From DSS to Data Science:
The Effect of Industry 4.0

Natheer Khleaf Gharaibeh
https://orcid.org/0000-0002-6566-9704
Taibah University, Saudi Arabia

INTRODUCTION

Major evolution of information systems taken place during the second half of the twentieth century during the shift from Industry 3.0 to Industry 4.0, accompanied with transitional period from 20th century to 21st century, jargon and metaphoric words emerged (Laudon and Laudon, 2022). The fields of data science and big data as an example of one of the most important sciences emerging from these transformations are full of jargon and metaphoric words. For example, data lakes represent a single repository of data that includes raw copies of source system data, sensor data through IoT (Internet of Things), social data, and converted data used for reporting, visualization, advanced analytics, and machine learning (Sharda et al., 2021). This is great, but things aren't as beautifully simple as the metaphor of the Lake. If we don't solve the problem of data silos, we may find ourselves facing a data swamp. However, to understand these metaphors we need to know the historical contexts behind these terms. For example, the concept of data silos is information dating back to 1988 (Trkman et al., 2015). The period in which enterprise application emerged, a new era of computing is appearing with more integration and holistic paradigm due to the invention of web in 1990s. This is known as Industry 4.0 technologies, which comes decades after Industry 3.0, in which the big data technologies thrived to integrate more data sources using NoSQL and Hadoop. Data lakes is a good example for of jargon that appeared in this era as an emerging approach to cloud-based big data, in which Enterprises across industries are starting to collect all the different types of data sources through a common ingestion framework like Hadoop-based repository (Stein & Morrison, 2014). This was considered to be 10 to 100 times less expensive to deploy than conventional data warehousing.

The problem of diverse concepts and ambiguous metaphors may be misleading if it is not understood in its proper context. In this chapter, it is proposed that the main reason of this problem is due to isolation between business and technical issues. The good news is that the concepts of DSS still work, and it can bring together all the contradictions. DSS evolved to include what became known as Analytics. The broad and diverse nature of DSS field makes it with high potential to consolidate and provide suitable interpretation for these diverse concepts that evolved through a long period of time.

This methodological suggestion came from the combining and adaptation of commonly used models and procedures in information system research, such as the socio-technical model for information systems. Socio-technical model for information systems views DSS as socio-technical systems with several viewpoints and potential conceptual framework (Hirschheim, Klein, & Lyytinen, 1995). For control or emancipation of effective design (Shim et al., 2002) requires knowledge of both the application domain and the solution domain. The methodology that is used to tackle this dual problem is based on dialectic theory, which was used several times in IS and DSS literature (Elgarah, Haynes, & Courtney, 2002;

DOI: 10.4018/978-1-7998-9220-5.ch085

Haynes, 2001). Further, it was used in similar and related fields, as Boehm did in his Hegelian dialectical (related the German philosopher Hegel) view of the history of the field of software engineering (SE). Boehm (2006) traced evolution of software development methodologies from the classic process models such as waterfall into a spiral model, prototyping and, to agile development methodologies. The latter models seem to have worked well for DSS (Gharaibeh & Soud, 2008; Gharaibeh et al., 2009). In this chapter DSS will be considered as a software system with unique characteristics, from both SE and DSS literature it is found that they have similar beginning, on the one hand power (Shim et al., 2002) in his historical review of DSS field [stated that SAGE (Semi-Automatic Ground Environment- air defense system for North America) is the first computerized DSS, on the other hand, Boehm stated that SAGE was the most ambitious information processing project of the 1950's. By tracing the evolution of concepts, we mitigate the Epistemological rupture (or epistemological break) according to the French philosopher Gaston Bachelard.

There is a general disconnection or gap of concepts between early concepts of DSS and current concepts of data science era. From this viewpoint, there was no research done to investigate the DSS evolution till data science age. This chapter is an extension to the researcher's efforts to push the boundaries of DSS research (Paradice, 2007). DSS evolved to include what became known as Analytics.

In this chapter, section 2 begins with the problem of data silos, and the need for DSS and enterprise applications in solving the problem of data integration and overcoming these silos. Then, section 3 displays the progress in industry which can be portrayed by Kondratieff waves and the four industrial revolutions, showing how to balance between these two concepts and focusing more on the role of IS discipline in industry 4.0. Section 4 and 5 focus on integration shift within the IS discipline and DSS integral position within the IS discipline, they present the solution provided by enterprise applications and how DSS played more comprehensive role within the progress of IS discipline. We finally end with section 6, dialectic of technology and business, which shows the dialectic of technology and business, explaining the thesis of technology-driven and its antithesis of business-driven, then synthesis: a more holistic paradigm in Industry 4.0.

THE PROBLEM AND METHODOLOGY

The Problem of Data Silos

Since the prevalent appearance of data science as a field, its practitioners have asserted that 80% of the work involved is acquiring and preparing data (Wilder-James, 2016). In today's digitized economy and from business perspective, incorporating data as a competitive advantage is necessary for Business, so why is it so hard to get access to the data we need? The biggest obstacle to using analyzing the data is the plain old access to the data in other words the Data Silos, "Organizational silos and a dearth of data specialists are the main obstacles to putting big data to work effectively for decision-making" (Jha, Jha, & O'Brien, 2016). These silos are data islands that make extracting data and putting it to other uses prohibitively expensive. They can occur for a variety of reasons. We need to reduce the impact of data silos on all company activities to shift to higher value uses and preserve a strategic advantage. The capacity to use data effectively is a significant competitive advantage. There's a lot of work to be done in integrating all the data sources to get to a future state of mature analytical proficiency. This is a strategic goal for all businesses, and if correctly addressed, will lead to the development of experience and a data infrastructure that unlocks every subsequent stage.