Model for Big Data Analytics in Supply Chain Management

Nenad Stefanovic*, Danijela Milosevic*
* Faculty of Technical Sciences Cacak, University of Kragujevac, Serbia
nenad@ftn.edu.rs; danijela.milosevic@ftn.kg.ac.rs

Abstract— **Supply chains are complex systems with silos of information that is very difficult to integrate and analyze. The best way to effectively analyze these composite systems is the use of business intelligence (BI). However, traditional BI systems face many challenges that include processing of vast data volumes, demand for real-time analytics, enhanced decision making, insight discovery and optimization of supply chain processes. Big Data initiatives promise to answer these challenges by incorporating various methods, tools and services for more agile and flexibly analytics and decision making. Nevertheless, potential value of big data in supply chain management (SCM) has not yet been fully realized and requires establishing new BI infrastructures, architectures, models and tools. The first part of the paper discusses challenges and new trends in supply chain BI and provides background research of big data initiatives related to SCM. In this paper, the methodology and the unified model for supply chain big data analytics which comprises the whole BI lifecycle is presented. Architecture of the model is scalable and layered in such a way to provide necessary agility and adaptivity. The proposed BI model encompasses supply chain process model, data and analytical models, as well as insights delivery. It enables creation of the next-generation cloud-based big data systems that can create strategic value and improve performance of supply chains. Finally, example of supply chain big data solution that illustrates applicability and effectiveness of the model is presented.**

I. Introduction

As the globalized business environment is forcing supply chain networks to adapt to new business models, collaboration, integration and information sharing are becoming even more critical for the ultimate success. Supply chains enterprise systems are experiencing a major structural shift as more organizations rely on a community of partners to perform complex supply chain processes. While supply chains are growing increasingly complex, from linear arrangements to interconnected, multi-echelon, collaborative networks of companies, there is much more information that needs to be stored, processed and analyzed than there was just a few years ago.

Supply chain business intelligence is a collection of activities to understand business situations by performing various types of analysis on the organization data as well as on external data from supply chain partners and other data sources (devices, sensors, social networks, etc.) to help make strategic, tactical, and operational business decisions and take necessary actions for improving supply chain performance. This includes gathering, analyzing, understanding, and managing high volumes of variety data about operation performance, customer and supplier activities, financial performance, production, competition, regulatory compliance, quality controls, device data and Internet.

Over the past few decades, the way in which companies need to collate, analyze, report and share their data has changed dramatically. Organizations need to be more adaptive, have increased access to information for decision-making, and effectively deal with a rapidly growing volume of data.

Today's business environment demands fast supply chain decisions and reduced time from raw data to insights and actions. Typically, supply chains are capturing enormous data volumes - including vast amounts of unstructured data such as files, images, videos, blogs, clickstreams and geo-spatial data, as well as data coming from various sensors, devices, and social networks.

Supply chain BI system proved to be very useful in extracting information and knowledge from existing enterprise information systems, but in recent years, organizations face new challenges in term of huge data volumes generated through supply chain and externally, variety (different kind of structured and unstructured data), as well as data velocity (batch processing, streaming and real-time data). Most of the existing analytical systems are incapable to cope with these new dynamics [1].

On the other hand, we have seen tremendous advancements in technology like in-memory computing, cloud computing, Internet of Things (IoT), NoSQL databases, distributed computing, machine learning, etc. Big data is a term that underpins a raft of these technologies that have been created in the drive to better analyze and derive meaning from data at a dramatically lower cost and while delivering new insights and products for organizations in the supply chain.

The key challenges for modern supply chain analytical systems include [2]:

- Data explosion – supply chains need the right tools to make sense of the overwhelming amount of data generated by a growing set of data internal and external sources.
- Growing variety of data – most of the new data is unstructured or comes in different types and forms.
- Data speed – data is being generated at high velocity which makes data processing even more challenging.
- Real-time analysis - in today's turbulent business climate the ability to make the right decisions in real-time brings real competitive advantage. Yet many supply chains do not have the infrastructure,

tools and applications to make timely and accurate decisions.

- Achieving simplified deployment and management – despite its promise, big data systems can be complex, costly and difficult to deploy and maintain. Supply chains need more flexible, scalable and cost effective infrastructure, platforms and services, such as those offered in cloud.

To succeed in a competitive marketplace, an agile supply chain requires a special big data analytical system to quickly anticipate, adapt, and react to changing business conditions. BI systems provide sustainable success in a dynamic environment by empowering business users at all levels of the supply chain and enabling them to use actionable, real-time information.

II. Background Reasearch

During the past two decades organizations have made large investments in SCM information systems in order to improve their businesses. However, these systems usually provide only transaction-based functionality and mostly maintain operational view of the business. They lack sophisticated analytical capabilities required to provide an integrated view of the supply chain. On the other hand, organizations that implemented some kind of enterprise business intelligence systems still face many challenges related to data integration, storage and processing, as well as data velocity, volume and variety. Additional issues include lack of predictive intelligence features, mobile analytics and self-service business intelligence capabilities [3].

The current approaches to BI has some fundamental challenges when confronted with the scale and characteristics of big data: types of data, enterprise data modeling, data integration, costs, master data management, metadata management, and skills [4].

The big data phenomenon, the volume, variety, and velocity of data, has impacted business intelligence and the use of information. New trends such as fast analytics and data science have emerged as part of business intelligence [5].

Sixty-four percent of supply chain executives consider big data analytics a disruptive and important technology, setting the foundation for long-term change management in their organizations [6]. Ninety-seven percent of supply chain executives report having an understanding of how big data analytics can benefit their supply chain. But, only 17 percent report having already implemented analytics in one or more supply chain functions [7].

While data science, predictive analytics, and big data have been frequently used buzzwords, rigorous academic investigations into these areas are just emerging.

Even though a hot topic, there is not many research related to big data analytics in SCM. Most of the papers deal with big data potential, possible applications and value propositions.

Wamba and Akter provide a literature review of big data analytics for supply chain management [8]. They highlighting future research directions where the deployment of big data analytics is likely to transform supply chain management practices. Waller and Facett examine possible applications of big data analytics in SCM and provide examples of research questions from these applications, as well as examples of research questions employing big data that stem from management theories [9].

Identifying specific ways that big data systems can be leveraged to improve specific supply chain business processes and to automate and enhance performance becomes crucial for ultimate business success. The information and analytics delivered would be used to improve supply planning, vendor negotiations, capacity planning, warehousing and transportation performance, productivity, shop floor performance, materials requirements planning, distribution and customer service.

Organizations can apply big data and BI in the following supply chain areas [10]:

- Plan Analytics — balancing supply chain resources with requirements.
- Source Analytics — improving inbound supply chain consolidation and optimization.
- Make Analytics — providing insight into the manufacturing process.
- Deliver Analytics — improving outbound supply efficiency and effectiveness.
- Return Analytics — managing the return of goods effectively and efficiently.

Some of the most important data-driven supply chain management challenges can be summarized as follows [11]:

- Meet rising customer expectations on supply chain management.
- Increase costs efficiency in supply chain management.
- Monitor and manage supply chain compliance & risk.
- Make supply chain traceability and sustainability a priority.
- Remain agile and flexible in volatile times and markets.

Supply chains which implemented certain big data systems have achieved the following benefits [12]:

- Improvement in customer service and demand fulfillment.
- Faster and more efficient reaction time to supply chain issues.
- An increase in supply chain efficiency.
- Integration across the supply chain.
- Optimization of inventory and asset productivity.

Despite new business requirements and technology innovations, big data methods, models and applications in SCM still need to be researched and studied.

In the subsequent sections we present the supply chain big data model, software architecture and example of big data analytical system.

III. Models and Methods

Big data analytics has to do more with ideas, question and value, than with technology. Therefore, the big data analytics methodology is a combination of sequential execution of tasks in certain phases and highly iterative execution steps in certain phases.

The big data analytics methodology is a combination of sequential execution of tasks in certain phases and highly iterative execution steps in certain phases. Because of the scale issue associated with supply chain big data system, an incremental approach is recommended, which include modifying and expanding processes gradually across several activities as opposed to designing a system all at once [13].

- Analyze and evaluate supply chain use case - frame the problem, gather sample data and perform data discovery and analysis.
- Develop business hypotheses - assemble illustrative supply chain use cases and perform fit-gap analysis.
- Build and prepare data sets - acquire data and understand data characteristics,
- Select and build the supply chain analytical models - test and validate with data, apply data visualization techniques and review results
- Build the production ready system - architect and develop the end state solution
- Measure and monitor - measure effectiveness of the big data analytics solution

Figure 1 provides a high-level view of the big data analytics methodology, and designers (i.e., architects, analysts, data modelers, etc.) are advised to iterate through the steps presented. Several cycles of design and experimentation during steps 2 through 5 should be completed. Each cycle should include additional and larger data samples and apply different analytics techniques as appropriate for data and relevant for solving the supply chain problem. The entire framework should be revised periodically after the go-live in order to maintain quality and accuracy of the supply chain analytical system.

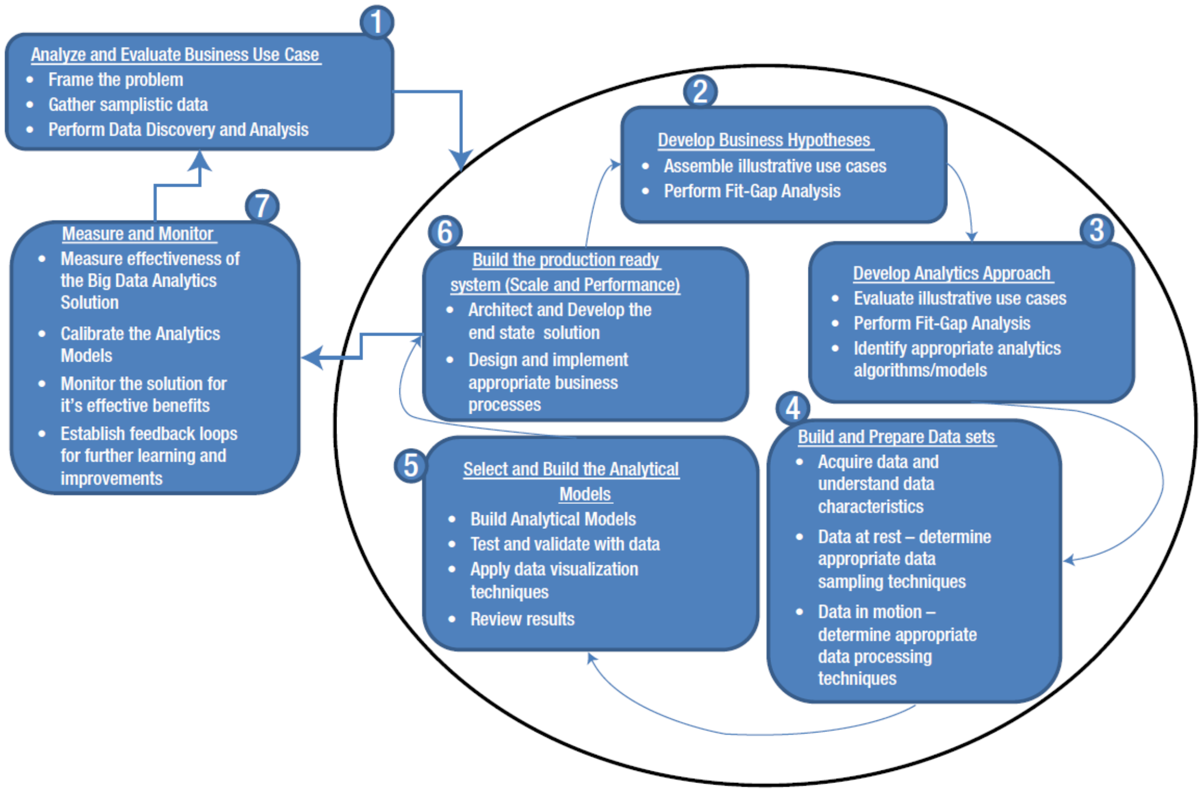

Figure 1. Big data analytics methodology

The methodology differs from others by the number of times the designer should execute the steps to solve design problems associated with processing at full scale. Knowledge gained during each pass through of the various steps should be reflected in the system design.

The real users of the analytics outcomes are business users but these users often do not understand the complex mathematical formulae, statistical analysis models, etc. Therefore, it is extremely important to equip the business users with easy-to-understand and highly intuitive tools through which they will understand what actions are to be performed.

IV. Solution and Discussion

In order to overcome main challenges of modern analytics and deficiencies of existing BI supply systems we propose a comprehensive multi-layered supply chain big data BI model that utilizes cloud-based big data services and tools for data extraction, transformation and loading (ETL), analysis, and reporting. Figure 2 shows the architecture with layers and services.

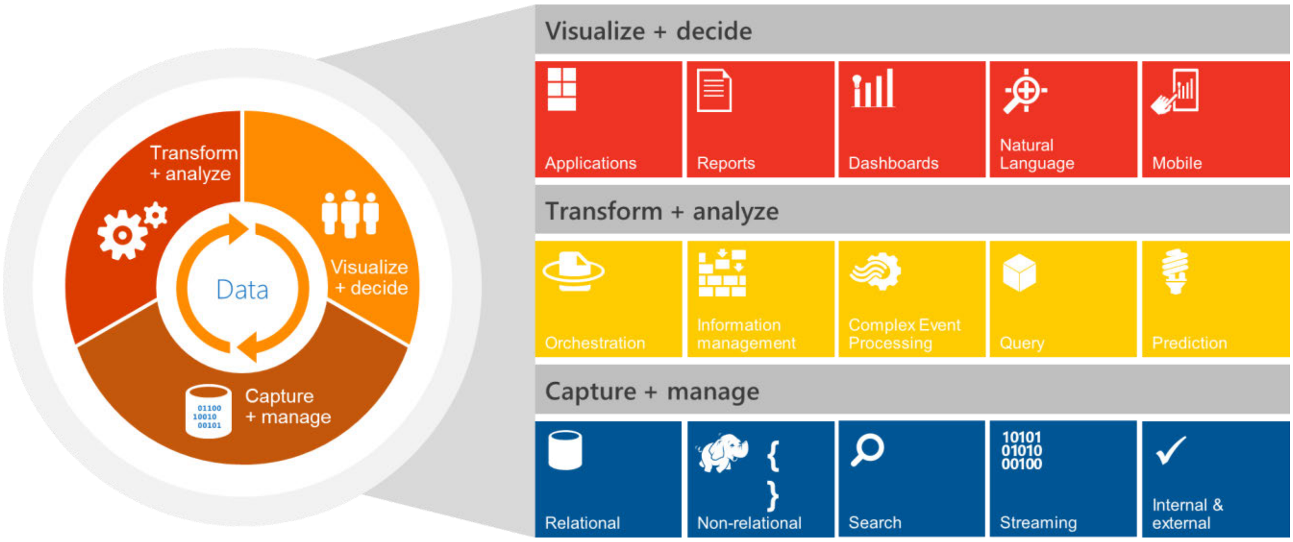

Figure 2. Supply chain analytical lifecycle model

The proposed big data model unifies processes, methodologies and tools into a single business solution. The model has been developed in such a way to seamlessly integrate within overall BI and collaboration framework [14]. It is process-centric, metrics-based and modular. It introduces the new supply network and data modeling approaches, as well as layered application architecture which enables creation of composite BI systems.

The data integration layer supports various data types (relational, unstructured, streaming, OLAP, etc.) via cloud ETL services. The data management layer is based on the Hadoop engine but with additional services which provide more flexible data models and querying. The analytical layer hosts various analytical models and schemas. These can be exploration, data mining, or performance monitoring models. The final insights layer provides insights to all users such as self-service BI, data search, collaboration and performance monitoring. The central component of this layer is specialized supply chain BI portal as the unifying component that provides integrated analytical information and services, and also fosters collaborative decision making and planning.

This approach utilizes various cloud data management services such as ETL jobs for data extraction, cleansing and import, as well as event hubs that acts as a scalable data streaming platform capable of ingesting large amounts of high-velocity data form sensors and IoT devices throughout supply chain.

Additionally, the supply chain-wide data catalog was created in the cloud. It is fully managed cloud service that enables users to discover, understand, and consume data sources. The data catalog includes a crowdsourcing model of metadata and annotations, and allows all supply chain participants to contribute their knowledge to build a various data models which can be integrated into specific applications and services.

In order to provide more flexibility, two big data stores are designed: supply chain enterprise multidimensional data warehouse and special in-memory tabular model for processing large amount of data. Combined with specific cloud analysis services, it is possible to design different analytical models. For example, stream analytical services can be used to set up real-time analytic computations on data streaming from devices, sensors, e-commerce sites, social media, information systems, infrastructure systems, etc. Another example is cloud machine learning service that enables supply chain participants to easily build, deploy, and share predictive analytics solutions (i.e. forecasting sales or inventory data) [15].

Finally, information derived from such analytical models need to be delivered to decision makers in timely and user-friendly way. For this purpose, a special web portal is used. It acts as a single point of data analysis, collaborative decision making.

In order to demonstrate our approach, we have designed a supply chain BI solution for analysis of supplier quality within the supply chain. Data from different sources (relational database, files, and web feeds) is integrated via cloud ETL job into the in-memory tabular data store.

Various analytical reports are designed using different technologies and services. All these BI artifacts (reports, charts, maps, etc.) are than integrated in the web dashboards as shown in Figure 3.

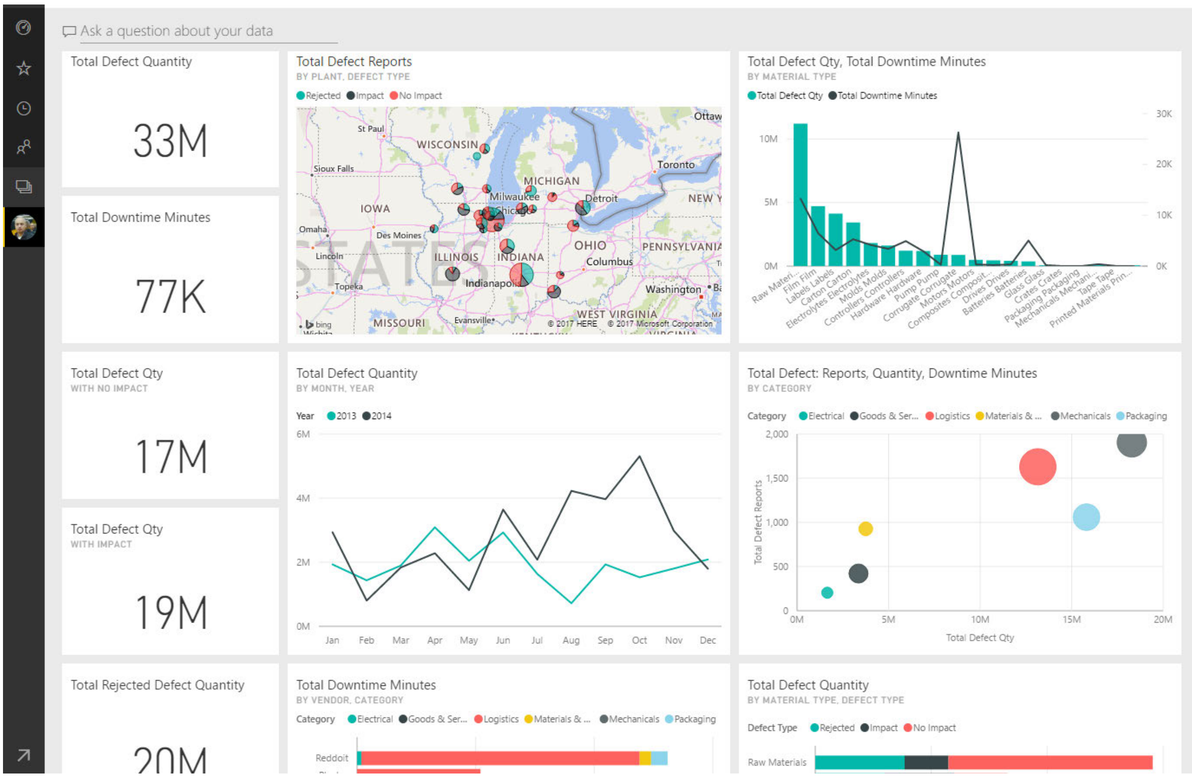

Figure 3. Supplier quality dashboard page

Supplier quality analysis is a typical supply chain task. Two primary metrics in this analysis are: total number of defects and the total downtime that these defects caused. This sample has two main objectives:

- Understand who the best and worst supply chain suppliers are, with respect to quality.
- Identify which plants do a better job finding and rejecting defects, to minimize downtime.

Each of the analytical segments on the dashboard can be further investigated by using various drill-down reports. For example, if user wants to analyze how plants dela with defective materials and the downtime, by clicking on the map segment, it opens supplier analysis dashboard (Figure 4) that can be used for deeper analysis and filtering in order to derive meaningful knwoledge about supplier management processes and take corrective actions.

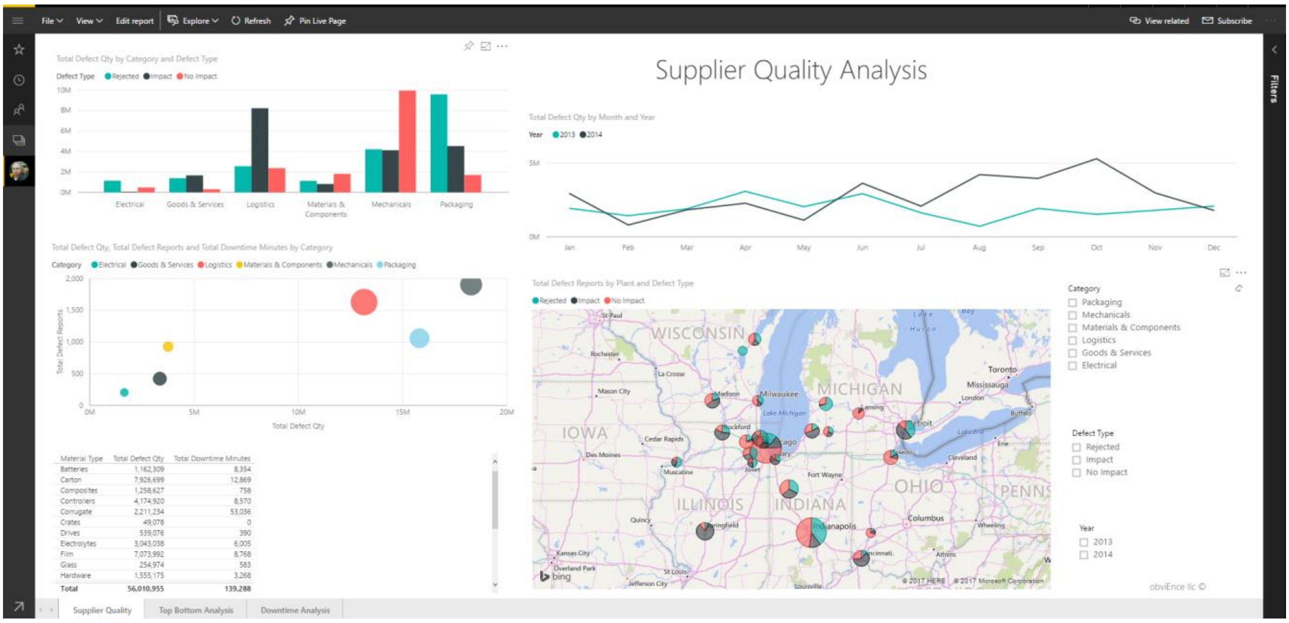

Figure 4. Supplier quality analysis

V. Conclusion

During the last several years there was an amazing progression in the amount of data produced within the supply chain information systems, but also externally. This poses many challenges related to data analysis specifically in terms of technology, infrastructure, software systems and development methods. The current business climate demands real-time analysis, faster, collaborative and more intelligent decision making.

The current approach to supply chain intelligence has some fundamental challenges when confronted with the scale and characteristics of big data. These include not only data volumes, velocity and variety, but also data veracity and value. One of the key aspects of leveraging big data is to also understand where it can be used, when it can be used, and how it can be used - how the value drivers of big data are aligned to supply chain strategic objectives.

Big data is starting to make inroads into logistics and supply chain management – large steps have certainly been taken over the past several years – but there is still a long way to go. Opportunities to create efficiency and savings through smart use of data are evident and concerted effort is being put into finding them.

Big data and advanced analytics are being integrated successfully as logistics management solutions such as: optimization tools, demand forecasting, integrated business planning, supplier collaboration and risks analytics.

The proposed layered supply chain big data model and architecture allows construction of the next-generation loosely-coupled analytical systems that combine different services, data sources, analytical models and reporting artifacts into a unified analytical. This will enable collaborative, efficient, responsive, and adaptive supply chains.

Acknowledgment

Research presented in this paper was supported by Ministry of Science and Technological Development of Republic of Serbia, Grant III-44010, Title: Intelligent Systems for Software Product Development and Business Support based on Models.

References

[1] D. Larsona and V. Chang, “A review and future direction of agile, business intelligence, analytics and data science,” *International Journal of Information Management*, vol. 36(5), pp. 700-710, October 2016.

[2] G. Wanga, A. Gunasekarana, E.W.T. Ngai, and T. Papadopolous, “Big data analytics in logistics and supply chain management: Certain investigations for research and applications,” *International Journal of Production Economics*, vol. 176, pp. 98–110, June 2016.

[3] N. Stefanovic and D. Stefanovic, “Supply Chain Business Intelligence – Technologies, Issues and Trends,“ in *Artificial Intelligence: An International Perspective*, M. Bramer Ed. Berlin/Heidelberg: Springer, pp. 217-245, 2009.

[4] D. Q. Chen, D. S. Preston, and M. Swink, “How the Use of Big Data Analytics Affects Value Creation in Supply Chain Management,” *Journal of Management Information Systems*, vol. 32(4), pp. 4-39, April 2016.

[5] T. Schoenherr and C. Speier-Pero, “Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential,” *Journal of Business Logistics*, vol. 36(1), pp. 120-132, March 2015.

[6] A. Renner, “Overcoming 5 Major Supply Chain Challenges with Big Data Analytics,” http://www.computerworld.com/article/3035144/data-center/overcoming-5-major-supply-chain-challenges-with-big-data-analytics.html.

[7] Accenture, “Big Data Analytics in Supply Chain: Hype or Here to Stay?,” https://www.accenture.com/t20160106T194441__w__/fi-en/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Digital_1/Accenture-Global-Operations-Megatrends-Study-Big-Data-Analytics-v2.pdf.

[8] S. Fosso Wamba and S. Akter, “Big Data Analytics for Supply Chain Management: A Literature Review and Research Agenda,” in *Enterprise and Organizational Modeling and Simulation*, Lecture Notes in Business Information Processing, vol. 231. Barjis J., Pergl R., Babkin E. Eds. Springer, Cham, 2015.

[9] M. A. Waller and S. E. Fawcett, “Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management,” *Journal of Business Logistics*, vol. 34(2), pp. 77-84, Jun 2013.

[10] N. Stefanovic and D. Stefanovic, “SCM 3.0 - Reshaping Supply Chains With the Next Generation Web,” in *Supply-Chain Management: Theories, Activities/Functions and Problems*, Regina M. Samson, Ed. Nova Publishers, pp. 313-333, 2011.

[11] L. Columbus, “Ten Ways Big Data Is Revolutionizing Supply Chain Management,” https://www.forbes.com/sites/louiscolumbus/2015/07/13/ten-ways-big-data-is-revolutionizing-supply-chain-management/#3edf715f69f5

[12] Advanced Fleet Management Concluting, “Top Challenges for Big Data in the Supply Chain Management Process,” http://en.advancedfleetmanagementconsulting.com/2016/07/13/1199/

[13] S. Mohanty, M. Jagadeesh, and H. Srivatsa, “Big Data Imperatives - Enterprise Big Data Warehouse, BI Implementations and Analytics,” Apress, 2013.

[14] N. Stefanovic, D. Stefanovic, and B. Radenkovic, “Integrated Supply Chain Intelligence through Collaborative Planning, Analytics and Monitoring,” in Mohhebi, S., I. Mahdavi and N. Cho (eds.) *Electronic Supply Network Coordination in Intelligent and Dynamic Environment: Modeling and Implementation*, Hershey: IGI Global, 2011, pp. 43-92.

[15] N. Stefanovic and D. Stefanovic, “Supply Chain Performance Measurement System Based on Scorecards and Web Portals,” *Computer Science and Information Systems*, vol. 8(1), pp. 167-192, 2013, doi:10.2298/CSIS090608018S.