

Chapter 112

Emerging Role of Big Data in Public Sector

Amir Manzoor
Bahria University, Pakistan

ABSTRACT

Technological advancements have made it easier to collect and store data. We are generating and storing data on a nearly pervasive basis and across multiple environments including work and home. Big data, a general term for the massive amount of digital data being collected from all sorts of sources, is too large, raw, or unstructured for analysis through conventional relational database techniques. For public managers, big data represents an opportunity to infuse information and technology into the design and management of organizations, personnel, and resources. Although the business sector is leading big-data-application development, the public sector has begun to derive insight to help support decision making in real time from fast-growing in-motion data from multiple sources. This chapter explores the big-data applications associated with the public sector and provide suggestions for follower governments.

1. INTRODUCTION

Collection and storage of data has now become easier with recent technological advancements. We generate and store vast amounts of data on daily basis. With dropping costs of storage devices, the costs of data storage have gone down. Today's organizations need to come up with strategies to leverage this data for developing insights that could help in evidence-based decision making.

Big data is general term that refers to the massive amounts of data collected from many sources including the web and the cloud. However, in its raw form, big data is not suitable for analysis using conventional relational database-based techniques of data analysis. It is estimated that around 2.5 quintillion bytes of data is added each day to the current amounts of data and approximately 90% of this new data is unstructured (Kim, Trimi, & Chung, 2014). This overwhelming amounts of big data offers new opportunities for discovery, value creation, and rich business intelligence for decision support in any organization. At the same, its use presents a new set of challenges involving complexity, security, risks to privacy, and need of new skills and technology. Big data has redefined the landscape of data manage-

DOI: 10.4018/978-1-4666-9840-6.ch112

Emerging Role of Big Data in Public Sector

ment. From the processes of extraction, transformation, and loading of data to cleaning and organizing this unstructured data we need new technologies (such as Hadoop) and skills.

Much of big data's growth has come from necessity. Data has been stored in various forms throughout history. One example is US census that was conducted in 1880. The US Census Bureau collected demographic information of 50 million people. The demographic information was collected about more than ten aspects. The data collected was logged by hand, microfilmed, and took approximately eight years for proper tabulation. Machine-readable punch cards were introduced in 1890 to streamline census data collection. With these punch cards, proper tabulation of census data took approximately one year. In the most recent census conducted in 2010, US Census Bureau employed a wide range of technologies to collect census data. These technologies include geographic information systems, social media, videos, intelligent character recognition systems, and sophisticated data-processing software. These technologies were instrumental in establishing a communication mechanism that reached every citizen.

There exist many big data success stories in private sector. In 2012, Merck, one of the largest pharmaceutical companies in the world, used specialized weather forecasts to anticipate hay fever in the year to come. These specialized weather forecasts provided both historical and current weather reports. Based on the analysis of this weather data, Merck concluded that the average pollen counts would be higher in upcoming warm weather in May. As such they anticipated an increased demand for allergy medication. This information was used in developing Merck's business strategies. Merck also used this information to prepare promotional material and establish a partnership with Wal-Mart. In this partnership, Merck developed personalized promotions based on zip codes. As a result of these personalized promotions, Merck's quarterly sales increased.

General motors used telematics for its OnStar technology. Telematics is the blending of computers and wireless telecommunications technologies, ostensibly with the goal of efficiently conveying information over vast networks to improve a host of business functions or government-related public services. OnStar is the leading provider of in-vehicle safety, security and connectivity services, and is now offered on nearly every GM vehicle sold in the U.S., Canada and China. OnStar technology provided various services (such as vehicle security) to assist drivers. OnStar technology also provided vast amounts of telemetry data gathered from the users. OnStar and GMAC insurance developed a joint program that utilized this telemetry data and offered low insurance premium to those users who travelled less. This program resulted in increased customer satisfaction. Top performing private companies use analytics five times more than the poorly performing companies. These firms are also able to make decisions based on rigorous analysis twice as fast as the poorly performing companies (Desouza, 2014). For public sector, big data is an emerging opportunity to streamline business processes, increased citizen engagement, innovate, and embrace evidence driven decision-making.

For public sector, big data is an opportunity to infuse technology into various aspects of organization including organizational design and management. The advancements in the fields data mining and information visualization has provided many resources to parse and traverse the vast amount of data available today that is spread across many networks and come from many resources (Stowers, 2013). It is expected that more than 1.5 million trained big data managers would be needed to leverage all the data generated globally on daily basis (Manyika et al., 2011). Data scientist is a new term that is now recognized as a key future skill organizations will seek (Desouza, 2014).

It can be said that big data in the public sector got the attention in the aftermath of the terrorist attacks on September 11, 2001. The 9/11 commission report identified many deficiencies in US ability to analyze the vast amount of data available. Such deficiencies were one significant cause that these

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/emerging-role-of-big-data-in-public-sector/150274

Related Content

TripRec: An Efficient Approach for Trip Planning with Time Constraints

Heli Sun, Jianbin Huang, Xinwei She, Zhou Yang, Jiao Liu, Jianhua Zou, Qinbao Song and Dong Wang (2015). *International Journal of Data Warehousing and Mining* (pp. 45-65).

www.irma-international.org/article/triprec/122515

Graph-Based Modelling of Concurrent Sequential Patterns

Jing Lu, Weiru Chen and Malcolm Keech (2010). *International Journal of Data Warehousing and Mining* (pp. 41-58).

www.irma-international.org/article/graph-based-modelling-concurrent-sequential/42151

An Analysis on E-Learning and Its Recommendations

Lalitha T. B. and Sreeja P. S. (2021). *New Opportunities for Sentiment Analysis and Information Processing* (pp. 166-187).

www.irma-international.org/chapter/an-analysis-on-e-learning-and-its-recommendations/286910

Automated Diagnostics of Coronary Artery Disease: Long-term Results and Recent Advancements

Matjaž Kukar, Igor Kononenko and Ciril Grošelj (2013). *Data Mining: Concepts, Methodologies, Tools, and Applications* (pp. 1043-1063).

www.irma-international.org/chapter/automated-diagnostics-coronary-artery-disease/73483

Machine Learning Based Admission Data Processing for Early Forecasting Students' Learning Outcomes

Nguyen Thi Kim Son, Nguyen Van Bien, Nguyen Huu Quynh and Chu Cam Tho (2022). *International Journal of Data Warehousing and Mining* (pp. 1-15).

www.irma-international.org/article/machine-learning-based-admission-data-processing-for-early-forecasting-students-learning-outcomes/313585