Chapter 41
A Simulation Model for Application Development in Data Warehouses

Nayem Rahman
Portland State University, Portland, OR, USA

ABSTRACT

Software development projects have been blamed for being behind schedule, cost overruns, and the delivery of poor quality product. This paper presents a simulation model of a data warehouse to evaluate the feasibility of different software development controls and measures to better manage a software development lifecycle, and improve the performance of the launched software. This paper attempts to address the practical issue of code defects in each stage of data warehouse application development. The author has compared the defect removal rate of their previous project to the newly proposed enhanced project development life cycle that uses code inspection and code scorecard along with other phases of software development life cycle. Simulation results show that the code inspection and code score-carding have achieved a significant code defect reduction. This has also significantly improved the software development process and allowed for a flawless production execution. The author proposes this simulation model to a data warehouse application development process to enable developers to improve their current process.

1. INTRODUCTION

Software development is laborious, expensive and unreliable. Hence, software development projects quite often encounter schedule slippage, cost overruns, and poor-quality software in both commercial and government sectors (Raffo & Wernick, 2001). To address this potential issue, we propose changes to the software development process. Smith and Rahman (2017) observe that "without efficient processes through which Information Technology (IT) builds and supports the technology, the full business-value potential will remain unrealized." Bringing the software project lifecycle under the radar of simulation

DOI: 10.4018/978-1-6684-3702-5.ch041

models could be good effort (Kellner et al., 2001). For the last four decades systems dynamics modeling and simulation techniques were applied in diverse disciplines of scientific, engineering and manufacturing processes (Richardson, 2013; Rashidi, 2016). According to the Merriam-Webster Online Dictionary, "simulation is the imitative representation of the functioning of one system or process by means of the functioning of another." Simulations run in simulation time, an abstraction of real-time (Imagine That Inc., 2014).

Simulation models are used to solve problems that arise in manufacturing (Barra Montevechi, 2016), business process design (Liu & Iijima, 2015), inventory management system (Cobb, 2017) and health care decision-making (Chick, 2006; Chen & Zhao, 2014). Martinez-Moyano and Richardson (2013) and others (Morrison, 2012; Mould & Bowers, 2013) listed 41 best practices of systems dynamics modeling and categorized them in terms of problem identification and definition, system conceptualization, and model formulation. Hughes and Perera (2009) argue that simulation could be integrated as a daily tool to solve problems. They present an easy-to-follow framework – consisting of five key stages, such as foundation, introduction, infrastructure, deployment and embedding - for enabling companies to embed simulation technologies into their business processes (Hughes & Perera, 2009). The work of Eatock et al. (2001) indicates that describing the dynamic behavior of IT could be very helpful for business process modelers in predicting the impact on organizational processes (Eatock et al., 2001). Software process simulation is suggested to be helpful to achieve higher Capability Maturity Model (CMM) levels in software development (Raffo et al., 1999).

In software engineering, simulation modelling has attracted considerable interest during the last decade (Ahmed et al., 2008). Software process simulation is used mainly to address the challenges of strategic management of software development and to support process improvements (Raffo & Kellner, 2000). In this work, we are making an attempt to leverage simulation modeling in a data warehouse application development. We developed a simulation model based on defined processes for the application development of a data warehouse reporting environment called Next Generation Capital Reporting (NGCR). In our recent project, we developed and implemented a Financial Reporting System (FRS) in the Enterprise Data Warehouse (EDW) environment. A data warehouse is used a central repository of data of medium and large business organizations. A data warehouse is considered as one of the key infrastructures of IT. And the capability of IT has a strong correlation between the agility and performance of an organization (Rahman, 2016a).

Industry and academic research suggest that when an EDW becomes successful at an initial stage more and more application development and reporting projects start to land each year. Research suggest that there is a correlation between the increase of the size of the data warehouse and increase of the value it brings to an organization (Rahman, 2016b). Engineering projects are complex (Rahman et al., 2016). From the standpoint of EDW projects complexity arise from multiple factors including different stakeholders' buy-in, software development effort, frequent changes of relevant tools and technologies, query performance issue, data validation, data quality, and production release timeliness. Therefore, it is quite challenging to maintain a stable data warehouse environment given it is an enterprise platform as well as a shared environment. Hence, some degree of discipline is needed in code development, code changes, testing, code performance optimization, system resource usage and configuration of integration specification (Rahman, 2016b). An EDW-specific simulation model can help in bringing discipline in data warehousing applications development and in improving overall quality of application development and reporting environment.