Information Sharing for Customized Dynamic Visual Analytics: A Framework

Alireza Khakpour [1[0000-0003-1961-1269]]

[1] Department of Computer Science, Østfold University College, B R A veien 4, 1757 Halden, Norway
alireza.khakpour@hiof.no

Abstract. Supply chain activities generate massive amount of data by several actors such as, suppliers, manufacturers, warehouses, distributers, and wholesalers. Visual analytics (VA) plays a key role in knowledge discovery and insight generation from this data and helps various players to enhance their operational and strategic decision making. This is more essential for Fast moving consumer goods (FMCG) industry, given the size of the industry and its sensitivity to the diverse market uncertainties. In this paper, we present a PhD research plan that responds to the requirements of a FMCG supply chain VA system by means of a comprehensive framework. In this regard, the information flow throughout the supply chain is a significant factor for developing a reliable and efficient VA solution and a proper information flow throughout the supply chain can be enhanced with the help of the framework consisting of modules including Data Generation, Data Integration and Management , Data Analytics, Data Visualization, and Data-driven decision making. The aim of the study is to explore the development of a VA framework that acts as a guideline for supply chain players to improve their analytical capabilities.

Keywords: Data Analytics, Decision support, Information flow, Supply chain, Visual analytics.

1 Introduction and Problem Statement

Fast moving consumer goods (FMCG) sector is among the largest industries in the world. This is where everyday consumer products are produced and supplied. Meanwhile, it requires a well operating supply chain to respond to the increasing needs of consumers. Indeed, flexible production and delivery of goods is significant to fulfil the needs of consumers in an efficient manner [1]. The flexibility helps both the consumers and enterprises to have a reliable and sustainable supply chain. At the same time, the enormous amount of data created from various FMCG supply chain sectors provide the ability to ground the supply chain processes based on the knowledge gained from these data [2]. In these regards, the need for a data-driven supply chain process through real-time analysis of the available data emerges [3].

One of the tools to analyse supply chain heterogenous data and generate insights and knowledge dynamically is Visual Analytics (VA) [4]. VA is defined as the “the formation of abstract visual metaphors in combination with a human information discourse (interaction) that enables detection of the expected and discovery of the unexpected within massive, dynamically changing information spaces" [5]. FMCG supply chain activities involve the procurement process, production of goods, transportation to distributers and wholesalers, and finally, retailing and delivery of products to the hands of consumers. Within these activities, market analysis and forecasting, selling strategy selection, pricing, supplier selection, production planning, inventory planning, logistic management, warehouse planning, and business ecosystem analysis are some of the applications of VA systems in FMCG supply chain. In general, Enterprises are using VA in their decision support systems and nowadays, it is becoming an integral part of every supply chain decision support systems.

However, to integrate VA into the decision support systems in various points of a supply chain, a proper information flow within partners is required. In this process, vast amount of data need to be ingested and consolidated from several sources and to be made ready for further analysis. Although the FMCG supply chain is data rich, the exchange of information within supply chain partners are poor due to the lack of an underlying standard framework providing and particularizing various aspects of the information flow. The existence of a reliable framework can facilitate the process of establishing an operational and tailored VA solution for FMCG supply chain by explicating the requirements for future development and implementations. Some of the aspects required for a useful supply chain VA system that the desired framework should cover are as follows:

- The ability to view and consolidate data from different sources to create what-if scenarios
- Big data analysis of trends, dynamics and performance indicators across multiple dimensions
- Root cause analysis to identify reasons for faults within an existing system - enhanced with Machine Learning (ML) techniques

In order to achieve these requirements the following three activities are planned based on which Research Questions (RQ) are also defined:

1. Exploring data collection and information sharing for VA: This activity satisfies the first aspect of the framework by collecting the sufficient amount of data from industrial partners of the project.
2. Exploring big data analytics techniques for VA: This activity is toward satisfying the second and third aspects of the framework by exploring the adequate data analytic techniques for the use in supply chain activities.
3. Exploring the development of a framework for VA: This activity satisfies the overall goals of the project by providing a conceptual reference model for development of a visual analytic system in the supply chain context.

Overall, in this paper, we outline a PhD work-in-progress in developing a conceptual framework for a customized VA for FMCG supply chain industry. In what follows, we presented the project background and scientific basis for the research area, we then formulated and presented the research questions with explanations on how to address them. Later on, we presented the planed methodologies for the research, followed by the current progress and the works lying ahead.

2 Project Background and Scientific Basis

Over the past two decades, with the advancements of data engineering and analytics, supply chain analytics has become an integral part of every supply chain management system [6]. Supply chain analytics is to extract meaningful information and insights from the available data sources within supply chain processes to support supply chain decision makers [7]. In this regard, VA plays a significant role in enhancing the analytical capabilities of the firms and it is crucial for companies to have an adequate and reliable visual analytic solutions to use in their supply chain management activities. Given the availability of massive amount of data from different sources of the supply chain [8], an established, integrated, and functional VA system helps both manufacturers and distributers to have a better understanding of their processes and to make better decisions in their marketing, sales, production, and procurement planning [9].

On the other hand, VA is even more essential for the Fast Moving Consumer packaging Good (FMCG) industry [2]. Indeed, FMCGs may have a short shelf life as well as being prone to various sale's uncertainties, therefore, requires its supply chain decision-makers to have reliable analytical tools to understand and analyse the supply chain data dynamically and efficiently [10]. This is where VA can be used.

VA at its core, is the combination of data visualization with advanced analytical techniques that support decision makers to make better judgements by making sense of related data visually [11]. The underlying disciplines that encompass activities around developing visual analytic solutions are varying as follows: Information Systems (IS), Data and Knowledge Management (DKM), Business Intelligence (BI), Artificial Intelligence (AI), and Human-Computer Interaction (HCI), hence, making it to be a multi-disciplinary domain. Therefore, it is required to study the topic from different aspects. The conjunction of supply chain VA and IS occurs when we are dealing with the information flow throughout various points of the supply chain which requires various stakeholders of the supply chain to share information both downstream and upstream, that is, information flow integration [12]. DKM activities involve with the data discovery to identify hidden knowledge using VA, which also requires extensible explanatory data analysis for predictive, prescriptive and descriptive analysis [13]. Activities regarding the conjunction of BI and VI deals with combining business information with business management decisions manifested through visualization [14]. AI includes the ML supported VA to provide automated analytical capabilities for features such as what-if scenario analytics [15]. Finally, HCI deals with questions of human perception, cognition, intelligence, decision-making and interactive techniques of visualization [16].

Therefore, there is a need to develop a framework that embody various aspects of a comprehensive supply chain VA system. In general, there are three categories of frameworks that can be considered in relation to the VA [11], including:

1. Frameworks that describe human sensemaking: These frameworks illustrate the complex procedures human use to gain knowledge from the data, which should be considered when developing sensemaking systems.
2. Frameworks that describe interaction and information design: These models demonstrate how data transformation and analytic computation are utilized and integrated in providing the visual representations of data analysis in embedded tools.

3. Frameworks that describe improvement of ML techniques: Basically, these frameworks use visualization of ML processes in various steps in order to provide the capability to the user to participate in the ML process and enhance the result of the ML with different type of interactions.

Since the objective of this project is to explore methods for consolidating, processing, and analysing the collected data from multiple sources for the purpose of VA, the second frameworks within abovementioned types should be considered. In this regard, such a framework conceptually consists of the following modules [17]:

1. Data Generation (DG): Data Generation relates to the ability of the VA system in an organization to obtain, ascertain, generate and access the required data from heterogeneous sources across the supply chain.
2. Data Integration and Management (DIM): DIM is the aspects corresponding to the underlying potentials of various actors of the supply chain in utilizing the tools and methods for the collection, transformation, storage, and integration of the heterogeneous data. In fact, it deals with both the potentials and limitations of the organizations to comply and handle the real-time data collection and processing.
3. Data Analytics (DA): Data analytics part of the framework deals with the analytical techniques and methods that organizations can utilize to analyze supply chain data to gain significant insights for providing knowledgeful information to decision makers.
4. Data Visualization (DV): This aspect of the framework considers the potential of the organizations for using the proper tools and techniques for visual presentation of the data-driven insights to the decision makers.
5. Data-driven Decision Making (DDM): These are the essential strategic thoughts and behaviors that forms the culture of the company to be data-driven decision making oriented.

Given the multidisciplinary nature of the problem, a conceptual framework consisting of interlinked concepts from various disciplines provides an understanding of the construct of a supply chain VA system and its processes [18]. Therefore, a holistic framework covering the underlying concepts and processes for developing a supply chain VA system is highly desirable. Eventually, the aim of the research is to propose a framework that provides a comprehensive guideline covering various activities and procedures required to integrate VA into supply chain analytics.

3 Research Agendas and Questions

Generally, a potential supply chain VA framework should be developed based on the previous literature, expert opinions of supply chain managers, business intelligent experts, and data analysts. The framework should cover various activities including, identification of decisions and tasks that should be supported with the help of data analytics, collection of required data for the corresponding data analytics tasks, characterization of the required team of visual analysts, defining the supply chain analytic task as a VA problem, implementation of VA technique and tactics, and interpretation of the VA results for decision support.

Based on the objectives of the study and the scientific basis previously presented, following research questions are formulated. Since the main objective of this project is to create a framework, the research questions are generated towards identifying the requirements in various modules of the framework presented in section II. Each research question formulated are presented along with a prospective plan towards answering them:

1. What is the desired data flow to integrate VA into supply chain decision making and what are the existing obstacles within the supply chain data flow and how to overcome those obstacles? This question can be answered with the help of investigating the current literature, interviewing the participating companies, and data mapping strategies. The answer to this question contributes to the data generation and data integration and management parts in the formulation of the framework design.
2. How can the big-data analysis techniques and tools be adapted for processing the supply chain big data for VA? This question can be answered by investigating the cons and pros of the current available tools and techniques with experiments and identifying the opportunities for tailoring and adapting them into the context of FMCG supply chain. The answer to this research question contributes to the Data Analytics and Data-driven decision-making parts of the framework.
3. How can the visualization tools and techniques be implemented for providing data-driven decision support? The answer to this question is handled by investigating the current visualization tools and techniques, both from the literature and the industry, and formulating a future road map perspective for the supply chain to deploy efficient and reliable visualization solutions. The answer to this question contributes to the Data Visualization and data-driven decision-making part of the framework.

4 Scientific Research Method

The main research strategy for this project is planned based on adopting the Design Science approach, which revolves around identifying the organizational problems and creating and evaluating artifacts that can address those problems [19]. It should be noted that, this PD project is part of a bigger project which has its objective in exploring the road towards a sustainable food supply chain in Norway, and some of the Norway's food product manufacturers are among the industrial partners of the projects, namely, Brynild Gruppen AS [20] and Mills AS [21]. In this regard, the industrial partners of the project will contribute with their potentials to act as focal companies for implementing and evaluating the artifacts.

In general, design science research method includes three main cycles of activities [22]: The relevance cycle that identifies the problems and requirements along with opportunities in a given problem domain and environmental context, the rigor cycle that identifies the grounding of the research based on the scientific theories, methods, and domain experts experience and expertise, and finally, the design cycle that involves with the design of artifacts and evaluating them through development and implementation of prototypes. Following this approach, the research process consists of the following five phases.

1. Phase 1 constitutes of mapping and analyzing the current situation to identify problems/weaknesses and improvement potentials. Here, the main constructs of the research are identified, and the research hypotheses and questions are formulated, by reviewing current literature and requirement analysis. A systematic literature review is intended to identify state of the art regarding the topic. This phase expresses the relevance cycle of the research.
2. Phase 2 focuses on the determination of conceptual solutions and research paradigms after a careful study of available techniques and solutions. In this phase, we will conduct interview studies to obtain the experience and expertise of the people involved with supply chain analytical activities. It is also intended to conduct a number of case studies [23] to better recognize the fit of a possible solution in the corresponding context. This phase corresponds to the rigor cycle of the research.
3. Phase 3 establishes the constructs for prospective specific solution in close collaboration with researchers and practitioners participating in the project. This phase includes the main activities toward developing the objective framework. Designing the framework as an artifact that addresses the problems identified in previous phases are carried out in this phase. The idea is to apply research through design methodology [24] towards proposing a design artifact for a VA framework. This phase is the first part of the design cycle.
4. Phase 4 includes the analysis and evaluation of the proposed framework by deploying and testing selected elements of the solution from phase 3 in the form of demonstrators and prototypes [25]. Here, we collaborate with our industrial partners of the project to implement and evaluate prototypes in real environments. This phase completes the design cycle of the research.
5. Phase 5 amalgamates the findings towards answering the research questions and fulfilling the project objectives.

5 Project progress and timeline

The progress of this project started with conducting a Systematic Literature Review (SLR) in connection with the intersection of VA and supply chain. The aim of the SLR was to investigate the state of the art in VA systems applied in supply chain decision making progress. The main objectives of the SLR were to identify the current types of VA used for supply chain activities, to recognize different supply chain activities and decisions that can be supported with VA, and to identify different types of data that can be used in VA system. As the result, a categorization of VA techniques and methods has been proposed based on various supply chain processes. It is then mapped to the need for specific data for each of the applications and the decisions they intend to support.

Based on the results of SLR, 8 decision areas have been identified to be the main objectives of VA systems to be addressed including, sales management, network design, collaborative forecasting, demand management, network integration and visibility, transportation management, and operation management. These decision areas are to be used to address the first part of the framework, that is, identification of decisions and tasks that should be supported with the help of data analytics. Adding to that, required visualization techniques and tactics towards achieving each of the analytical goals have been identified. The results of the SLR, provides the ability to understand the decision areas that VA can support, following with specific data requirements for each of the tasks along with the appropriate visualization techniques need to be implemented. One of the gaps that needs to be further considered are providing the ability to analysts to interact with the system and logging the interactions for future reviews and comparisons. Furthermore, given that sales data are the main data used for analysis and the spatio-temporality nature of such data, the visualization techniques to be used should possess the ability to provide analysis against time and location attributes.

The 1st Year of the of the project was dedicated to the conducting of the literature review study and investigating existing commercial solutions to identify various fields of study in the context of the project. This year will cover the phase 1 and 2 of the project by identifying the main constructs of the research based on the examining of a broad range of literature and available commercial solutions and mapping them with an understanding of the requirements of the project.

The next step of the research in progress is to conduct an interview study with supply chain manager, business intelligent experts, and data analysts within our project partners and some external participants. The aim of this study is to understand the requirements of the industry in the data analysis tasks they perform, the current tools they use, the gaps in data availability, and the extent to which data-driven decision making are integrated into their activities based on analysis. Eventually, we aim at providing a mapping between the requirements and gaps from the industry and proposals of the literature. This participates extensively to the development of the framework. Currently, we are conducting interviews with stakeholders from the industrial partners of the project to identify the available data sources, the existing limitations of their current solutions and to identify the opportunities for future development of the framework. Later, we will map the requirements of the industry identified through the interview study to the proposals found in the systematic literature review.

The 2nd Year of the project devotes to the design and development of the framework based on the knowledge, requirements, and solutions acquired in the 1st year. A number of scientific publications are to be expected in this year regarding various aspects of the framework.

The 3rd Year of the project mostly devotes to the implementation of the framework by deploying one or more case studies either for whole the framework or some of its element and evaluating the performance of the framework. The framework is to be employed in the case study with the help of our project partners.

The documentation and dissemination of the findings and the project progress are to be carried out during all the project stages scientifically and professionally in various publication channels such as scientific journals and magazines, national and international conferences, workshops, seminars, and media. All of which will contribute to the production of the final PhD thesis.

6 Acknowledgment

This work is supported by DigiMat project granted by Research Council of Norway (Project Number: 296686). I would like to thank Professors Ricardo Colomo-Palacios and Antonio Martini for supervising this PhD project. I also want to thank Brynild Gruppen AS and specifically, Mathias Holm for organizing and managing the DigiMat project.

References

1. Manders, J.H.M., Caniëls, M.C.J., Ghijsen, P.W.Th.: Exploring supply chain flexibility in a FMCG food supply chain. Journal of Purchasing and Supply Management. 22, 181–195 (2016). https://doi.org/10.1016/j.pursup.2016.06.001.
2. Gupta, D.: FMCG Case Study. In: Gupta, D. (ed.) Applied Analytics through Case Studies Using SAS and R: Implementing Predictive Models and Machine Learning Techniques. pp. 345–396. Apress, Berkeley, CA (2018). https://doi.org/10.1007/978-1-4842-3525-6_7.
3. Tiwari, S., Wee, H.M., Daryanto, Y.: Big data analytics in supply chain management between 2010 and 2016: Insights to industries. Computers & Industrial Engineering. 115, 319–330 (2018). https://doi.org/10.1016/j.cie.2017.11.017.
4. Sacha, D., Stoffel, A., Stoffel, F., Kwon, B.C., Ellis, G., Keim, D.A.: Knowledge Generation Model for Visual Analytics. IEEE Transactions on Visualization and Computer Graphics. 20, 1604–1613 (2014). https://doi.org/10.1109/TVCG.2014.2346481.
5. Pak Chung Wong, Thomas, J.: Visual Analytics. IEEE Computer Graphics and Applications. 24, 20–21 (2004). https://doi.org/10.1109/MCG.2004.39.
6. Trkman, P., McCormack, K., de Oliveira, M.P.V., Ladeira, M.B.: The impact of business analytics on supply chain performance. Decision Support Systems. 49, 318–327 (2010). https://doi.org/10.1016/j.dss.2010.03.007.
7. Sahay, B.S., Ranjan, J.: Real time business intelligence in supply chain analytics. Information Management & Computer Security. 16, 28–48 (2008). https://doi.org/10.1108/09685220810862733.
8. Rozados, I.V., Tjahjono, B.: Big data analytics in supply chain management: Trends and related research. Presented at the January 6 (2015).
9. Huang, T., Fildes, R., Soopramanien, D.: The value of competitive information in forecasting FMCG retail product sales and the variable selection problem. European Journal of Operational Research. 237, 738–748 (2014). https://doi.org/10.1016/j.ejor.2014.02.022.
10. Doganis, P., Alexandridis, A., Patrinos, P., Sarimveis, H.: Time series sales forecasting for short shelf-life food products based on artificial neural networks and evolutionary computing. Journal of Food Engineering. 75, 196–204 (2006). https://doi.org/10.1016/j.jfoodeng.2005.03.056.

11. Endert, A., Ribarsky, W., Turkay, C., Wong, B.L.W., Nabney, I., Blanco, I.D., Rossi, F.: The State of the Art in Integrating Machine Learning into Visual Analytics. Computer Graphics Forum. 36, 458–486 (2017). https://doi.org/10.1111/cgf.13092.
12. Patnayakuni, R., Rai, A., Seth, N.: Relational Antecedents of Information Flow Integration for Supply Chain Coordination. Journal of Management Information Systems. 23, 13–49 (2006). https://doi.org/10.2753/MIS0742-1222230101.
13. Chehbi-Gamoura, S., Derrouiche, R., Damand, D., Barth, M.: Insights from big Data Analytics in supply chain management: an all-inclusive literature review using the SCOR model. Production Planning & Control. 31, 355–382 (2020). https://doi.org/10.1080/09537287.2019.1639839.
14. Jun, S.: Business Intelligence Visualization Technology and Its Application in Enterprise Management. In: Proceedings of the 2020 2nd International Conference on Big Data Engineering and Technology. pp. 45–48. Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3378904.3378909.
15. Diez-Olivan, A., Ser, J. [Del, Galar, D., Sierra, B.: Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. Information Fusion. 50, 92–111 (2019). https://doi.org/10.1016/j.inffus.2018.10.005.
16. Holzinger, A.: Human-Computer Interaction and Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together? In: International Conference on Availability, Reliability, and Security. pp. 319–328. Springer (2013).
17. Arunachalam, D., Kumar, N., Kawalek, J.P.: Understanding big data analytics capabilities in supply chain management: Unravelling the issues, challenges and implications for practice. Transportation Research Part E: Logistics and Transportation Review. 114, 416–436 (2018). https://doi.org/10.1016/j.tre.2017.04.001.
18. Jabareen, Y.: Building a Conceptual Framework: Philosophy, Definitions, and Procedure. International Journal of Qualitative Methods. 8, 49–62 (2009). https://doi.org/10.1177/160940690900800406.
19. Hevner, A., Chatterjee, S.: Design Science Research in Information Systems. In: Hevner, A. and Chatterjee, S. (eds.) Design Research in Information Systems: Theory and Practice. pp. 9–22. Springer US, Boston, MA (2010). https://doi.org/10.1007/978-1-4419-5653-8_2.
20. Brynildgruppen - Hjem, https://www.brynildgruppen.no, last accessed 2019/11/22.
21. Mills AS - Mat- og merkevarehuset på Grünerløkka, https://www.mills.no/, last accessed 2020/02/17.
22. Hevner, A.R.: A Three Cycle View of Design Science Research. Scandinavian Journal of Information Systems. 87–92 (2007).
23. Runeson, P., Höst, M.: Guidelines for conducting and reporting case study research in software engineering. Empir Software Eng. 14, 131 (2008). https://doi.org/10.1007/s10664-008-9102-8.
24. Zimmerman, J., Forlizzi, J.: Research Through Design in HCI. In: Olson, J.S. and Kellogg, W.A. (eds.) Ways of Knowing in HCI. pp. 167–189. Springer New York, New York, NY (2014). https://doi.org/10.1007/978-1-4939-0378-8_8.
25. Houde, S., Hill, C.: Chapter 16 - What do Prototypes Prototype? In: Helander, M.G., Landauer, T.K., and Prabhu, P.V. (eds.) Handbook of Human-Computer Interaction (Second Edition). pp. 367–381. North-Holland, Amsterdam (1997). https://doi.org/10.1016/B978-044481862-1.50082-0.