Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 57 (2016) 276 – 280

www.elsevier.com/locate/procedia

49th CIRP Conference on Manufacturing Systems (CIRP-CMS 2016)

Usage of analytical services in industry today and tomorrow

Colangelo, Eduardo[a*], Bauernhansl, Thomas[a,b]

[a]*Institute of Industrial Manufacturing and Management IFF, University of Stuttgart, Nobelst. 12, Stuttgart, 70569, Germany*
[b]*Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelst. 12, Stuttgart, 70569, Germany*

* Corresponding author. Tel.: 0049 711/ 970 1912, fax: 0049 711/ 970 1927, E-mail address: eduardo.colangelo@ipa.fraunhofer.de

Abstract

Data is everywhere. Both, machines and men leave a digital shadow behind, which, for some means the success or failure of their business. Enterprises strive to make the most of this scattered, diverse and ever growing data, in order to obtain information they can apply to the decision-making processes. But, apart from the known and researched technical issues of volume, variety and velocity; more essential issues have to be addressed. Namely, how does an enterprise find the analytical model it needs to obtain the information it desires? From simple regression analyses to artificial intelligence, the variety in which data can be analyzed is immense. Involving specialist and consultants is time-consuming, needs effort and is usually too expensive, especially for SMEs.

This paper discusses the current options in the usage of analytics by enterprises as well as the existing challenges and elaborates recommendations for the future. Special focus is put on customer-oriented analytics by means of analytical services. In these, the building blocks of analytics are modularized in three layers: Data Interpretation & Cleansing Layer, Data Processing Layer, and Data Visualization Layer. This modularization allows building analytics in a standardized manner. Such services aim at reducing the gap between the holders of expert knowledge and the users of analytics. This is achieved by placing the attention on obtaining the desired information (choosing from a portfolio of analytics) instead of solving fundamental challenges, already addressed by the respective modules.

Peer-review under responsibility of the scientific committee of the 49th CIRP Conference on Manufacturing Systems

Keywords:

1. Introduction

The presence of data in everyday aspects of current societies is a well-known and accepted fact. It is estimated that by 2020 there will be 28,1 billion via internet connected devices [1], generating more than 44 zettabytes (IDC, 2014) of data. With a projected market growth up to 48.6 billion USD in 2019 [2], the importance of the sector is obvious.

Analytics provide the means to interpret this data. The term “Analytics” is used to describe statistical and mathematical data analysis that clusters, segments, scores and predicts which scenarios are most likely to happen [3]. Analytics are implemented through the employment of models, which allow a simplified representation of the phenomenon to be analyzed.

Examples of analytics are the influenza epidemic spread model created by Google in order to predict the development of this disease (with mixed results), or the developments of John Deere to increase the efficiency in crops production by interconnecting equipment, owners, operators, dealers and agricultural consultants.

The increasing complexity in business scenarios creates a fertile field for the application of analytics in business decisions. Thus, complex systems are denominated by the influence of four dimensions: variety, heterogeneity, dynamic and opacity [4]. The equivalence between these dimensions and the challenges of analytics, data volume, data variety and data velocity (also called the 3 Vs of Big Data) is clear.

On top of that, the implementation of the so-called Industrie 4.0 (Integrated Industry) foments a total data availability and connectivity, both horizontal (across the supply chain and product life cycle) and vertical (from the shop floor to the strategic level).

Enterprises should, therefore, find themselves in an ideal context to use analytics to profit from the data available.

Peer-review under responsibility of the scientific committee of the 49th CIRP Conference on Manufacturing Systems
doi:10.1016/j.procir.2016.11.048

Nevertheless, in practice it is not so simple for the users, especially SMEs, to implement analytical solutions. The reasons for these problems as well a solution approach will be explained in detail in this paper.

2. Variety turned into confusion

"Essentially, all models are wrong, but some are useful". This famous phrase by George E. P. Box expresses not only that models are incomplete per definition, but also makes clear that there is an intrinsic difficulty in finding the right solution to the problem at hand.

According to studies, although 75% of enterprises in Germany intend to make decisions based on Data Analytics, they don't possess an adequate strategy [5]. Amongst the main causes for this problem we find:

1. Implementation costs and efforts
2. The benefit is not apparent
3. Analytics can only be operated by experts
4. Poor price/performance ratio [6]

Indeed, a lot has been said and done about Data Analytics. Terms like Big Data, Small Data, Data Mining, Neuronal Networks, Machine Learning, etc.; find their way in the media and specialized literature. The abundance of concepts, together with the fact that more than one approach is valid for a particular problem, create a situation in which the benefit of analytics is not visible, as the selection turns extremely difficult. Moreover, new analytical models are created every day, mostly as part of research initiatives, as a further development of analytical technologies and not as the solution of specific problems

Specialized knowledge is required since enterprises need to find their way through this overwhelming offer of solutions. This causes not only an increase in costs, but also time required until the solution is implemented.

On top of this, finding the adequate approach in this kind of projects is an iterative process, in which variables are analyzed and selected. Depending on their characteristics and influence on the objective function, models are created, tested and the most convincing is selected [7]. Such approach is difficult for enterprises, who conceive the extended project times and iterations as a deterioration of the price/performance ratio.

The effect is particularly noticeable by SMEs who, because of their characteristics, do not possess the financial means to perform long and costly projects.

Lastly, the number of vendors from analytics has grown exponentially (as one might expect of an attractive market). In Germany alone, more than 120 vendors were identified in the Big Data domain [8]. Without a differentiated offer or selection criteria (in contrast, for example, to ERP or MES solutions) choosing between vendors is, in itself, an additional challenge.

3. First step: Classification

The first step towards understanding a problem, is being able to describe it. In this case, the problem has to do with the lack of transparency regarding the analytical solutions available.

A first approach towards a classification can be made, following the categorization proposed by Gartner, as described in figure 1.

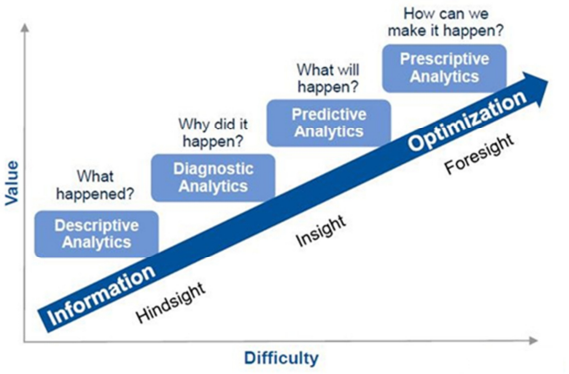

Fig. 1. Classification of analytical solution [9]

This classification possesses the advantage of not only being synthetic, but also goal-oriented. Users are, therefore, capable of matching their expectations with what the analytical model should deliver.

Nevertheless, this view presents three big problems:

1. Analytics are presented as increasingly difficult when the degree of prediction increases. This, though true in many cases, generates the false believe that descriptive or diagnostic analytics should be easier to create and implement, which is not the case. Some solutions regarded as complicated, like neuronal networks or machine learning, are known to find their way in the lower level of analytics.

2. The assumption that the way towards "optimization" is achieved only through higher degrees of prediction is also false. ISO defines quality as "ability to satisfy stated or implied needs" and this is also true for analytics. Many enterprises may find the solution to their problems just by applying diagnostic tools. It is the case, for example, by many enterprises, who require a causality analysis regarding their poor performance and quality.

3. The clear separation between analytical solutions doesn't correspond to reality (although the representation does not intend to create a fixed dissociation). Many models profit from the result of others, creating a "chain" of related results. One of the most common cases is the use of clustering solutions to achieve a first classification which reduces the