market and consumer information from big data enables to assess, in a more holistic manner, the variables that affect the decisions on distribution and location [25].

4.3.4. Communication
Big data analysis enables the emergence of new forms of communication research through the observation on how the audience interacts with the social networks. From their behavior analysis, new insights on their preferences and idols [3] may emerge to define the concepts and adjust details on the campaign execution. Moreover, the online interaction while displaying offline actions of brands enables the creation and follow up of indicators to monitor the communication [3] [26], whether quantitative or qualitative.

The increase of information storage, processing and availability enables the application of the CRM concept to B2C clients, involving the activities of gathering, processing and analyzing information on clients, providing insights on how and why clients shop, optimizing the company processes, facilitating the client-company interaction, and offering access to the client’s information to any company.

4.3.5. Price
Even offline businesses will be strongly affected by the use of online prices information. A research by Google Shopper Marketing Council [27], published in April, 2013, shows that 84% of American consumers consult their smartphones while shopping in physical stores and 54% use them to compare prices. According to Vitorino (2013) [4], the price information available in real time, together with the understanding of the consumers’ opinion and factors of influence (stated opinions, comments on experiences, browsing history, family composition, period since last purchase, purchase behavior), combined with the use of predictive algorithmics would change the dynamics, and could, in the limit, provide inputs for a customized decision making on price every time.

5. Limitations
Due to the lack of a culture that cultivates the proper use of information and to a history of high costs for storage
A. B. Salvador, A. A. Ikeda

space, a lot of historical information was lost or simply not collected at all. A McKinsey study with retail companies observed that the chains were not using all the potential of the predictive systems due to the lack of: 1) historical information; 2) information integration; and 3) minimum standardization between the internal and external information of the chain [28]-[30]. The greater the historical information, the greater the accuracy of the algorithm, provided that the environment in which the system is implemented remains stable. Biesdorf, Court and Willmott (2013) [12] highlight the challenge of integrating information from different functional systems, legacy systems and information generated out of the company, including information from the macro environment and social networks.

Not having qualified people to guide studies and handle systems and interfaces is also a limiting factor for research [23], at least in a short term. According to Gobble (2013) [23] McKinsey report identifies the need for 190,000 qualified people to work in data analysis-related posts today. The qualification of the front line should follow the development of user-friendly interfaces [12]. In addition to the people directly connected to the analytics, Don Schults (2012) [31] still highlights the need for people with “real life” experience, able to interpret the information generated by the algorithms. “If the basic understanding of the customer isn’t there, built into the analytical models, it’s really doesn’t matter how many iterations the data went through or how quickly. The output is worthless (SCHULTZ, 2012, p. 9).”

The management of clients in a different manner through CRM already faces a series of criticism and limitations. Regarding the application of CRM for service marketing, its limitations would lie in the fact that a reference based only on the history may not reflect the client’s real potential; the unequal treatment of clients could generate conflicts and dissatisfaction of clients not listed as priorities; and ethical issues involving privacy (improper information sharing) and differential treatment (such as redlining). These issues can be also applied in a bigger dimension in discussions about the use of information from big data in marketing research and its application on clients and consumers.

The predictive models are based on the fact that the environment where the analyzing system is implemented remains stable, which, by itself, is a limitation to the use of information. In addition to it and to the need of investing in a structure or expending on outsourcing, the main limitations in the use of big data are connected to three main factors: data shortage and inconsistence, qualified people, and proper use of the information. The full automation of the decision due to predictive models [5] also represents a risk, since that no matter how good a model is, it is still a binary way of understanding a limited theoretical situation. At least for now, the analytical models would be responsible for performing the analyses and recommendations, but the decisions would still be the responsibility of humans.

Nuan and Domenico (2013) [5] have also emphasized that people’s behavior and their relationships in social networks may not accurately reflect their behavior offline, and the first important thing to do would be to increase the understanding level of the relation between online and offline social behavior. However, if on one hand people control the content of the intentionally released information in social networks, on the other hand, a great amount of information is collected invisibly, compounding their digital trail. The use of information without the awareness and permission of the studied person involves the ethics in research [15]-[17]. Figure 2 shows a suggestion of continuum between the information that the clients would make available wittingly and the information make available unwittingly to the predictive systems. The consideration of the ethics issues raised by Kozinets (2006) [17], Nunan and Domenico (2013) [15], and reinforces the importance of increasing the clients’ level of awareness regarding the use of their information or ensuring the non-customization of the analysis of information obtained unwittingly by the companies.

6. Final Considerations

This study discussed the use of big data in the context of marketing information system, and, what was clear is that we are still in the beginning of a journey of understanding its possibilities and use, and we can observe the great attention generated by the subject and the increasing ethical concern. As proposed by Nunan and Domenico (2013) [5], the self-governance via ESOMAR (European Society for Opinion and Market Research) [32] is an alternative to fight the abuses and excesses and enable the good use of the information. Nunan and Di Domenico (2013) [5] propose to include in the current ESOMAR [32] rules the right to be forgotten (possibility to request deletion of history), the right to have the data expired (complementing the right to be forgotten, the
Figure 2. Continuum between the awareness and non-awareness regarding the use of information. Source: authors.

transaction data could also expire), and the ownership of a social graph (an individual should be aware of the information collected about them). In marketing communication, the self-governance in Brazil has showed positive results, such as the examples in the liquor industry and kid’s food industry, which, upon the pressure of public opinion, have adopted restrictive measures to repress abuses and maintain the communication of categories [33]. Industries such as the cigarette are opposite examples of how the excess has led to great restrictions to the categories. As in the prisoners’ dilemma [34], the self-governance forces a solution in which all participants have to give up on their best short-term individual options for the good of the group in a long term (Figure 3).

On the other hand, if the consumer’s consent in releasing the use of their information would solve the ethical issues, the companies would never have so much power to create value for their clients and consumers. Recovering the marketing application proposed in “Broadening the concept of marketing” [35], the exchange of consent release could be performed by offering a major non-pecuniary value. This value offer could be the good use of the information to generate services or new proposals that increase the value perceived by the client [10]. Currently, many mobile applications offer services to consumers, apparently free of charge, in exchange for their audience for advertisements and access to their information in social networks. By understanding which service, consistent with its business proposal, the consumer sees the value in, and making this exchange clear, the service and consent of the information use could be a solution to access information in an ethical manner.

From the point of view of marketing research, the development of recovery systems and the analyses of great volumes of non-structured information could lead to the understanding of consumer behaviors. Issues regarding the findings and understanding of the consumers in marketing research are addressed qualitatively. However, due to the volume of cases, could the studies, through big data, provide at the same time the understanding on the consumer and the measurement of the groups with this behavior? A suggestion for the following research
would be to study the combination of the qualitative and quantitative research objectives with the use of big data and analytical systems in understanding consumer behavior and measurement of group importance.

References


http://dx.doi.org/10.2307/1248740
Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.