Managing Innovation and Diversity in Knowledge Society Through Turbulent Time
25–27 May 2016
Timisoara • Romania

make learn
TIIM

Management, Knowledge and Learning
Joint International Conference 2016
Technology, Innovation and Industrial Management

MODEL OF HANDLING BIG DATA AND KNOWLEDGE MANAGEMENT IN AUTOMOTIVE INDUSTRY

Horatiu Constantin Palade
Lucian Blaga University of Sibiu, Romania
extern at AUDI AG
Horatiu.Palade@ulbsibiu.ro

Sergiu Stefan Nicolaescu
Lucian Blaga University of Sibiu, Romania
Continental Automotive Systems S.R.L., Sibiu, Romania
Sergiu.Nicolaescu@ulbsibiu.ro

Claudiu Vasile Kifor
Lucian Blaga University of Sibiu, Romania
Claudiu.Kifor@ulbsibiu.ro

Abstract:
The amount of knowledge in the entire world and inside organizations already reached tremendous dimensions and is constantly increasing. In order to handle it, Big Data approaches and algorithms needs to be introduced. Without a strong management of knowledge and deployment of Big Data, the business cannot evolve in the current competitive environment and will slowly vanish. The paper is presenting the importance of Knowledge Management and Big Data inside automotive industry and is proposing a model of integration of this two emerging concepts. The vast amount of knowledge elements (data, information, knowledge, wisdom), referred as “generic awareness”, are gradually increasing and proving classical methods - like structuring - to become obsolete. As an original solution, the paper analyses the mapping of potential Big Data stages for automotive industry onto the DKIW pyramid. Based on the bibliographic research and observational study inside an automotive company, Big Data approaches are investigated and structured in order to find new ways that can prove valuable for the customer, business and innovation.

Keywords: Big Data, knowledge management, automotive industry, DKIW, Generic awareness

1. INTRODUCTION

The hallmarks of Big Data are therefore often referred to as Volume, Variety and Velocity (The 3 V's), recently emphasized as the 5Vs (Volume, Velocity, Variety, Value and Veracity), put into perspective, the emergence of BD. Big Data has been compared, from a scientific perspective to the invention of the microscope (Higginbotham S., 2011).

Organizations are facing a great risk for their business; the risk is the unknown impact that BD has over the intangible assets like intellectual capital and intellectual property. Automotive industry has entered into an era of innovation, without it, the products are not competitive and business not profitable; main element of innovation are intellectual capital and an efficient management of it.

There are estimations that the automotive industry will incline to be the second-largest generator of data (McKinsey, 2011), seconded only by the IT Industry (with Internet-based companies such as Google, Facebook, and LinkedIn). This assessment is not surprising at all, considering that some plug-in hybrid vehicles (the Ford "Energi") generate 25 GB of data in only one hour (Gardner D., 2013). The main objective of the research is to analyze the impact that Big Data has on the future activities on R&D centers of automotive industry and to emphasize methods of handling BD.

The article is structured in three parts:

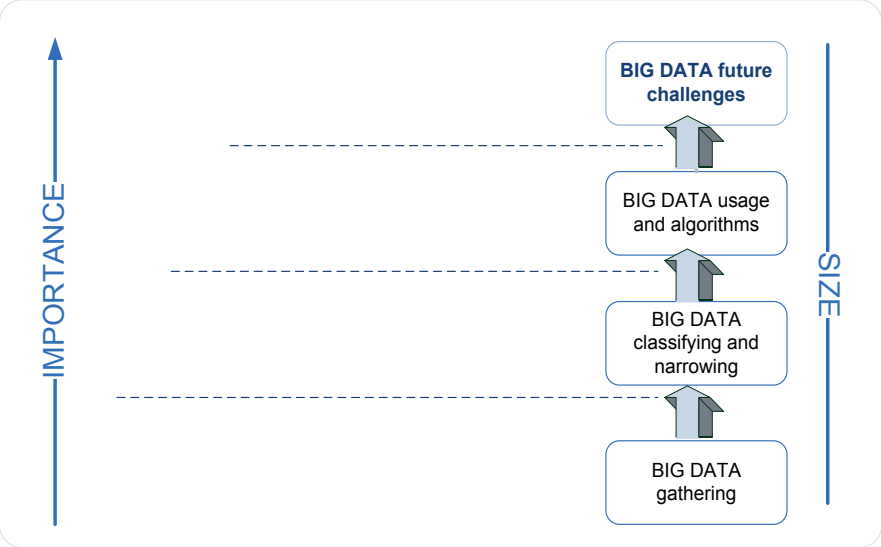
- *Big Data and Knowledge Management* presents the relation of Big Data and the DIKW pyramid so that a perspective vision of potential BD stages for automotive industry is established;
- *Big Data research - interest and results* shows our findings on the growing interest in this new topic, from academic point of view to the applied industry;
- *An approach on handling Big Data in automotive industry* starts by presenting an overview of well-known Big Data solutions and displays afterwards a model of systemizing Big Data into four, industry specific areas.

2. BIG DATA AND KNOWLEDGE MANAGEMENT

Knowledge management is also a big concern inside automotive industry, main connections between knowledge and big data can be visualized in Picture 1 and are explained in this chapter. In order to process the exponential rise of the "*Generic awareness"* (generic naming used to describe Wisdom, Knowledge, Information and Data and the different levels of the DIKW pyramid), BD programs are becoming a mandatory tool for all organizations. The pyramid is providing an epistemological outlook for the structure of knowledge and has become statutory to the field of organization knowledge (Bernstein, 2011).

The hierarchy of knowledge has been first published in 1998 by American organizational theorist Russell Ackoff (Ackoff, 1989), who explained the concept but did not represent the hierarchy graphic, even if the representation as a pyramid is credited to him by subsequent authors as the "original articulation" (Rowley J., 2007).

Picture 1: Big Data and DKIW - integrated solution



Data can exist in any form, it simply exists and has no significance beyond its existence; Big data gathering techniques are the first step into encompassing the raw material into a digital form, it is a process of “datafication”(ex: recording all the project based conversation from all the meeting room into a cloud support).

Information is data that is combined with other information or data and has a meaning, that can be useful or not. In order to structure a large amount of information the Big Data classifying and refining techniques are used (ex: now, this conversations get to be classified by date, speaker or topic, with some initial algorithms being applied, so access can be made in an facile manner).

Knowledge is the compilation of information in a manner that it can be useful and by this being a “deterministic process” (Bellinger, Durval & Mills, 2004); on the other hand, information lacks this quality. At this point, using BD algorithms, an interrogation on the conversation database on how to solve a specific client problem will already have a response and become useful with a list of solutions, after second level applied algorithms.

“Wisdom is an extrapolative and non-deterministic, non-probabilistic process” (Bellinger et al.,2004) and dissimilar to the previous levels, it poses questions without a clear, out-of-the-box answer, or some that even lack humanly-known answers until this date. Wisdom relies also on the mechanism of discerning between right and wrong, actionable or non-actionable. The next logic step in evolution, after autonomous BD, is proposing the *Ethics* concept into the “thinking mind”, what we can call Big Data Future Challenges (ex: finally at this step, it would be possible to generate a tailored solution to content all the stakeholders alike, clients, employees, suppliers, civil society).

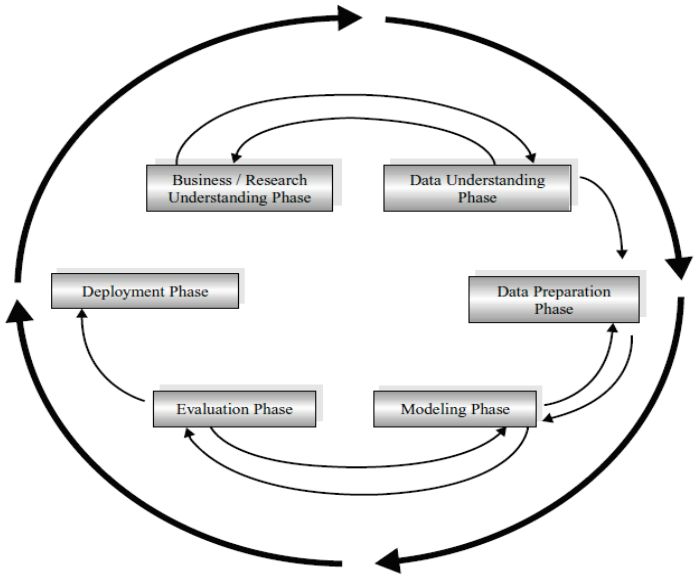
As the importance factor decreases with every level you descend the pyramid, so does the increasing size of the *Generic awareness* at the bottom reaching its full potential and size of Big Data. The *Generic awareness* can prove to be a possible metric for naming and evaluation the quality or *usefulness* of all the four previous described concepts. In this way, Big Data can resuscitate Knowledge Management initiatives with a well-balanced *fusion*, needed after the KM fall of interest due to setbacks in practical implementation (Palade, Nicolaescu & Kifor, 2015).

Big Data gathering refers, but not exhaustively, to data mining algorithms - a well-established field, coupled recently with the BD term. The Big Data terminology first appeared in 1998 in a Silicon Graphics (SGI) slide deck by John Mashey (Fan & Bifet, 2013), but started its fulminant rise only later on, after 2010, as it is presented in the next chapter.

Big Data pure algorithms (*classifying and narrowing*) that can be found at the next step are mainly referring to clustering, decision trees and networks, all derived from the field of Statistics/ Probability and expressed by Bayesian networks, naive Bayes or Markov chains.

At the IEEE 2006- International Conference on Data Mining, the top 10 most important data mining algorithms were determined, without the previous separation, all based on the reputed experts opinions, community survey and citation counts: C4.5, k-means, SVM (support vector machine), Apriori, EM (expectation maximization), PageRank, AdaBoost, kNN (k-nearest neighbors), Naïve Bayes, and CART. The identified algorithms range from classification, clustering, regression, association analysis to network analysis (Chen, Hsinchun, Chiang & Storey, 2012).

Picture 2: Cross-Industry Standard Process for Data Mining



Source: Larose, 2014

All of the above should not be applied without a clear methodology and within an integrated workflow. CRISP-DM, Cross Industry Standard Process for Data Mining**,** developed in 1996 by a group of analysts representing DaimlerChrysler, SPSS, and NCR provides a nonproprietary and open standard process for integrating data mining into the general business practice of an organization (Larose, 2014), but adaptation are in order to achieve the Knowledge Management mapping. Picture 2 details this standard and can be considered the starting point for a quality integrated approach. The focus should be on applying iterative concepts, same as in neural networks or common organization "lesson learned" scenarios, in order to create a living standard that can be addressed to a continuous process, always *expanding* and *fit* to the newest challenges.

Quality activities are defining the strength level of a supplier or customer in the automotive industry. A low quality will surely push a company to lose the trust of the stakeholders and face unforeseen failures. When the automotive organizations are starting a Big Data project approach they will face all the risks that a new field, commonly IT proprietary, encompasses. A *quality wall* and standards of implementation must be adopted from the beginning; ones that respect the desiderates to be easily deployed and rapidly adapted.

Big Data usage and algorithms refers to the recent rising number of tools that promise: *automatic, user –free, intuitive approaches,* with the comment that these *could never be separated* (or at least for now*)* from the human element. As a parallel we can note the words of Georges Grinstein of the University of Massachusetts that stated the following (Ankerst, 2003): "Imagine a black box capable of answering any question it is asked [...] Will this eliminate our need for human participation..? Quite the opposite. The fundamental problem still comes down to a human interface issue. How do I phrase the question correctly? [...] How do I get the results in reasonable time and in a form that I can understand? *All the questions connect the discovery process* to me, for my human consumption."

The future challenges reside to the ones of moving Big Data from the domain of probability and statistics to the one of certainty. The right steps are made in this direction, algorithms already predicting the user intention with a degree of over 99% percent (with variations in the retrieval time,

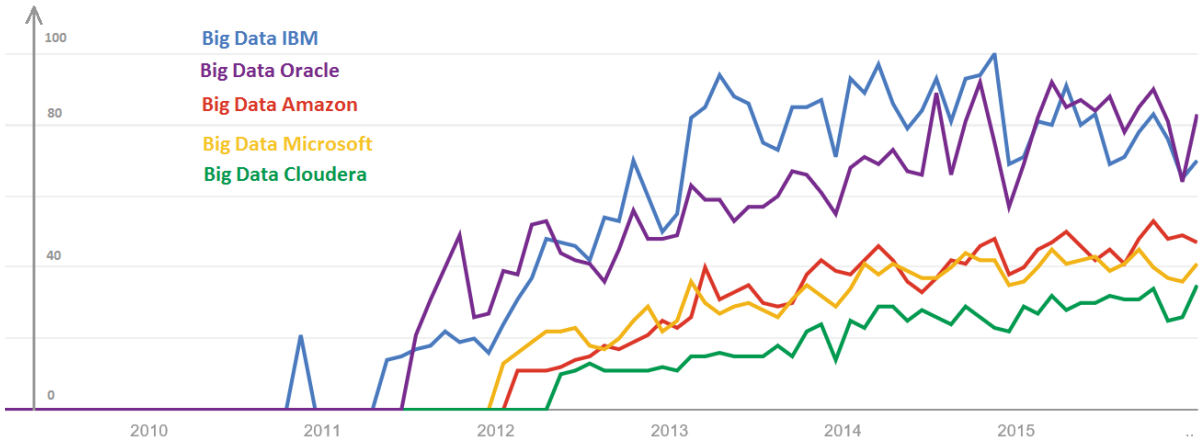
and additional hit mechanism). But this final step, in our opinion, will not be a leap of one's revolutionary algorithm but will truly reside on the raw quantity; as Alon Halevy and Peter Norvig stated in 2009 in their widespread article "The Unreasonable Effectiveness of Data": "follow the data. [..] because with very large data sources, the data holds a lot of detail." (Halevy, Norvig & Pereira, 2009). Following their research on semantic web and machine learning we can expand and extrapolate the pivotal point: a probable statement in the near future refers to the fact that *Knowledge,* as a whole concept derived from Information and Data after effort or algorithms, could become obsolete; *generic awareness* (or other naming) will replace the quality factor with the quantitative sufficient one.

3. BIG DATA RESEARCH – INTEREST AND RESULTS

The chapter presents the popularity of well-known Big Data platforms, based on Google Trends statistic engine and concludes with figures on the academic interest, articles and conferences, ending with a positive survey on applying Big Data programs in organizations.

The algorithm behind Google Trends measures interest in a search topic (Big Data) by counting many different search queries that may relate to the same topic (The Big Data, IBM Big Data, Big Data Analytics) and generated the graphs using Big Data! Real -time analysis.

Picture 2: Big Data Solution trends



Source: Google Trends

Picture 3 above presents the acceptance of some well-known Big Data solutions, relatively measured to one another based on the search term popularity. Statistics of Google Trends starts from year 2004, but the analysis showed that Big Data solution research started much later between year 2010 and 2011. It can be observed a high increase in the interest of most Big Data approaches between 2012 and 2013 and some high peaks in 2013 and 2014. The Y axe represents the number of searches that have been done for the particular term (Big Data 'Solution') relative to the total number of searches done on Google over time. The used scale is from 0 to 100 and don't represent the absolute search volume numbers, the real numbers being normalized. In case Google does not have enough data on a term, like for Cloudera until mid- 2012, it plots the value zero.

Knowledge Management had the same rapid rise in the 90', but some of the KM initiatives were not as initially promised, having failure rate of 50%, but this number could be even higher if failure is defined more broadly so as to include all projects that did not live up to their expectations (Akhavan et al., 2005). We consider that "Big Data" will not suffer the same and that it may even reinvigorate Knowledge Management initiatives with the introduction of these user-friendly, process-ready solutions that can monetize the return rate and prove their economic viability.

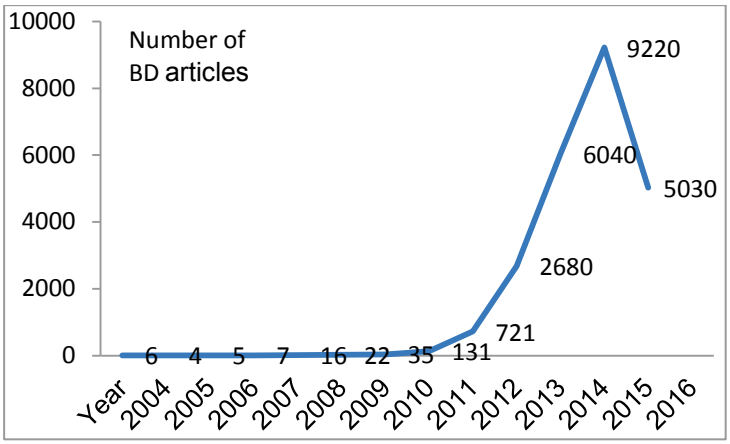
Some researches considered in 2013 Big Data to be just a redesigned old concept for marketing purpose, "buzzword to stimulate the IT providers' sales" (Buhl, Röglinger, Moser & Heidemann, 2013), but conclude that some industries (guided by the technological development and future increase of the

quality of data), like the automotive one, will be able to benefit from the usage of BD for better fulfilling customer's needs and in respect of their privacy concerns. If the IT arena is dominated by USA giants, the German automotive manufacturers set the pace of the global markets and realize sales records; here BD could prove to be a be a forced chance for these companies, where stricter privacy rules create a harsh environment (Buhl *et al.*, 2013), to innovate and develop business models that create value for the company and surpass all limitations.

Scholars' research interested in Big Data has been growing quickly: from 7 published articles/editorials in 2012 (with "Big Data" in their title), the numbers rose to 59 in 2013, doubling to 112 in 2014 and over 300 in 2015 and with already 82 articles in 2016; based on a query in Science Direct journal list database (Sciencedirect.com, 2016).

The rate of growth remains the same with the increase number of articles per database; when running queries on Thomson Reuters Web of Science or Google Scholar. This last research platform, indexing more than any others, return the first Big Data articles already from 2013 (6 titles), remaining under 10 until 2008; a modest increase to 35 in 2009 and a surge in 2010 to 131 results. The research method remains the same as the previous one, with the exact BD syntax in the title, separating yearly results and excluding auto citations and automatic mentions. While in 2011 we have a number of around 700 titles and 2680 in 2012, each year signals a new doubling to the point of reaching about 10.000 results in 2014 (Scholar.google.com, 2016).

Picture 4: Number of Article on Big Data by Google Scholar



Picture 4 presents this evolution, with the remark that 2015 and 2016 articles are not fully indexed, new articles are continuously added (for the last five years, newest years being the most subject to change) and we cannot state that a decline tendency is to be observed.

We also detected a number of increase conferences and research articles in 2015 from the one's in 2014, both new entries and continuing older editions: The First IEEE International Conference on Multimedia Big Data (BigMM), Beijing China; The inaugural INNS Big Data conference 2015 USA, San Francisco; The Second International Conference on Data Mining, Internet Computing, and Big Data, 2015 IEEE International Conference on Big Data (IEEE BigData 2015); The 8th International Conference on Knowledge Science, Engineering and Management (KSEM 2015); 4th International Congress on Big Data May 15 - May 16, 2015 (Taipei 101, Taiwan), and the list can continue with the upcoming events in 2016: IEEE International Conference on Big Data Analysis (ICBDA 2016, Hangzhou-China) or IEEE BigDataService 2016, April Oxford- England.

4. AN APPROACH ON HANDLING BIG DATA IN AUTOMOTIVE INDUSTRY

The technology behind BD is qualitatively different from previous data analytics technology so far as it can operate on vastly larger amounts of data, can tap data from any number of sources, public and private (including social media, for example), and can do so in real time, as the data is produced or recorded.

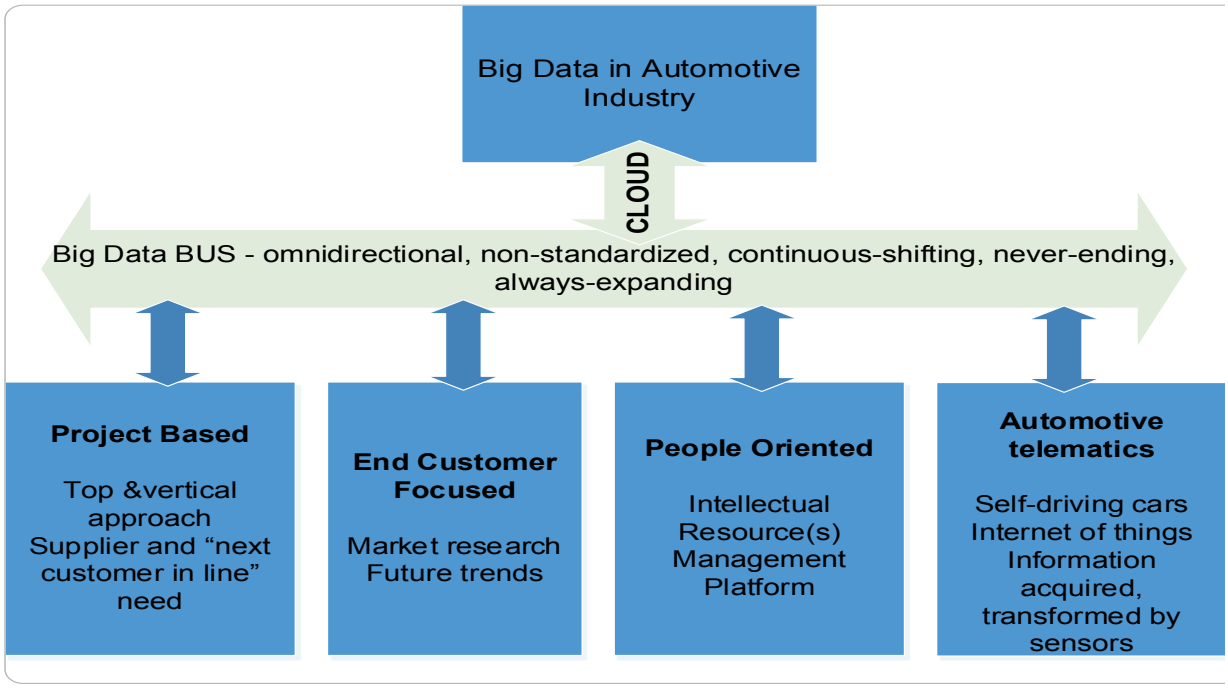
The amount of "*Generic awareness*" in the entire world, and inside organizations is increasing, in order to handle it, Big Data solutions, free or paid service platforms were and are developed. The table below presents an overview on the Big Data services with highlights on the used technology and their applicability in the industry field (including automotive).

Table 1: Big Data top used services

Name	Type	Description
IBM	Paid service	The solution offers an approach for automotive industry, which is promoted through best practices of well-known OEMs. The focus of the approach is to reduce costs and increase competitive advantage through the exploitation of data.
Microsoft	Paid service	HDInsight is a comprehensive set of managed Apache big data projects: batch, sql, nosql, script, streaming and in-memory. It is common used in businesses, being integrated with other Microsoft products; Excel, a Big Data analytic tool for all employees with basic IT knowledge.
Oracle	Paid service	Oracle Big Data Appliance has one branch designed for industry and offers a unified architecture of integration, management, analytics, and applications that fulfills the business requirements. It's suggested to be used in various domains like: communication, financial services, health science, high technology, public sector and retail. The platform uses Hadoop and NoSQL systems and supports flexible customer needs.
Amazon	Paid service	Amazon Web Services is an easy-to-use cloud computing platform that performs big data analytics and meets the challenges of the three big data Vs - volume, variety, and velocity - in digital information. Public data is offered through AWS by many organizations around the globe for the purpose of research. This is increasing the chances of scientific discovery and builds efficiencies across many industries..
Celonis	Paid service	Celonis is a recent German initiative which offers a complete enterprise-ready Process Mining solution; it is designed for business process mining for different industry (including automotive), with the option of personalized solutions
Cloudera	Open Source / Paid service	An Apache Hadoop solution designed by employees of Google, Yahoo, Facebook and Oracle that provides solutions for big data in both open source and commercial approach. Cloudera Enterprise is a specific integrated solution for storing analyzing all data an metadata with compliance-ready security and data governance.
Kaggle	Open Source	An approach that converted data prediction in a competition. Companies, organizations or even researchers can post their data and have it inspected by the world's best researchers. It also offers a free of charge self-service competitions for academic institutions, creating a community around Big Data.
Mojito	Open Source	The solution consists of a dongle designed to be plugged on the ODB car port that reports vehicle performance data to a cloud server. The approach integrates all vehicles which lack of internet connectivity (most of current cars).

Based on the previous presented research and the observational study on an automotive company, Big Data was structure in four types of domains that cover the whole automotive industry area: Project Based, End Customer focused projects, People oriented and Automotive Telematics and Sensors, described below in Picture 5.

Picture 5: Big Data structured in automotive industry



Project based – Big Data will enhance the results and will improve the Supply Chain and reduce the lean time of quality approaches like Six Sigma, Design for Six Sigma, Kaizen, Lean manufacturing; no more statistic results, clear data is provided.

Supply Chain is predicting which car will be in demand by using Customer Relationship management database or public data. Also it's monitoring dealer/ "end of line customer" satisfaction like Hertz (American car-rental Company with worldwide locations in 145 countries).

End customer focus - Through Big Data customer becomes part of the Control Board of project decision, the needs of the customer will influence the future cars' functionalities. Customized Driving Experience will provide personalized driving tips and suggestions for improving everyone driving style; such as a break warning or unnecessary acceleration message for inexperienced drivers to reduce consumption. Destination suggestions or alternative routes could inform drivers of probable points of interest linked with their Smartphone's recent search queries.

Looking on the *management role* to decide the future of the business, it can be observed that the decision power will gradually be harnessed directly from the Client, through BD decision making algorithm; in this way the Client can be seen as a share-manager, a future level of CRM (Customer relationship management), the Client needs will directly influence the business decision.

People oriented – Measurement of intellectual capital & knowledge management; in this way the business profitability will be increase with the growth of an intangible resource, people know-how. Based on the processed data every employee can be tracked during its development and its current performance can be properly measured. This Big Data management platform can be used to check the know-how distribution inside companies or inside departments, team; in this way interchanges can be made for an efficient resource allocation. *How much would the performance grow if the management would know exactly how to balance the teams? If an employee is well integrated inside a team?*

A big issue of the organization management is the employee inadaptability, with BD it can be measured and the management will notice if the person is suitable or not for the team. In this way decision can be made based on the suggestions that Big Data algorithms are providing. Until 2018 it is predicted that 6,400 organizations with more than 100 staff will implement big data analytics; all this from a SAS Study on more than 1200 business, conducted in 2013 (SAS.com, 2016).

Automotive telematics – Interconnected car, locally and globally alike, will be used for saving lives, time and money by avoiding accidents, traffic jams and even avoid theft; algorithms that map the

personal driving behavior of a vehicle's owner /'s, can also be used to identify car theft as an unexpected behavior. BD has its place also in the Connected Car environment, being used to deliver real time traffic and weather information to drivers, all this for improving traffic efficiency. Connected Cars will drive in "Intelligent Cities" and exchange real-time Exabyte of BD with the "Intelligent Infrastructure".

The idea that makes the foundation of automated driving and the need of big data is simple and comes from the nature. Some complex devices can make a good interpretation of the world around the car, but even with the best one there are limitations and some exceptions cannot be caught. If more than one device is used (same parallel as the natural senses) the probability to handle all specific exceptions is higher.

In the second International VDI Conference on Automated Driving it was discussed a Comprehensive Environment Model that will gather the information from all car devices and sensors in order to have redundant information and can take safe decisions. This is where the Big Data will intervene to process all existing data inside the automobile and data from the surrounding environment extracted by the car's sensors.

Big Data will prove to be of a great effectiveness in the fields of supply-chain and into pursuing customer oriented policy of reducing warranty costs (detecting and remediating defects before they become a safety threat), integrate with cities government agencies ("Intelligent cities") and centralized traffic authorities to resolve road network issues, improve traffic flow and cut emergency response time. By doing so, this will ensure new opportunities to promote accident prevention, improve vehicles safety and enrich the driving experience by having the same level of connectivity in a vehicle whether driver or passenger as one would enjoy in its home.

BD allows organizations to combine their data (e.g., on customer behavior) with data from other sources, public and private, to produce useful insights. BD has contribution to public and private sector and has proven its utility in numerous applications like public health, fraud detection, cost reduction, service improvement and transportation efficiency (Schroeder, 2014). Automakers will continue to invest in Big Data and Big Data Analytics to pursue the goal of the Connected Car, a binding development in this competitive market and fast changing environment.

5. CONCLUSIONS

Large quantities of raw data, even so unstructured could provide the evolutionary leap to the classical DIKW pyramid; where a 99.999% percent on large population will begin to feel like 100% for almost any, normal user, and the rare, unattended 0.001% percent rare events will begin to pile up, as data grows and group up, adding another decimal 9 after the comma, another precision point.

There should be no concerns, we are only on the brink of autonomous data transforming (with the previous detailed supervised use guidance) into information; knowledge "needs to have patience" and wait for its turn.

Here Moore Law comes into place and could further indicate that if we observe the rate in which Big Data is created by the products of automotive industry (like Connected cars sensors) we can safely assume that the rate of growth will surpass the law in electronics - regarding the doubling every two years, of transistors per square inch of integrated circuit meaning/ translated into processing speed (Keyes, 2006) - in the near future for all automobile industry.

REFERENCE LIST

1. Ackoff, R. L. From data to wisdom. Journal of Applied Systems Analysis 15: 3-9. (1989)
2. Akhavan, P., Jafari, M., and Fathian, M,'Exploring Failure-Factors of Implementing Knowledge Management Systems in Organizations, Journal of Knowledge Management Practice, [electronic], vol. 6, May, pp.1-8, (2005)
3. Bellinger, Gene, Durval Castro, and Anthony Mills. "Data, information, knowledge, and wisdom." (2004)
4. Bernstein, J. (2011). The Data-Information-Knowledge-Wisdom Hierarchy and its Antithesis. NASKO, 2(1). http://dx.doi.org/10.7152/nasko.v2i1.12806

5. Buhl, Hans Ulrich; Röglinger, Maximilian; Moser, Florian; and Heidemann, Julia "Big Data - A Fashionable Topic with(out) Sustainable Relevance for Research and Practice?," Business & Information Systems Engineering: Vol. 5: Iss. 2, 65-69. (2013)
6. Chen, Hsinchun, Roger HL Chiang, and Veda C. Storey. "Business Intelligence and Analytics: From Big Data to Big Impact." *MIS quarterly* 36.4 (2012): 1165-1188.
7. E-skills UK, Big Data Analytics. Adoption and Employment Trends, 2012-2017; retrieved from http://www.sas.com/offices/europe/uk/downloads/bigdata/eskills/eskills.pdf in 30.01.16.
8. Fan, W., & Bifet, A. (2013). Mining big data: current status, and forecast to the future. *ACM sIGKDD Explorations Newsletter*, *14*(2), 1-5.
9. Gardner, Dana. "Ford scours for more big data to bolster quality, improve manufacturing, streamline processes." Briefings Direct on ZDNet.com. Jan. 9, (2013)
10. Google.com,. "Google Trends - Web Search Interest - Worldwide, 2004 - Present". N.p., 2016. Web. 2 Feb. 2016.
11. Greg Schroeder (2014), "Advanced Information Technology Solutions: An Engine of Innovation"
12. Halevy, A., Norvig, P., & Pereira, F. (2009). The Unreasonable Effectiveness of Data. IEEE Intell. Syst., 24(2), 8-12. http://dx.doi.org/10.1109/mis.2009.36
13. Keyes, R.W., "The Impact of Moore's Law," Solid-State Circuits Society Newsletter, IEEE , vol.11, no.5, pp.25,27, doi: 10.1109/N-SSC.2006.4785857, Sept. (2006)
14. Larose, D. T. (2014). *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons.
15. McKinsey Global Institute. "Big data: The next frontier for innovation, competition, and productivity." June, (2011)
16. Mihael Ankerst, The perfect data mining tool: Interactive or automated? Report on the SIGKDD-2002 Panel, *SIGKDD Explorations*, Vol. 5, No. 1, July 2003.
17. Palade, H. C., Nicolaescu, S. □., & Kifor, C. V. (2015). The Impact of Big Data and Knowledge Management on R&D Projects from Automotive Industry. In *Knowledge Science, Engineering and Management* (pp. 183-189). Springer International Publishing.
18. Rowley, Jennifer. "The wisdom hierarchy: representations of the DIKW hierarchy". Journal of Information Science 33 (2007)
19. S. Higginbotham, "For science, big data is the microscope of the 21st century" (2011)
20. SAS.com (2016), Big Data Analytics Adoption and Employment Trends, 2012-2017. Retrieved 10 February 2016, from http://www.sas.com/offices/europe/uk /downloads/bigdata/eskills/eskills.pdf
21. Scholar.google.com,. (2016). allintitle: "big data" - Google Scholar. Retrieved 12 February 2016
22. Sciencedirect.com, (2016). 594 Search Results - TITLE ("big data") - ScienceDirect. Retrieved 12 February 2016