

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/303363742>

A SWOT analysis of big data

Article in *The Journal of Education for Business* · May 2016

DOI: 10.1080/08832323.2016.1181045

CITATIONS

16

READS

10,766

3 authors, including:



Mohammad Ahmadi

University of Tennessee at Chattanooga

25 PUBLICATIONS 380 CITATIONS

[SEE PROFILE](#)



Parthasarathi Dileepan

University of Tennessee at Chattanooga

33 PUBLICATIONS 513 CITATIONS

[SEE PROFILE](#)

A SWOT ANALYSIS OF BIG DATA

Abstract

This is the decade of “Data Analytic” and “Big Data”; but not everyone agrees with the definition of Big Data. Some researchers see it as the future of data analysis, while others consider it as hype and foresee the demise of it in the near future. No matter how it is defined, Big Data for the time being has its glory moments. The most important factor about Big Data is the analysis of mammoth amounts of data. This paper addresses “Big Data,” will provide an explanation of it, will look at its strengths, weaknesses, opportunities and threats, and will provide a “SWOT” analysis of these fast growing topics.

Key Words for Indexing: Big Data, Data Analytic, SWOT Analysis,

Introduction

This is the decade of “Data Analytic” and “Big Data.” In recent years, there has been a tremendous emphasis on “Big Data” and “Business Analytics.” Not everyone agrees with what Big Data is, nor there is an agreement on its definition. Some researchers define Big Data as (1) an immense amount of data in real time and (2) degrees of structure of data (Sigma 2014). Gartner avers that “Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization” (<http://www.gartner.com/it-glossary/big-data>). Others see the concept of Big Data as hype and foresee the demise of it in the near future. No matter how it is defined, Big Data for the time being has its glory moments. The most important factor about Big Data is the analysis of mammoth amounts of data. This paper addresses “Big Data” and will provide a “SWOT” analysis of these fast growing topics. SWOT is an acronym that stands for “strengths, weaknesses, opportunities, and threats.” The research aims to see if there is a consensus in the definition of “Big Data” among scholars. Massive amounts of data from a wide variety of sources are being collected every second by businesses and organizations and analyzed in near real-time. Organizations are using these data and techniques to help them improve their management decisions and strategies to gain a competitive advantage over their competitors.

Big Data Defined

Big Data is a term used to describe the volume, velocity, and variety of data which has been coined as the three Vs. First V represents the volume of data created each day has become so immense that it is measured in exabytes (an exabyte is one billion gigabytes). To visualize what this amount of data represents, McAfee & Brynjolfsson provide an example regarding Walmart’s customer transactions. Walmart collects more than 2.5 petabytes of data every hour,

which is equivalent to about 20 million filing cabinets worth of text (McAfee & Brynjolfsson, 2012). The second V represents the velocity, which indicates how quickly data can be accumulated. In the business world where competitive advantage plays an important role, speed by which data can be gathered will give a company a major competitive advantage. The Big Data industry is trying to create tools to process information quick enough to be nearly real-time (McAfee & Brynjolfsson, 2012). Finally, the last V represents variety, which is incredibly large. The variety of data from social networks to daily telephone conversation makes Big Data and data analysis a major task. With a traditional database, all the data is structured. The big change that Big Data will bring is that we will be able to use, interpret, and store any type of data whether it is structured or not. This will help companies tap into a lot of really powerful data contained within social media sites, blogs, and any other type of unstructured data. In short, this will allow companies access to every data source they can think of, no matter where it comes from, in order to see a more detailed picture of their business, customer base, employee base, supplier base, and brand image (Cvijanovic, 2012). Steve Lohr from the New York Times defined Big Data as, “shorthand for advancing trends in technology that open the door to a new approach to understanding the world and making decisions”. (Lohr, 2012) Simply put, it is any form of data that a company can collect and analyze to solve a problem.

Literature Review

Big Data provides colossal amount of information for companies. Leaders, managers, and executives want to strengthen their business and try to gain an advantage over the competition. According to a study done by the McKinsey Global Institute, the United States needs at least 140,000 more people that have a “deep analytical” skill set and over a million managers that can understand the data. (Lohr, 2012) These employees are not needed just for

Fortune 500 companies. Data analytics can be used in everything from the financial markets to healthcare to even sports teams. Anyone and everyone can use collected data as long as they have someone with the expertise to analyze and make sense of the data so that it can be used effectively.

Economic benefits to Big Data have been identified in three main areas: business efficiency, business innovation, and business creation. ("Data equity", 2012) Efficiency gains can be realized through customer intelligence, fraud detection, and supply chain improvement. Customer intelligence efforts can be augmented by social media analysis to profile and segment customers and big data can be used to derive predictive models of customer behavior. Business innovations are improvements such as data-driven R&D and new product development. Big data also allows for business creation by decreasing the barriers to entry and better identification of profit signals in the marketplace. ("Data equity", 2012)

In healthcare, Big Data techniques have created new ways for researchers, pharmaceutical companies, and patients to manage conditions. Vast amounts of data are being released for analysis, such as clinical data, cost and claims data, and patient sentiment and behavior data. For example, Asthmapolis, with the help of GPS-enabled asthma inhalers have allowed the collection of population and location data that is merged with CDC data about known asthma causing substances to create individualized treatment plans and other mitigation opportunities. However, this new era of data openness has raised serious questions with regard to patient privacy. The old collective mindset of how patient data is treated is being shifted from “protect” to “share, with protections”.

As data comes more readily available and shared healthcare providers will be under more scrutiny to protect private patient data and may be at a greater risk for lawsuits or other penalties.

(Groves, Kayyali, Knott & Van Kuiken, 2013) Also, players in healthcare data may try to exploit data for their own benefit without regard to what is best clinically. For example, MRI machine owners may use Big Data to identify potential business opportunities that may not be medically necessary. (Groves, Kayyali, Knott & Van Kuiken, 2013)

Big Data is also changing the way in which healthcare is paid. Outcome-based reimbursement models that incentivize healthcare providers to make positive changes to patients' conditions are replacing the old "fee for service" model of healthcare reimbursement that rewarded volume of patients seen. Under the "fee for service", outcome data was never readily available, but over the past ten years, risk-sharing arrangements between payers and providers have incentivized positive patient outcomes or total cost controls. (Groves, Kayyali, Knott & Van Kuiken, 2013)

Potential uses of Big Data could be enormous. "Smart cities" have been envisioned where sensors throughout the city gather data on the environment, financial transactions, transportation and population movements in real time to give a new insight on how cities function and how best to implement services and foster growth. On demand transportation strategies, energy management, and better health initiatives have all been potential areas where Big Data can play in the city of the future. ("Big Data", 2012).

Another area where Big Data could play a huge role is in our education system. McKinsey Global did a study on some certain sectors that could potentially have gains from the use of big data using historical productivity growth data from 2000 to 2008. After their analysis, they concluded that the education sector had room to improve by using Big Data, but that there were obviously some barriers that would need to be worked through before becoming effective. An example of what could be used in increasing the productivity in the education sector would

be to evaluate teachers' effectiveness in increasing the students' academic performance.

(Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh & Byers, 2011)

Although there is extreme potential in Big Data techniques, there is the risk of "data dredging". "Data dredging" is a derogatory term related to technique of "data mining" where large datasets are scoured for connections and relationships that ultimately prove to be spurious and give false insights. The importance of understanding the fundamental subject matter is as important as it is in traditional statistical and data analysis, if not more so: "No matter how much data exists, researchers still need to ask the right questions to create a hypothesis, design a test, and use the data to determine whether that hypothesis true". (Shaw, 2014) The vast size of Big Data datasets increase the chances that some relationships are identified and without a good understanding of the underlying fundamental subject matter, those will not get ruled out.

Methodology

A SWOT-style analysis was performed on Big Data drawn from the analysis presented in the previous literature review. A SWOT analysis is a framework for evaluating the strengths, weaknesses, opportunities, and threats that an idea or concept may have in a simple and straightforward manner.

Strengths

Organizations are just now starting to incorporate Big Data into their processes and realize benefits. The authors of the report "Big data, analytics and the path from insights to value" found that the use of Big Data and its analytics would make the business twice as likely to be a top performer in their market. (LaValle, Lesser, Shockley, Hopkins & Kruschwitz, 2011)

Another strength that Big Data has is the amount of customer feedback data that has become available through social media networks such as Facebook and Twitter. TDWI Research

found that better targeted social influence marketing was at the top of the list when they asked what benefits would come from implementing some sort of Big Data analytics. (Russom, 2011) This demand for Big Data analytics will be even greater as more and more people join the social media networks in the future. For example, there were more than 30 billion pieces of content shared in 2011 on Facebook alone each month. (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh & Byers, 2011)

Weaknesses

As Big Data analytics becomes more popular and a more standard part of modern business processes, there will need to be more training and knowledge transfer for the small to medium enterprises such that they are able to analyze the data they collect to gain a better insight into what their customers want and need. (“Big Data”, 2012) “There aren’t enough people comfortable dealing with petabytes of data”, a quote from Nathan Eagle. Nathan also believes that Big Data should start being incorporated into all aspects of an undergraduate degree so that more graduates have at least a high level of understanding in the field. (Shaw, 2014)

Another major weakness of Big Data is the risk of poor quality conclusions and insights gathered from the data. With the large amounts of data available in this kind of analysis, data scientists could draw false associations between datasets from chance occurrences of correlation. Nathan Eagle quotes the “data dredging” problem as: “you don’t get good scientific output from throwing everything against the wall and seeing what sticks”. (Shaw, 2014) So understanding the data that you are looking at and the quality of the data that you are analyzing could be a major weakness, especially if you don’t have much experience in the field from which the dataset was collected.

Opportunities

Big Data has an exciting set of opportunities in many different industries. A company called PASSUR Aerospace has a service that collects data from sources such as flight schedules, weather, and other information it collects about all of the other planes in the air over 10 times a minute. PASSUR then uses this data to help make more accurate estimations on when flights are going to arrive at airports so that there is static time where money isn't being made. PASSUR only had 155 installations in 2012, so there is still plenty of room for expansion of this technology. (McAfee & Brynjolfsson, 2012)

Another area where Big Data could benefit a good majority of the population would be in the education sector. Having the ability to process the data to see if teachers are being effective in making the performance of students increase would not only increase the testing scores to make the school system look better, but it would make for a more productive, educated future workforce.

Threats

As more and more data is collected, there is a risk that some of this data could be used inappropriately. For example, in the healthcare field, if a third party were analyzing data, the data would have to be striped clean of certain identifying information. Leaving someone's name or other personally identifiable information in a dataset that you send outside of the company could not only endanger the client, whether it be from identity theft or some type of fraud scheme, but it could also have an impact on the company that released the information under the HIPAA Act.

Another threat is that the data collected could become siloed and segregated from other datasets and organizations. In the past, databases have been structured along with appropriate models and software created to analyze this data efficiently. The limited amount of data

collected and stored made the data sets manageable and maintainable by an organization.

However, most, if not all, of the data sets used in Big Data analytics is extremely unstructured.

Having unstructured data means that there is a need to spend resources to clean and “scrub” data before any processing can be completed. Table I summarizes the SWOT components of the analysis.

Table I

Strengths	Weaknesses	Opportunities	Threats
<ul style="list-style-type: none">● Growing interest in the use of Big Data● Large amounts of data is being collected and liberated	<ul style="list-style-type: none">● Lack of knowledge and understanding● Data quality● Analysis quality	<ul style="list-style-type: none">● Smarter products and services● Benefits for more than just businesses	<ul style="list-style-type: none">● Privacy concerns may cause public/private opposition to Big Data● Siloed data

Conclusions and Implications

Based on the previous literature review and SWOT-style analysis, Big Data is an exciting field with many potential applications. From its roots in the traditional fields of statistics and data analytics, Big Data combines modern information technology advancements to gather, store and analyze volumes of data in near real time. However, while the amount of data available and technology allows for never-before realizable applications, care must be taken to ensure that analysis is done soundly with appropriate safeguards to protect sensitive data.

References

- Betsy Page Sigman et al. "Teaching Big Data Experiences, Lessons Learned, and Future Directions." Decision Line, January 2014.
- Big Data - A World of Opportunities. (2012, December). Retrieved from http://www.nessi-europe.com/Files/Private/NESSI_WhitePaper_BigData.pdf
- Big data defined. (n.d.). Retrieved from http://www.sas.com/en_us/insights/big-data/what-is-big-data.html
- Cvijanovic, V. (2012, May 10). The three most common questions I get asked about big data analytics [Web log message]. Retrieved from <http://www.datameer.com/blog/uncategorized/the-three-most-common-questions-i-get-at-datameer.html>
- Data equity - unlocking the value of big data. (2012, April). Retrieved from <http://www.sas.com/offices/europe/uk/downloads/data-equity-cebr.pdf>
- Groves, P., Kayyali, B., Knott, D., & Van Kuiken, S. (2013, January). The Big Data Revolution in Healthcare. Retrieved from http://www.mckinsey.com/insights/health_systems_and_services/the_big-data_revolution_in_us_health_care
- LaValle, S., Lesser, E., Shockley, R., Hopkins, M., & Kruschwitz, N. (2011, February 20). Big data, analytics and the path from insights to value. Retrieved from <http://www.analytics-magazine.org/special-reports/260-big-data-analytics-and-the-path-from-insights-to-value.pdf>
- Lohr, S. (2012, February 11). The age of big data. Retrieved from http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?pagewanted=all&_r=0
- Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. (2011, May). Big data: The next frontier for innovation, competition, and productivity. Retrieved from http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation
- McAfee, A., & Brynjolfsson, E. (2012, October). Big data: The management revolution. Retrieved from <http://hbr.org/2012/10/big-data-the-management-revolution/ar>
- Russom, P. (2011). Big data analytics. Retrieved from <http://public.dhe.ibm.com/common/ssi/ecm/en/iml14293usen/IML14293USEN.PDF>

Shaw, J. (2014, March-April). Why “big data” is a big deal. *Harvard Magazine*, Retrieved from <http://harvardmagazine.com/2014/03/why-big-data-is-a-big-deal>