Original Research Paper

Big Data in Design and Manufacturing Engineering

[1]**Lidong Wang and** [2]**Cheryl Ann Alexander**

[1]*Department of Engineering Technology, Mississippi Valley State University, USA*
[2]*Technology and Healthcare Solutions, Inc., USA*

Abstract: Big Data helps facilitate information visibility and process automation in design and manufacturing engineering. It also helps analyze trends through analytics and predict inventory, manufacturing output and equipment lifespan and cycles, etc. This paper introduces Big Data, its characteristics and a number of issues of Big Data in design and manufacturing engineering. These issues include design and manufacturing data, Big Data benefits and impacts and its applications and opportunities. Methods, technologies and some technology progress around Big Data are presented in this study. General challenges of Big Data and Big Data challenges in design and manufacturing engineering are also discussed.

Keywords: Big Data, Design, Manufacturing, Sensors, Big Data Analytics, Machine Learning, Data Mining, Map Reduce Algorithm, Hadoop

Introduction

Big Data and Characteristics

The McKinsey study defines Big Data as "datasets whose size is beyond the ability of typical database software tools to capture, store, manage and analyze." Big Data in many sectors ranges from a few dozen Terabytes (TB: Approximately 10^{12} bytes) to multiple Petabytes (PB: Approximately 10^{15} bytes) (Minelli *et al*., 2013). It is too big, moves too fast, or does not fit the strictures of conventional database architectures (Dumbill, 2013). Characteristics (Russom, 2011; Eaton *et al*., 2012; ORRT, 2012; Zikopoulos *et al*., 2011; Demchenko *et al*., 2013) of Big Data can be categorized into "6 Vs". They are: Volume, Velocity, Variety, Value, Variability and Veracity.

Volume

It means data size such as Terabytes (TB), Petabytes (PB), Exabytes (EB: Approximately 10^{18} bytes), Zettabytes (ZB: Approximately 10^{21} bytes) and Yottabyte (YB: Approximately 10^{24} bytes).

Velocity

This relates to how frequently the data is generated. It can be batch, near real time, real time, or streams.

Variety

It represents all types of data such as streamed video, streamed audio and Radio Frequency Identification (RFID) sensor readings. The data type can be structured, unstructured, or semi-structured (Bellini *et al*., 2013; IWT, 2014). Structured data has fixed fields such as spreadsheets or relational databases; unstructured data does not reside in fixed fields-text from articles, email messages, untagged audio or video data, etc.; and semi-structured data does not reside in fixed fields, but it uses tags or other markers to capture elements of the data such as Extensible Markup Language (XML) and Hyper Text Markup Language (HTML)-tagged text (Nedelcu, 2013). Data in variety or different formats makes data integration difficulty or very expensive.

Value

It is defined by the added-value that the collected data can bring. It refers to the value that the data adds to creating knowledge. There is some valuable information somewhere within the data. The valuable information is golden data if it is extracted, although most of the pieces of data individually may seem valueless. Big Data consists of hidden gold (high-valued data) mixed with dirty (noise, erroneous and

Science Publications

raw) data. Big Data technologies can process massive amounts of dirty data and extract the gold information from it. Data value is related to data volume and data variety. The economic value of different data varies depending upon both the source and its end use (Zaslavsky *et al*., 2012; Megahed and Jones-Farmer, 2013; Rajpathak and Narsingpurkar, 2013).

Variability

This refers to the fact that data can be changed at times. It also means data unpredictability and how data may change. Increasing variety and variability also increases the attractiveness of data and the potentiality in providing unexpected, hidden and valuable information (Bellini *et al*., 2013).

Veracity

Veracity (Dove *et al*., 2012; Megahed and Jones-Farmer, 2013; IWT, 2014) includes two aspects: Data consistency (or certainty) and data trustworthiness. Data can be in doubt: Uncertainty due to data inconsistency and incompleteness, ambiguities, latency, deception and model approximations. The following aspects help ensure data veracity:

- Integrity of data and linked data (e.g., for complex hierarchical data, distributed data)
- Data authenticity and (trusted) origin
- Identification of both data and source
- Computer and storage platform trustworthiness
- Availability and timeliness
- Accountability and reputation

Big Data is often dynamic, heterogeneous, inter-related, noisy and untrustworthy. However, even noisy Big Data could be more valuable than tiny sample data because general statistics obtained from frequent patterns and correlation analysis usually overpowers individual fluctuations. Interconnected Big Data forms large heterogeneous information networks; therefore, information redundancy can be explored to compensate for missing data, validate trustworthy relationships, crosscheck conflicting

cases, disclose inherent clusters and uncover hidden models and relationships (Agrawal *et al*., 2011).

Big Data means more information, but it also means more false information. Its focus is on correlations, not causality. It is about what, not why (Bottles *et al*., 2014). In addition, the data we consider big today may not be considered big tomorrow because of the advances in data processing, storage and other system capabilities (Zaslavsky *et al*., 2012).

Design and Manufacturing Engineering Data

Industrial operations and systems often produce a continuous stream of sensor data, event data and contextual data through sensors, smart machines and instrumentation. In a factory, data sources possibly include Computer-Aided Design (CAD) models, Computer-Aided Manufacturing (CAM) models, Computer-Aided Engineering (CAE) models, sensors, instruments, Internet transactions and simulations. The data is often large, fast-moving and complex. The data is often in large variety, including documents, test data, product failure data, CAD/CAM/CAE data, unstructured CAD drawings and specifications and product and process performance data, etc. The increasing volume of data with different types needs to be stored, managed and analyzed. Big Data technologies, driven by innovative analytics, can process large sets of heterogeneous data; and help extract value and hidden knowledge from large and diverse data streams (Noor, 2013; Rajpathak and Narsingpurkar, 2013; Dayal *et al*., 2014).

Some examples of Big Data in manufacturing are shown in Table 1. Structured data has the advantage of being easily entered, stored, analyzed and queried. Examples include manufacturing data stored in relational databases, data from manufacturing execution systems and data from enterprise systems. Unstructured data such as log files and human-operator-generated shift reports may be in a raw format that requires decoding before data values can be extracted. Semi-structured data does not conform to the models of relational databases or other data tables, but contains tags or other markers to separate semantic elements and demonstrate hierarchies of field sand records (IC, 2014).

Table 1. Manufacturing data examples (IC, 2014)

Structured data	Unstructured data	Real-time, semi-structured data
Spreadsheets	Operator shift reports	RFID
Relational databases	Machine logs, error logs	XML
Enterprise data warehouse	Texts, images, audio/video	Machine builder standards like
Files stored in manufacturing PCs	Manufacturing collaboration social platforms	Sensors (vibration, pressure, valve and acoustics), relays
		Manufacturing historians (time series data structures)

Benefits and Impacts of Big Data in Design and Manufacturing Engineering

Mining Big Data can offer benefits in design, such as better detection of defects in design and design improvement, saving design time and costs, fast response to market, developing innovative products and matching customers' needs and gaining customer satisfaction through extracting crucial customer requirements from the customer-generated data to update and refine existing designs (Wu *et al.*, 2015).

The manufacturing sector generates a great deal of text and numerical data in product development processes. Big Data offers the following benefits in manufacturing (Brown *et al.*, 2011; McGuire *et al.*, 2012; Nedelcu, 2013; Noor, 2013; Wu *et al.*, 2015):

Defect tracking and product quality:

- Perform predictive diagnostics for product/part failure
- Monitor product data quality
- Early detect quality problems
- Better detect product defects
- Provide real-time alerts based on analyzing manufacturing data
- Reduces defects during manufacturing processes by tracking every detail about every part that goes into a product
- Boost quality

Improvements in supply planning:

- Unlock significant value and unearth valuable insights by performing Big Data analytics and making information transparent
- Better forecast products, production and manufacturing output
- Better forecast sales volumes through semantic-based Big Data analytics
- Improve relationship with suppliers and conduct better contract negotiations according to collected supplier performance data
- Improve decision-making and minimizes risks in supply

Improved product manufacturing processes:

- Provide an infrastructure for transparency in manufacturing
- Analyze sensor data from production lines, creating self-regulating processes that cut waste, avoid costly (and sometimes dangerous) human interventions and ultimately lift output
- Better monitor and control manufacturing processes by tracking every detail about every part

and procedure, better information visibility and Big Data analytics for data in motion
- Perform predictive manufacturing and optimize manufacturing processes
- Better simulate and test new manufacturing processes

Driven efficiency across the extended enterprise:

- Increase the efficiency of the manufacturing processes
- Increase energy efficiency
- Enable effective and consistent collaboration through integrating datasets from multiple systems and divisions
- Facilitate innovative design for manufacturing and integration of CAD/CAE/CAM; reduce unnecessary iterations in product development cycles; and finally reduce production and development costs
- Offer further opportunities to accelerate product development; increase product innovation and development of next-generation products

Improved service:

- Determine what manufacturing parameters most influence customer satisfaction
- Develop new products and make products better match customers' needs through sentiment analysis and recommendation systems for Big Data
- Enable mass-customization in manufacturing
- Better correlate manufacturing and business performance information together
- Reduce warranty costs through warranty analysis based on Big Data analytics
- Better perform remote intelligent services

Table 2 (MKGI, 2010) shows the degree of potential benefits that Big Data could generate in the areas of greatest benefits for manufacturing/operations.

Applications and Opportunities of Big Data in Design and Manufacturing Engineering

Big Data in General Electric, General Motors and the Automotive Industry

Some manufacturing firms, such as General Electric, view Big Data from sensors in manufactured products (e.g., locomotives, jet engines and gas turbines in GE's case) as key to effective and efficient servicing strategies. In the same mode, automobile manufacturers such as General Motors created self-

driving cars based on the analysis of Big Data from sensors and machine vision technologies (Davenport, 2013). Big Data has become the key asset for the whole production and manufacturing cycle, as well as the provision of services in the automotive and mobility space. Big Data is actually at heart of how the extracted sensor data and location data are combined to provide services (Camilli and Duisberg, 2013).

Big Data in Semiconductor Manufacturing and Integrated Circuits

Semiconductor companies have seen significant opportunities for Big Data and analytics to optimize semiconductor manufacturing (Hattori, 2013). For higher product quality, semiconductor companies conduct extensive tests and collect terabytes of data. Semiconductor vendors mine Big Data for product quality (Neison, 2014).

With the continuous shrink of integrated circuits in feature sizes due to nanotechnologies, extracting manufacturing information and mining valuable intelligence from automatically collected Big Data in the wafer fabrication facilities for real-time decisions and a higher yield has become very important to support intelligent manufacturing, enhance the service quality and maintain competitive advantages for high-tech companies in the global competition (Hsu *et al*., 2012).

Big Data at Work for a Missile Plant

A missile plant (Fig. 1), for instance, Raytheon Corp. in Huntsville, Alabama, USA, monitors its assembly operations down to the turn of ascrew. If a screw in a missile fails to complete its full count of turns, Raytheon will know about it immediately and be able to take corrective steps. Raytheon's monitoring technology is often called "Manufacturing Execution Software (MES)," and several manufacturers have used MES to collect and analyze factory-floor data. The systems enable the real-time control of multiple elements of the production process (Noor, 2013).

Table 2. Greatest benefit areas for manufacturing/operations; (1: No benefits; 3: Moderate benefits; 5: Very high benefits)

Areas of greatest benefit for manufacturing/operations	Degree of benefits
Product quality/defect tracking	3.37
Supply planning	3.34
Manufacturing process defect tracking	3.32
Supplier/supplier component/parts defect tracking	3.11
Collecting supplier performance data to inform contract negotiations	3.08
Forecasting manufacturing output	3.03
Increasing energy efficiency	2.97
Simulation and testing of new manufacturing processes	2.88
Enable mass-customization in manufacturing	2.75

Fig. 1. Big Data at a missile plant (Noor, 2013)

Big Data in Cloud-based Design and Manufacturing

Cloud-Based Design Manufacturing (CBDM) is a service-oriented networked product development model. Based on this model, service consumers can configure, select and utilize customized product realization resources and services. CBDM uses the Internet of Things (IoT) (e.g., RFID), smart sensors and wireless devices (e.g., smart phone) to collect real-time design- and manufacturing-related data. IoT allows engineers to have access to data such as equipment condition, machine utilization and the percentage of defective products from any location. Engineers can use Big Data analytics for forecasting, automation and proactive maintenance (Wu *et al*., 2015).

Technology Integration Based on Big Data for More Value

Manufacturing sector generates data from a multitude of sources such as instrumented production machinery (process control). Most manufacturing companies have Information Technology (IT) systems to manage the product data generated via CAD, CAM, CAE and Product Development Management (PDM) systems. However, the large datasets generated by these systems often remain trapped within their respective systems. Manufacturers can create a significant opportunity to create more value through effective and consistent collaboration, the integration of datasets from multiple systems and Big Data analytics for the integrated datasets (Nedelcu, 2013).

Medical Device Design and Manufacturing

Simulation-based engineering methods involve Finite Element Analysis (FEA), Finite Difference Analysis (FDA), Computational Fluid Dynamics (CFD) and/or multi-physics simulations to work toward an optimal design. Advancements in medical device design and manufacturing that is data-driven and simulation-based have drawn upon and combined emerging work in the areas of regulatory science and Big Data. Big Data technologies enable a broader set of device materials, anatomical configurations, delivery methods and tissue interactions to be evaluated. Computational methods and Big Data technologies can play an important role in medical device design and manufacturing (Erdman and Keefe, 2013).

Big Data and Additive Manufacturing

Understanding the use and implications of Big Data and predictive analytics will be very important as additive manufacturing (also called 3D printing) makes traditional models of production, distribution and demand obsolete in some product areas. Some companies no longer need to invest in dedicated, expensive milling machines or injection molding equipment. A single 3D printer can produce a wide range of parts-without expensive dies and jigs. People can even modify designs to meet special, one-of-a-kind needs and then use an affordable 3D printer to "manufacture" parts (Waller and Fawcett, 2013b).

Production Process Monitoring, Maintenance, Quality Assurance and Logistics for Manufacturers

Production sensors generate vast volume of data at sub-second speeds. Big Data and advanced analytics analyze this vast amount of data to monitor production processes and identify when an event will affect production quality and when maintenance is required, before production quality is actually affected. Big Data provides the ability to ensure Quality Assurance (QA) tests are confirmations of high product quality (Software AG, 2013).

Table 3 (Waller and Fawcett, 2013a) provides some examples of potential applications of Big Data within logistics for manufacturers.

Big Data in CAD/CAE/CAM and CAD Educational Assessment

Big Data working with3D software in CAD/CAE/CAM can greatly help companies, especially companies in the aerospace industry (e.g., the areas of aeronautics and astronautics) because these companies have struggled to manage the constantly growing volume of data. The datasets are large, complex and often fast-moving. It means that these datasets are often in large volume, velocity, variability and variety. A lot of these datasets are unstructured, for example, CAD drawings and CAD/CAE/CAM specifications. Big Data and analytics are powerful in processing and managing these kinds of heterogeneous information, improving data veracity and creating more values.

Table 3. Examples of potential applications of Big Data in logistics for manufacturers

Forecasting	Early response to extremely negative or positive customer sentiment
Inventory management	Reduction in shrink, efficient consumer response, quick response and vendor managed inventory
Transportation management	Improved notification of delivery time and availability; surveillance data for improved yard management
Human resources	More effective monitoring of productivity; medical sensors for safety of labor in factories

A computational method that is based on time series analysis was proposed to assess engineering design processes using a CAD tool. Educational data mining and learning analytics were studied to assess student performance in learning and designing in a project-based setting. The time series process data can be as fine-grained as the 'atomic' design steps (meaning that they cannot be logically divided further). Data at this level of granularity has all the four characteristics of Big Data (IBM, 2012; Xie *et al*., 2014a):

- *High volume*: A large amount of process data in a complex open-ended project
- *High velocity*: The data can be collected, processed and visualized in real time to instantaneously provide to students and teachers
- *High variety*: Many types of information provided by a rich CAD system such as all the learner actions and artifact properties
- *High veracity*: The data must be accurate and comprehensive to ensure fair and trustworthy assessments of student performance

CAD logs are instructionally sensitive and can serve as an effective instrument for assessing complex engineering design processes. High-volume and high-variety software logs can be used to detect the effects of what happens outside the computer on individual students. CAD logs were used in performance assessment because their fine-grained, temporal nature can provide more comprehensive, more reliable and more personalized process data for finding evidence of deep learning related to design creativity and problem solving. Deep learning generates a large amount of datasets. Big Data can be used to analyze and visualize these large datasets (Xie *et al*., 2014b).

'Mobile Access to CAD' is a growing with above average importance and usage. 'Cloud Based CAD' currently has low average importance and below average usage. 'Big Data' is not yet 'big' in CAD-very low awareness (Turner, 2014).

Methods, Technologies and Technology Progress around Big Data

Big Data analytics uses analysis algorithms running on powerful supporting platforms to uncover potentials concealed in Big Data, such as unknown correlations or hidden patterns (Hu *et al*., 2014). Advanced modelling, analysis, feedback and visualization are the techniques of Big Data analytics. These techniques help manufacturing companies eliminate waste and create value in the design and production of the products (Papanagnou, 2014). In addition, data mining, text mining, opinion mining, social network analysis, cluster analysis are Big Data analytics methods (Cho and Hwang, 2015).

Machine learning has been used in Big Data. Massive Parallel-Processing (MPP), distributed file systems and cloud computing, etc. are supporting technologies of Big Data (Zaslavsky *et al*., 2012). Besides general cloud infrastructure services (storage, compute, infrastructure/Virtual Machine (VM) management), the following services are also required to support Big Data (Turk, 2012):

- Cluster services
- Hadoop related services and tools
- Specialist data analytics tools (logs, events, data mining, etc.)
- Databases/Servers SQL, NoSQL
- MPP (Massively Parallel Processing) databases
- Registries, indexing/search, semantics, namespaces
- Security infrastructure (access control, policy enforcement, confidentiality, trust, availability, privacy)

Hadoop is an open source framework for writing and running distributed applications that are capable of batch processing large sets of data. Hadoop framework is mainly known for Map Reduce and its distributed file system. The Map Reduce algorithm consists of two basic operations: map and reduce. It is a distributed data processing model that runs on alarge cluster of machines. Hadoop includes three parts: Hadoop Distributed File System (HDFS), Hadoop Map Reduce and Hadoop Common (Chardonnens, 2013).

To process and manage Big Data with parallel and distributed data mining algorithms, a Cloud-Based Design Manufacturing (CBDM) system should employan open-source software/programming framework that supports data-intensive distributed applications (Ren *et al*., 2012). Map Reduce, a parallel programming model, is one of widely used programming models in cloud computing. It enables CBDM systems to process large datasets. Hadoop is one of the open source implementations of the Map Reduce model. Hadoop divides computationally extensive tasks into small fragments of work and each work unit is processed on a computer node in a Hadoop cluster (Dean and Ghemawat, 2008).

New methods for compact visualization of data with ranging *variety* and *veracity* are constantly being developed to present correlations across bases more effectively. Visualization may be standalone or may cross-filter with other views across feature bases. Visualization may be measured in many ways, including the bases spanned continuously, number of points drawn, level of over plot and precision/correctness (Schwartz *et al*., 2014).

Challenges of Big Data

General Challenges of Big Data

Traditional Statistical Process Control (SPC) methods only focus onnumeric datasets. However, most Big Data applications are required to process non-numeric data obtained from different databases. The modeling of this kind of data is often based on disciplines that are not in the areas of statisticians and quality engineers. As for text data, models draw from linguistic sciences, computer science and psychology. The models maybe integrate data input in different languages. The arrival rate of the data fluctuates depending on factors that are often not understood before analyzing it. This phenomenon is called trending/viral for online content (Megahed and Jones-Farmer, 2013). Big Data has challenges in capture, storage, search, analysis and virtualization (Zaslavsky et al., 2012).The specific information of some general challenges is provided as follows:

- Challenges in Big Data management can be categorized into two types: Engineering and semantic. Engineering challenges lie in performing data management activities such as query and storage efficiently. Semantic challenges lie in extracting the meaning of the information from massive volumes of unstructured data, even dirty data (Bizer et al., 2012)
- It is difficult to collect and integrate data with scalability from distributed locations because of the variety of disparate data sources and the sheer volume (Hu et al., 2014)

- Data quality is a focal point. During the process of data capture, sources of data are often heterogeneous, geographically distributed and unreliable, being susceptible to errors. Therefore, a number of data preprocessing techniques, such as data cleaning, data reduction, data integration and data transformation, are often used to remove noise and correct inconsistencies (Han and Kamber, 2006)
- Big Data systems manage and store the gathered heterogeneous and massive datasets, while providing function and performance guarantee, in terms of scalability, fast retrieval and privacy protection (Hu et al., 2014). Privacy and information security are concerns in Big Data
- Exponential growth of data volume is generated from different instruments and/or collected from sensors; it is necessary, but not easy to consolidate e-Infrastructures as persistent platforms to ensure continuity and cross-disciplinary collaboration (Demchenko et al., 2012)

Big Data Challenges in Design and Manufacturing Engineering

It is important and necessary to integrate CAD/CAE/CAM and cyber-physical systems with Big Data systems to make design and manufacturing more competitive. How to fulfil the integration and how to extract the right information from Big Data in design and manufacturing for the right purpose at the right time are major challenges.

Table 4. Big Data challenges for manufacturing; (1: Not at all a challenge; 3: Moderate challenge; 5: Very high challenge)

Areas of greatest challenges for manufacturing/production	Degrees of challenges
Building high levels of trust between data scientists who present insights on Big Data and functional managers	3.31
Determining what data to use for different business decisions	3.29
Being able to handle the large volume, velocity and variety of Big Data	3.25
Getting business units to share information across organizational silos	3.22
Finding the optimal way to organize Big Data activities in a company	3.20
Getting functional managers to make decisions based on Big Data, rather than on intuition	3.14
Putting the analysis of Big Data in a presentable form for making decisions	3.12
Getting top management in a company to approve investments in Big Data and is related investments	3.11
Determining what to do with the insights that are created from Big Data	3.09
Getting the IT function to recognize that Big Data requires new technologies and new skills	3.08
Finding and hiring data scientists who can manage large amounts of structured and unstructured data and create insights	3.02
Determining which Big Data technologies to use	3.02
Keeping the data in Big Data initiatives secure from external parties	2.98
Understanding where in a company people should focus Big Data investments	2.98
Reskilling the IT function to be able to use new tools and technologies of Big Data	2.95
Other	2.80
Keeping the data in Big Data initiatives secure from internal parties	2.71

Machine learning is an important method of Big Data; however, it has challenges in implementation. One challenge is the availability of the right data from different operations and processes. Machine systems such as Programmable Logic Controllers (PLC) and Supervisory Control And Data Acquisition (SCADA) often capture a lot of machine data, but this data may not be relevant. PLC and SCADA do not store all the data that is required for a Big Data predictive analytics solution based on machine learning (Joseph *et al.*, 2014).

Other major challenges of Big Data in design and manufacturing include building high levels of trust between data scientists and managers; confidence in analyzing and managing data with large volume, velocity and variety; deciding what methods and technologies will be used; and maintaining the consistency of managing and using Big Data, etc. (Nedelcu, 2013).

Table 4 (MKGI, 2010) shows some areas of greatest challenge for manufacturing/production and degrees of these challenges.

Conclusion and Future Research

Big Data is large in volume, velocity, variety, value, variability and veracity. Big Data helps integrate various types of datasets in design and manufacturing engineering; uncover hidden correlation patterns through analytics; improve design and production processes; and create more values. Mining Big Data helps improve design in quality, time, costs and mass-customization. Big Data also offers greatest benefits for manufacturing engineering such as detecting product defects, boosting quality and improving supply planning, etc. It has had a lot of applications or great opportunities in design and manufacturing engineering. These applications or opportunity areas include electricity, automotive, missile plants, integrated circuits, semiconductor manufacturing, additive manufacturing, medical device design and manufacturing and cloud-based design and manufacturing, etc. Generally, Big Data has challenges such as data capture, date integration, data visualization, extracting values from all of heterogeneous data and privacy and information security, etc. Specifically in design and manufacturing engineering, besides the above challenges, other major challenges lie in: Trust between data scientists and managers; confidence in analyzing big data, choice of methods and technologies and consistency of managing and using big data, etc. All the aspects of these challenges can be future research. The authors of the paper will focus on Big Data in CAD/CAE/CAM of medical devices as further research.

Acknowledgment

Funding Information

Author's Contributions

All authors equally contributed in this work.

Ethics

This article has not been published elsewhere. The corresponding author confirms that both authors have read and approved the manuscript and no ethical issues involved.

References

Agrawal, D., P. Bernstein and E. Bertino, 2011. Challenges and opportunities with Big Data. Purdue University Purdue e-Pubs.

Bellini, P., M. di Claudio, P. Nesi and N. Rauch, 2013.Tassonomy and Review of Big Data Solutions Navigation. In: Big Data Computing, Akerkar, R. (Ed.), CRC Press, ISBN-10: 1466578386, pp: 57-101.

Bizer, C., P. Boncz, M.L. Brodie and O. Erling, 2012. The meaningful use of Big Data: four perspectives--four challenges. SIGMOD Rec., 40: 56-60. DOI: 10.1145/2094114.2094129

Bottles, K., E. Begoli and B. Worley, 2014. Understanding the pros and cons of Big Data analytics. Physic. Executive, 40: 6-12. PMID: 25188972

Brown, B., M. Chui and J. Manyika, 2011. Are you ready for the era of 'Big Data'? McKinsey Global Institute.

Camilli, R. and A. Duisberg, 2013. Big Data and the automotive industry. Report, Bird and Bird.

Chardonnens, T., 2013. Big Data analytics on high velocity streams. MSc Thesis, University of Fribourg.

Cho, Y. and S. Hwang, 2015. A study on LBS technique makes for CRETA LBS platform service. Int. J. Signal Process. Syst., 3: 82-87.

Davenport, T., 2013. The future of the manufacturing workforce, report three: Contemporary manufacturing technologies and the talent to use them. Technical Report, Manpower Inc.

Dayal, U., M. Akatsu, C. Gupta, R. Vennelakanti and M. Lenardi, 2014. Expanding global Big Data solutions with innovative analytics. Hitachi Rev., 63: 333-339.

Dean, J. and S. Ghemawat, 2008. MapReduce: Simplified data processing on large clusters. Commun. ACM, 51: 107-113. DOI: 10.1145/1327452.1327492

Demchenko, Y., Z. Zhao, P. Grosso, A. Wibisono and C. de Laat, 2012. Addressing Big Data challenges for scientific data infrastructure. Proceedings of the IEEE 4th International Conference on Cloud Computing Technology and Science, Dec. 3-6, IEEE Xplore Press, Taipei, pp: 614-617. DOI: 10.1109/CloudCom.2012.6427494

Demchenko, Y., P. Grosso, C. de Laat and P. Membrey, 2013. Addressing Big Data issues in Scientific Data Infrastructure. Proceedings of the International Conference on Collaboration Technologies and Systems, May 20-24, IEEE Xplore Press, San Diego, CA., pp: 48-55. DOI: 10.1109/CTS.2013.6567203

Dove, E.S., S.A. Faraj, E. Kolker and V. Ozdemir, 2012. Designing a post-genomics knowledge ecosystem to translate pharmacogenomics into public health action. Genome Med., 4: 91-91. DOI: 10.1186/gm392

Dumbill, E., 2013. Making sense of Big Data. Big Data.

Eaton, C., D. Deroos, T. Deutsch, G. Lapis and P. Zikopoulos, 2012. Understanding Big Data. McGraw-Hill Companies.

Erdman, A.G. and D.F. Keefe, 2013. Grand challenge: Applying regulatory science and Big Data to improve medical device innovation. IEEE Trans. Biomed. Eng., 60: 700-706. DOI: 10.1109/TBME.2013.2244600

Han, J. and M. Kamber, 2006. Data Mining, Southeast Asia Edition: Concepts and Techniques. 2nd Edn., Morgan Kaufmann, Amsterdam, ISBN-10: 0080475582, pp: 800.

Hattori, R., 2013. Big Data and analytics for semiconductor manufacturing. Technical Report, IBM, pp: 1-25.

Hsu, C.Y., C.F. Chien and P.N. Chenb, 2012. Manufacturing intelligence for early warning of key equipment excursion for advanced equipment control in semiconductor manufacturing. J. Chin. Inst. Ind. Eng., 29: 303-313. DOI: 10.1080/10170669.2012.702135

Hu, H., Y. Wen, T.S. Chua and X. Li, 2014. Toward scalable systems for Big Data analytics: A technology tutorial. IEEE Access, 2: 652-687. DOI: 10.1109/ACCESS.2014.2332453

IBM, 2012.What is Big Data? IBM.

IC, 2014. Optimizing manufacturing with the internet of things. Intel Corporation.

IWT, 2014. From predictive maintenance to predictive manufacturing: Big Data challenge for industry. POM Prognostics for Optimal Maintenance.

Joseph, J., O. Sharif, A. Kumar, S. Gadkari and A. Mohan, 2014. Using Big Data for machine learning analytics in manufacturing. Tata Consultancy Services.

McGuire, T., J. Manyika and M. Chui, 2012. Why Big Data is the new competitive advantage. Ivey Bus. J.

Megahed, F.M. and L.A. Jones-Farmer, 2013. A Statistical Process Monitoring Perspective on "Big Data", Frontiers in Statistical Quality Control. 11th Edn., Springer, New York.

Minelli, M., M. Chambers and A. Dhiraj, 2013. Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Business. 1st Edn., John Wiley and Sons, New Jersey, ISBN-10: 111814760X, pp: 187.

MKGI, 2010. The internet of things. McKinsey Global Institute.

Nedelcu, B., 2013. About Big Data and its challenges and benefits in manufacturing. Database Syst. J., 4: 10-19.

Neison, R., 2014. Semiconductor vendors mine Big Data for quality. Evaluat. Eng., 53: 24-24.

Noor, A.K., 2013. Putting Big Data to work. Department of Mechanical Engineering.

ORRT, 2012. Planning for Big Data. O'Reilly Radar Team.

Papanagnou, C., 2014. The role of Big Data exploration and cloud-based technologies in manufacturing. High Speed Sustainable Manufacturing Institute.

Rajpathak, T. and A. Narsingpurkar, 2013. Managing knowledge from Big Data analytics in product development. White Paper, Tata Consultancy Services.

Ren, L., L. Zhang, F. Tao, X. Zhang and Y. Luo *et al.*, 2012. A methodology towards virtualisation-based high performance simulation platform supporting multidisciplinary design of complex products. Enterprise Inform. Syst., 6: 267-290. DOI: 10.1080/17517575.2011.592598

Russom, P., 2011. Big Data analytics. TDWI Best Practices Report, Fourth Quarter.

Schwartz, J., C. Joshua and E. Poore, 2014. Measuring the value of Big Data exploitation systems: Quantitative, non-subjective metrics with the user as a key component. Parsons J. Inform. Mapp., 6: 1-12.

Software AG, 2013. The value of Big Data in the supply chain. White Paper, Software AG.

Turk, M., 2012. A chart of the Big Data ecosystem, take 2.

Turner, C., 2014. Business advantage announces results of worldwide CAD trends survey. Report, Business Advantage Company.

Waller, M.A. and S.E. Fawcett, 2013a. Data science, predictive analytics and Big Data: A revolution that will transform supply chain design and management. J. Bus. Logist., 34: 77-84. DOI: 10.1111/jbl.12010

Waller, M.A. and S.E. Fawcett, 2013b. Click here for a data scientist: Big Data, predictive analytics and theory development in the era of a maker movement supply chain. J. Bus. Logist., 34: 249-252. DOI: 10.1111/jbl.12024

Wu, D., D.W. Rosena, L. Wang and D. Schaefer, 2015. Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation. Computer-Aided Des., 59: 1-14. DOI: 10.1016/j.cad.2014.07.006

Xie, C., Z. Zhang, S. Nourian and E. Hazzard, 2014a. Time series analysis method for assessing engineering design processes using a CAD tool. Int. J. Eng. Educ., 30: 218-230.

Xie, C., Z. Zhang, S. Nourian, A. Pallant and S. Bailey, 2014b. On the instructional sensitivity of CAD logs. Int. J. Eng. Educ., 30: 760-778.

Zaslavsky, A., C. Perera and D. Georgakopoulos, 2012. Sensing as a service and Big Data. Proceedings of the International Conference on Advances in Cloud Computing, (ACC' 12), Bangalore, India, pp: 1-8.

Zikopoulos, P.C., C. Eaton and P. Zikopoulos, 2011. Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data. 1st Edn., McGraw-Hill, New York, ISBN-10: 0071790535, pp: 176.