Important of data management in 4 industrial revolution (4 IR)

Wael Mohamed Shaher Yafooz

Al-Madinah International University – Malaysia

Abstract

The concept of data management is the practice of storing, validating and processing required data to accessibility and reliability of data for its users. The data sources come across web and social services, IoT, sensing, transactions of any online organizations, machines, so on. These huge amounts of data can be found in the servers into structured, unstructured and semi-structured. Moreover, these data are stored few categories such, graph based, documents based, key value based and column based. The purpose of data management is not a goal in itself, rather than the key to innovation and knowledge discovery and to integration and reuse after the data publication process. Many organizations and governmental agencies are beginning to require data management and plans for various experiments. Beyond the collection of data and archival, it includes to 'long-term care' that is valuable digital assets .

Organizations gather unstructured data such internal sources (e.g., sensor data) and external sources (e.g., social media). Therefore, from the emergence of data management technologies and analytics enabled the organizations to process data in their business and innovative processes. One of the techniques is facial recognition technology that empowers to acquire intelligence about store traffic, composition of customers, and store movement patterns. These information's are invaluable of leveraged to decisions product promotions, staffs and for placement. In fact, the traditional data management systems assuming by a user query, that they have enough knowledge of the schema, contents and meaning, and certain the query they wanted to pose, thereafter, the system tries to produce complete and correct results. To handle the sensor data in structural monitoring applications, traditional relational database management systems (RDBMS) employs, however little efforts devoted of data management for fundamental issues. For storing, managing and retrieving large scale of data, Apache H-Base, Apache Cassandra, and MongoDB noted as NoSQL (Not only SQL) database tools have designed to handle unstructured data. NoSQL database systems are significant rather than RDBMS for flexibility and scalability. For sensor network data to handling and managing, Apache Cassandra shown a better performance of scalability from massive IoT data which is NoSQL system . Apache Cassandra supports large scale of data management and processing as well. In this talk, will talk about the important of data management and the techniques that help in data mining and discover the insight from huge amount of data.

Biography

Wael Mohamed Shaher Yafooz is associate professor in information technology department at Al-Madinah international university (MEDIU)-Malaysia. He is a dean of faculty of computer and information technology. He received his bachelor degree in area of computer science from Egypt 2002 while Master of Science in Computer Science from university of MARA Technology (UiTM) 2010 as well as PhD in Computer Science in 2014 from UiTM. He awarded many Gold and silver Medals for his contribution in local and international expo of innovation and invention in area of computer science. Besides, he awarded Excellent Research Award from UiTM. He served as member of various committees in many international conferences. Additionally, he chaired IEEE international conference on smart computing and electronic enterprise 2018. Moreover, he supervised many students at master and PhD level. Furthermore, He delivered and conducted many workshops in the research area and practical courses in data management and visualization. His research interest includes Big Data, Data Mining, Machine Learning and Data Management.