Chapter 117 A Hybrid Approach to Big Data Systems Development

Anil Aggarwal University of Baltimore, USA

ABSTRACT

Data has always played a critical part in business decision making. The digital economy is generating Tsunami of data which must be analyzed and used by both the public and private sector. Survival and citizen satisfaction may depend on how governments use big data to develop citizen-centric services. Big data analysis can lead to better transparency, less corruption and citizen satisfaction. Big data is an emerging area where models and applications are still emerging. Currently there are few, if any, models that provide guidance in developing applications. This chapter proposes a hybrid approach which can be used as a starting point for future development.

INTRODUCTION

Big data is an emerging area which has only recently caught the attention of researchers. Data has always played a critical role in decision making, and these analyses typically use structured (or well-behaved) data. In the digital age, however, there is an abundance of untapped unstructured data that needs to be mined and analyzed for useful information. This data is generated via social networking, sensors, mobiles, apps and many smart devices. Much of this data contains important information that is still unexplored. The volume, speed and variety with which this data is generated is prompting both the public and private sectors to think about new ways of managing and analyzing data--big data analytics is an outgrowth of this thinking. Literature has typically focused on social networks and data mining but has not merged the areas to provide meaningful insights. Big data combines many disciplines such as analytics, statistics, database, sociology, etc., in order to provide useful information. Many isolated big data applications have been reported in literature but there is no model that provides guidance in developing these applications. This chapter proposes a hybrid approach to developing big data applications, as it combines both structure and unstructured data in its model. This chapter uses the ethical lens theory proposed by William May in developing the hybrid model, and also describes several current applications of big data.

DOI: 10.4018/978-1-5225-7501-6.ch117

What is "Big" Data?

Big data, by definition, is "big" and that's where the definition ends, but what is big is still being debated. Big is a moving target that changes with time as more and more data is being generated at lightning speed. By some estimates (Gartner group), big data will grow at a 45% rate to 35 zettabytes annually by 2020. According to estimates from emarsysglobal.com, 21 billion Short Message Services (SMS) are sent and 1 billion users visit YouTube every day, while 80% of online content is user generated. There are 1 billion Facebook members and almost 700 million Twitter users, which generates almost 600,000 tweets every second. Figure 1(a) and 1(b) shows big data and HADOOP searches from 2005-2015 and beyond from both the general and government perspectives. It is clear from Figures 1(a) and 1(b) that interest in these two concepts started to rise in 2011, with the trends continuing a steep upward for both big data and HADOOP.

Figure 1.



b)

16 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/a-hybrid-approach-to-big-data-systemsdevelopment/217942

Related Content

A Service Recommendation Algorithm Based on Modeling of Dynamic and Diverse Demands

Yanmei Zhang, Tingpei Leiand Zhiguang Qin (2018). *International Journal of Web Services Research (pp. 47-70).*

www.irma-international.org/article/a-service-recommendation-algorithm-based-on-modeling-of-dynamic-and-diversedemands/193861

Drivers of Cloud Computing Adoption for E-Government Services Implementation

Fathey Mohammedand Othman Bin Ibrahim (2019). Web Services: Concepts, Methodologies, Tools, and Applications (pp. 1444-1459).

www.irma-international.org/chapter/drivers-of-cloud-computing-adoption-for-e-government-servicesimplementation/217896

Quality Measures for Semantic Web Application

Adiraju Prasanth Rao (2019). Web Services: Concepts, Methodologies, Tools, and Applications (pp. 1907-1916).

www.irma-international.org/chapter/quality-measures-for-semantic-web-application/217921

Anomaly Detection Algorithm Based on Subspace Local Density Estimation

Chunkai Zhangand Ao Yin (2019). *International Journal of Web Services Research (pp. 44-58).* www.irma-international.org/article/anomaly-detection-algorithm-based-on-subspace-local-density-estimation/231449

Data Mining in Web Services Discovery and Monitoring

Richi Nayak (2008). *International Journal of Web Services Research (pp. 63-81).* www.irma-international.org/article/data-mining-web-services-discovery/3115