International Journal of
Advances in Scientific Research and Engineering (ijasre)

E-ISSN : 2454-8006

DOI: http://dx.doi.org/10.7324/IJASRE.2017.32526 Vol.3 (10) November-2017

BIG DATA IN CHEMICAL INDUSTRY

Matthew N. O. Sadiku[1], Sarhan M. Musa[1], and Osama M. Musa[2]

[1]Roy G. Perry College of Engineering
Prairie View A&M University
Prairie View, TX 77446

[2]Ashland Inc.
Bridgewater, NJ 08807
U.S.A

ABSTRACT

The overwhelming amount of computerized information that organizations and businesses generate and the process is growing so large that the term "big data" is commonly used to describe the situation. Big data is being used by chemical companies to enhance manufacturing, fine-tune pricing, improve marketing, and support innovation. It is creating great market opportunities in the chemical industry. This paper presents a short essay on how big data is currently being used in the chemical industry.

Key Words: Big Data, Big Data Analytics, Chemical Industry.

---

1. INTRODUCTION

It is widely known that we are experiencing an explosion of data as digitization, automated measurement, and the recording of all types of activities increases. The amount of digital information that organizations, companies, and governments collect and process is growing exponentially. The chemical industries are experiencing the same forces and are contributing to this computer-generated information explosion.

Today, chemical companies use information technology (IT) and data gathered from different sources to improve the way they do business, support growth, and compete in the marketplace. Like other industries, the chemical industry is gathering more data (volume) from different sources (variety). A large volume of data is available to characterize products, customers, and operations. Using the right data and the right tools (analytics) can help companies make the right decisions in real time (velocity).

In recent years, the chemical industry has been facing increasingly severe data analytics challenges, which partly have to do with extracting valuable information from a wealth of raw data. The raw data is stored in different formats across multiple databases and is of limited utility until it is curated and transformed into manageable and useable information, which is not an easy step.

2. OVERVIEW OF BIG DATA

Big data refers to massive amounts of data that are so large that traditional processing tools cannot cope. This data is produced by emails, online transactions, social media (Facebook, YouTube, Google, etc), digital sensors, medical records, and so on. Thus, big data refers to huge amounts of data, often arriving from diverse sources, which is unstructured in nature. The units of data are illustrated in Figure 1 [1]. For example, Twitter generates more than 7 Terabytes (TB) of information daily; the Library of Congress holds about 10 TB of information [1].

Big data can be distinguished from regular data by the 4 V's(Volume, Velocity, Variety, and Veracity) [2]:

(1) Volume: Scale of ever-increasing amount of data generated and collected

(2) Velocity: acquisition and analysis speed of the data stream

(3) Variety: Different forms of data (structured and unstructured) such as text, audio, video, etc.

(4) Veracity: Uncertainty of the data and noise

Thus, big data is high-volume, high-velocity, and high-variety information of variable veracity that requires specialized information processing tools.

Big data applies to data sets of extreme size (e.g. exabytes, zettabytes) which are beyond the capability of the commonly used software tools. The data is too big, arrives and changes too quickly, or does not fit regular database architectures. It may require different strategies and tools for profiling, measurement, assessment, and processing.

The process of examining big data is often referred to Big Data Analytics. It is an emerging field since massive computing capabilities have been made available by e-infrastructures. Analytics include statistical models and other methods that are aimed at creating empirical predictions. Data-driven organizations use analytics to guide decisions at all levels [3]. Big data is used with the right analytic tools to recognize patterns, trends and associations not apparent in smaller data sets.

3. OPPORTUNITIES

Several domains can benefit from the big data phenomena: medicine, education, manufacturing, communication, government, and industry. Big data is becoming a new technology focus in science, healthcare, business, and industry.

Big data can effectively integrate multiple data sources to provide real time information on asset utilization, which enhances operational decisions. It enables companies to transfer products from manufacturing facilities at the right time to various destinations, thereby bringing down the inventory holding costs. Leading chemical companies are beginning to revolutionize the decision-making process with big data by dramatically improving productivity and profitability.

- ***Manufacturing*:** Manufacturing is the production of products using raw materials, labor, machines, and tools. This sector is flooded with massive amounts of data from on-line sensors and other digital devices used on the production lines and in factories [4]. Most chemical manufacturing processes are complex. In some global chemical companies, the manufacturing sector is spread across the globe and the many steps in production can be decentralized [5]. To meet the growing demands in the market and maintain a competitive edge, manufacturers have started looking for ways in which big data can open a new array of opportunities. The chemical manufacturing industry can use big data to address existing challenges and simultaneously create new business opportunities. By applying advanced analytical tools to raise their productivity, manufacturers can increase operating efficiency and enhance product quality and consistency.

- ***Drug development*:** Big data provides an alternative "big data" approach based on the development of electronic databases of chemical substances. The use of big data has facilitated achieving personalized precision medicine.

- ***Smart Technology/Production*:** The chemical industry is moving toward sustainability using smart technology. Big data is used to analyze which components cause the most emissions or pollution and replace them with new products yielding less pollution impact. By installing smart technology to monitor energy consumption, the data collected can be analyzed and used to reduce overall energy utilization. In developed markets, chemical companies can use big data to reduce costs and deliver greater innovation in products and services. Data analytics further helps bring down the volume of waste and boost the overall ROI.

- ***Market analysis:*** The chemical industry is often looking for ways to predict future trends. The industry can leverage big data and advanced analytics to make volume forecasting and pricing decisions. Using big data allows companies to quickly identity the needs and desires of their customers [6], anticipating needs.

Other chemical industry-related opportunities include smart production, improved pricing, supply chain management, marketing, innovation, and enhanced management of the workforce. Big data tools have been used in many areas of chemistry – pharmacology, cheminformatics, molecular modeling, medicinal, environmental, industrial, and toxicological chemistry.

4. CHALLENGE

Although big data has evolved from an "emerging technology" into a mainstream phenomenon impacting virtually every sector of the economy, transforming this into practical action is quite challenging for most industries. Big data has some inherent challenges including volume, velocity, variety, and veracity. The sheer size of the data sets is a challenge. Another challenge is the presence of data "noise," which refers to complex information that makes drawing confident conclusions difficult [7].

The rise of big data also raises fundamental challenges in privacy, security, and data ownership. Concerns are being expressed over the impact that collecting, storing, and processing large amount of data could have on security. Security of big data is a primary concern in many applications. From a security point of view, big data may seriously weaken confidentiality since the key information is often not explicit, it is implicitly embedded in multiple portions of the overall data set. Security is a concern because of the variety and heterogeneity of big data; there is access to data from multiple and diverse domains. Although modern chemical processing is mainly automatic, abnormal situations still rely on human operators to intervene. Abnormal situations are costly and can lead to injury and loss of lives [8].

The chemical industry is one of the most complex and high-risk industries in the world. It is characterized by inherently hazardous processes and volatile markets. These factors increase risk and add to the unpredictability of chemical production, transport, and marketing. The chemical industry has always been subject to numerous regulations. Compliance with these regulations can be costly as a simple change to a product formulation could result in an unintended violation.

5. CONCLUSION

Big data is a broad concept that is influencing the chemical industry in a big way. It is creating block-bluster opportunities for chemical companies and making the industry more sustainable. It has emerged as a vital means for enhancing business activities, achieving better execution of operations, and improved marketing strategy. However, the chemicals industry must realize that big data is more than an incremental change in technology; it is a force which is posed to bring fundamental changes in the global chemicals industry.

Chemistry students should participate in this concept of big data. The chemistry departments should introduce big data tools to their undergraduate and graduate students. It is important that they include big data in their curriculum since it will prove to be the next frontier for virtually every economic sector in the next future.

REFERENCES

1. H. E. Pence and A. J. Williams, "Big data and chemical education," *Journal of Chemical Education*, vol. 93, 2016, pp. 504-508.
2. L. Chiang, B. Lu, and I. Castillo, "Big data analytics in chemical engineering,"
3. *Annual Review of Chemical and Biomolecular Engineering*, vol. 8, June 2017, pp. 63-85.
4. M. N.O. Sadiku, M. Tembely, and S.M. Musa, "Big Data: an Introduction for engineers," Journal *of Scientific and Engineering Research*, vol. 3, no. 2, 2016, pp. 106-108.
5. R. Y. Zhong et al., "Big data for supply chain management in the service and manufacturing sectors: challenges, opportunities, and future perspectives," *Computer & Industrial Engineering*, vol. 101, 2016, pp. 572-591.
6. M. Nino, J. M. Blanco, and A. Illarramendi, "Business understanding, challenges and issues of big data analytics for servitization of a capital equipment manufacturer," *Proceedings of IEEE International Conference on Big Data*, 2015, pp. 1368-1377.
7. B. Stringer, "Big data in the chemicals industry*," ICIS Chemical Business*, Dec. 2013/Jan. 2014, pp. 26-27.
8. Y. Chen, E. Argentinis, and G. Weber, "IBM Watson: How cognitive computing can be applied to big data challenges in life sciences research," *Clinical Therapeutics*, vol. 38, no. 4, April 2016, pp. 688-701.
9. Y. Shu et al., "Abnormal situation management: challenges and opportunities in the big data era," *Computers and Chemical Engineering*, vol. 91, 2015, pp. 104-113.

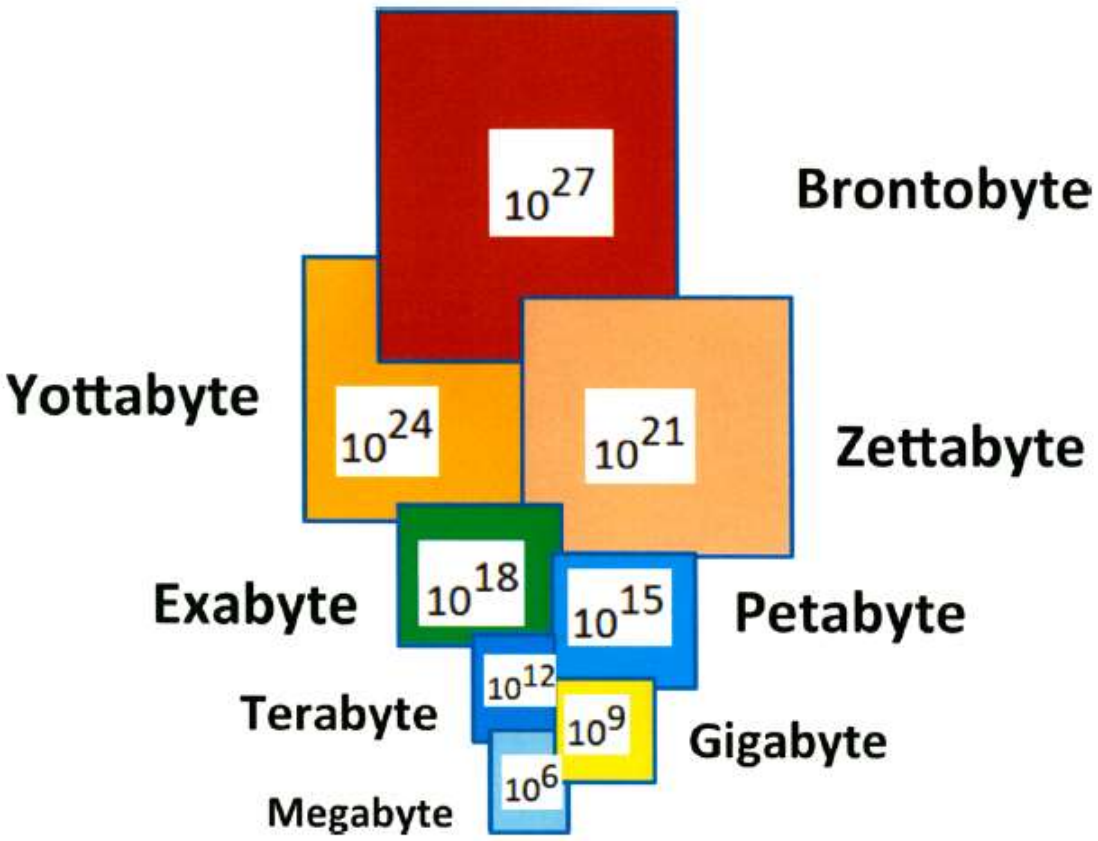

Figure 1. Data units in terms of bytes [1].

AUTHORS

Matthew N.O. Sadiku is a professor in the Department of Electrical and Computer Engineering at Prairie View A&M University, Prairie View, Texas. He is the author of several books and papers. His areas of research interest include computational electromagnetics and computer networks. He is a fellow of IEEE.

Sarhan M. Musa is a professor in the Department of Engineering Technology at Prairie View A&M University, Texas. He has been the director of Prairie View Networking Academy, Texas, since 2004. He is an LTD Sprint and Boeing Welliver Fellow.

Osama M. Musa is currently Vice President and Chief Technology Officer for Ashland Inc. Dr. Musa also serves as a member of the Advisory Board at Manhattan College's Department of Electrical and Computer Engineering as well as a member of the Board of Trustees at Chemists' Club of NYC. Additionally, he sits on the Advisory Board of the International Journal of Humanitarian Technology (IJHT).