Challenges and Issues in Unstructured Big Data: A Systematic Literature Review

Nur Syafiqah Mohd Nafis[1], Suryanti Awang[1]
[1] Soft Computing & Intelligent System Research Group (SPINT), Faculty of Computer System & Software Engineering (FSKKP), Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300, Kuantan, Pahang, Malaysia
Corresponding author Email: suryanti@ump.edu.my
Received:13 June 2017 Accepted:18 September 2017

In past years, data mining had been a standard approach to extract hidden knowledge or pattern known as Knowledge Discovery in Database (KDD). However, KDD approach is unable to handle a terabytes or petabytes of 'Big Data'. The presence of Big Data term arises due to the growth of data volume, velocity, variety, veracity, and value. The objectives of this study are to explore the various data types available in Big Data and current challenges in Big Data. A systematic literature review is conducted to fulfill the research objectives. Several research questions arise in this study for example what type of data available in Big Data as well as what are the major challenges and issues emerged from previous studies. The outcomes of this review identified three data types in Big Data which are unstructured, semi-structured and structured data. Besides, there are also three greatest challenges in Big Data which are data, processing and management challenge. The result also reveals that process challenge is the most critical challenge in Big Data. This extensive systematic literature review can contribute knowledge to the Big Data area as well as for future research.

Keywords: Data Mining, Big data, Unstructured Data, Structured Data, Computational Intelligence