Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 21 (2018) 141–148

Procedia MANUFACTURING

www.elsevier.com/locate/procedia

15th Global Conference on Sustainable Manufacturing

A road map for applied data sciences supporting sustainability in advanced manufacturing: the information quality dimensions

Ron S. Kenett[a,b,c]*, Avigdor Zonnenshain[b], Gilead Fortuna[b]

[a] *KPA Ltd., Raanana, Israel*
[b] *Samuel Neaman Institute, Technion, Israel*
[c] *University of Turin,, Turin, Italy*

Abstract

Data science is a multidisciplinary blend of data inference, algorithm development, and technology in order to solve analytically complex problems. Sustainability is a critical asset of a manufacturing enterprise. It enables a business to differentiate itself from competitors and to compete efficiently and effectively to the best of its ability. This paper is a review of data analytics, and how it supports advanced manufacturing with an emphasis on sustainability. The objective is to present a context for a roadmap for applied data science addressing such analytic challenges. We start with a general introduction to advanced manufacturing and trends in modern analytics tools and technology. We then list challenges of analytics supporting advanced manufacturing and sustainability aspects. The information quality (InfoQ) framework is proposed as a backbone to evaluate the analytics needed in advanced manufacturing. The eight InfoQ dimensions are: 1) Data Resolution, 2) Data Structure, 3) Data Integration, 4) Temporal Relevance, 5) Chronology of Data and Goal, 6) Generalizability, 7) Operationalization and 8) Communication. These dimensions provide a classification of advanced manufacturing analytics domains. The paper provides a roadmap for the development of applied analytic techniques supporting advanced manufacturing and sustainability. The objective is to motivate researchers, practitioners and industrialists to support such a roadmap.

Peer-review under responsibility of the scientific committee of the 15th Global Conference on Sustainable Manufacturing (GCSM).

Keywords: Advanced manufacturing, Industry 4.0, Information Quality (InfoQ), Data Science, Analytics.

1. Introduction

In this paper, we present a roadmap for data science and applied statistics with an emphasis on industrial applications and challenges posed by Industry 4.0. Data is increasingly cheap and ubiquitous. The rise of "big data"

* Corresponding author. Tel.: +97297408442.
E-mail address: ron@kpa-group.com

Peer-review under responsibility of the scientific committee of the 15th Global Conference on Sustainable Manufacturing (GCSM).
10.1016/j.promfg.2018.02.104

has the potential to deepen our understanding of phenomena ranging from physical and biological systems to human social and economic behavior. Working with this data requires distinctive new skills and tools. The data is also more heterogeneous than the highly curated data of the past. Digitized text, audio, and visual content, like sensor and blog data, is typically messy, incomplete, and unstructured; it is often of uncertain provenance and quality; and frequently must be combined with other data to be useful. Working with user-generated data sets also raises challenging issues of privacy, security, and ethics. A system approach to the design for sustainability is presented in [27]. The paper provides an overview of the challenges of systems engineering in designing for sustainability by involving all interested stakeholders, dealing with the entire life cycle value chain of products, practicing corporate social responsibility and managing the relevant risks and opportunities. In parallel with these developments, the last several years have seen a significant growth in the number of courses and new programs titles “Data Science”, “Business Analytics”, “Predictive Analytics”, “Big Data Analytics”, and related titles. Different programs have a different emphasis depending on whether they are housed in a business school, a computer science department, or a cross-departmental program. What is however common to all of them, is their focus on data (structured and unstructured), and specifically, on data analysis. Many Statistics and Operations Research programs and departments have been restructuring, revising, and rebranding their courses and programs to match the high demand for people skilled in data analysis (see [3]).

Applied statistics is about meeting the challenge of solving real world problems with mathematical tools and statistical thinking. In a paper discussing the expanded role of statistics in research, business, industry, and service organizations. The paper in [9] proposes a life cycle view that combines an inductive-deductive learning process. The main point in this proposal is that effective statistical work is much more than properly applying statistical methods. One needs to emphasize that statistical analysis is a collaborative venture whose success depends essentially on the effectiveness of the communication between the statistician and the client. The statistician needs to exercise social and communicative skills for his work to have an impact. In addition, a systematic assessment of the impact and the quality of information generated by the statistical analysis requires additional activities not usually performed by statisticians. A particularly important area of application to data science and applied statistics is industry and manufacturing operations. The first industrial revolution was triggered by the introduction of the steam engine and the mechanization of manual work in the 18th century. Electricity drove mass production in the second industrial revolution in the early 20th century. The third revolution in manufacturing was due to the use of electronics and computer technology for manufacturing and production automation. We are now entering the fourth phase, labeled advanced manufacturing or Industry 4.0. For a perspective on the evolution of production and quality conceptual frameworks, from inspection to process improvement to quality by design, see chapter 1 in [13].

2. Data Science and Applied Statistics in Advanced Manufacturing

Advanced manufacturing requires analytics and operational capabilities to interface to devices in real time, at an individual level. Software development has become agile and commonly applied DevOps operations provide continuous delivery (see [14]). Moreover, processing and analytic models evolve to provide a high level of agility as organizations realize data agility, the ability to understand data in context and take the right business action, is the source of competitive advantage (see [19]). The emergence of agile processing models enables the same instance of data to support batch analytics, interactive analytics, global messaging, database and file-based models. More agile analytic models are also enabled when a single instance of data can support a broader set of tools. The result is an agile development and application platform that supports the broadest range of processing and analytic models. Tools and methods implemented to verify if process behavior is consistent with normal operating conditions include functionalities such as: i) Detection - rapidly detect abnormalities in process operation ,ii) Diagnosis - look for the root cause of abnormal behavior ,iii) Fault criticality assessment - assess potential severity of the fault and iv) Decision - stop the process and fix the problem or accommodate the fault and proceed.

In this paper, we focus on applications of statistics and data science to “Sustainable Manufacturing”. In the US, a president team on sustainable manufacturing, emphasized the need to maximize every atom of matter and joule of energy using technologies and systems that enable optimal raw material, energy, and resource utilization. The