Chapter 2
Data Warehouse and Data Virtualization

Kanishka Mukherjee
GVK Biosciences Private Limited, India

ABSTRACT

This chapter contains as well as illustrates different innovations that is changing the way Clinical Drug Development and Safety organizations perceives IT and the ways and means through which these innovations are facilitating the change in business itself. The main content contains illustrations of two structurally different means to create data warehouses, the benefits of the approaches and the difficulties. It also explains the importance of data virtualization technology when implemented in the Clinical and Safety Organizations.

INTRODUCTION

Life Sciences Companies are not facing the lack of data as a problem any longer; their problem is absence of '*right*' data at '*right*' time. Business leaders not only require '*right*' data, they require it at '*right*' *(real)* time in order to effectively define the strategic direction of the organization. The problem is, most of the currently available of applications and/or software was designed for functionality, most of the data warehouses and the associated means of information dissemination are based on reports, with a built in time-lag. This is in order to either execute complex algorithms on voluminous data or due to traditional ETL approach for data transportation.

The innovative future generation of applications will require real time access to data and timely analytics as well. While strategically this would require a change of corporate mindset; technically it will require establishment of correct data structures, in order to not only analyze internal data but also to absorb external (structured/unstructured) data, so that algorithms can execute seamlessly to provide uniformly right output to help in deducing right strategy.

In this chapter, we will go through the details of innovative means and mechanisms of utilizing optimal Clinical and Safety aligned data structures; from which and to which, appropriate data can be extracted and retained, for various forward looking as well as past action analysis. We will also look very briefly

into importance and integration aspects of external unstructured data (including but not limited to social media) which not only provides voluminous relevant information but also how can they be segregated and analyzed with various parameters to provide extremely worthy and pin-pointed inputs about geographies, markets, competition, products, usage, user satisfaction, adverse reactions etc. We will also discuss to some detail on methods used for data virtualization and the advantages of data virtualization

The chapter is divided into multiple sections and subsections describing various innovations in data warehouses in Clinical and safety areas. First two sections of this chapter deals with Business Case and Resources (human and infrastructure/software etc.) though these two topics are not must for this chapter, I felt, it is important to understand these topics for innovations and successful implementation of a data warehouse project.

1. DEVELOPMENT OF BUSINESS CASE

Multiple studies has shown historically Data Warehouse project don't have high success rate. A study published by Amin & Arefin (2010), list top 20 reasons on why 50% of the Data Warehouse projects fail, 6 out of 20 reasons listed, deals with failure to acquire right Business Case and failure to perform right stakeholder mapping at the start of the project. The main reason to include this section in detail in this chapter is to put further emphasis on the fact when building an innovative application where results could be radically beautiful on one side, on the other side, chances of failure are also high. In such conditions, the business case assumes greater importance.

The most important factor, of the whole sequence, would be for the organization and its leaders to jointly decide, on exactly what they perceive as outcome and what the measurable benefits of the upcoming system are and hence a business case should be developed. It is essential that along with the objectives of the project the Business Case includes as key success criteria; the right accountability structure within the organization as a whole (*business owner*) and for the project organization (*project manager*), right business needs, cost, and estimated business benefit to be derived, on a clear time scale. Who (*business teams*) gets what (*functionality, reports etc.*) when (*timeline*), for how much (*cost, resource/ SME allocation*) should be documented. Finally the risks associated with the achievement of the project objectives is crucial to be identified and documented albeit at a high level at the very beginning, since if kept fluid, this is observed to be the single most important cause of failure of data warehouse projects. A feasibility study could be undertaken during the development of Business case. It allows for a more detailed analysis and assessment of the business problem/s to be dealt by the project team. It also support the team to identify risks and visualize possible pitfalls in advance and finally it present options to mend the solution design envisaged for the project.

It is also vital that one keeps in mind the right template of the Business case which aids to record the essentials, I will not delve into the details and the formats of the business case templates used in the industry however the reader should remember that all the relevant information gets identified, agreed upon, documented and signed off before starting to plan the project execution. One has to remember that this document is used as reference multiple times during the project. At each Quality Gateway the Business Case is used to validate if the benefits, costs and risks currently projected are matching to what was projected in the Business Case. During the Project Closure Meeting the Business Case is discussed with all the stakeholders and the Business Realization Metrics (BRM) is base lined; in this regard, the

Related Content

Ethnobotanical, Phytochemical, and Pharmacological Aspects of Hemidesmus indicus: A Herbal Bliss for Mankind
Minakshi Rajput, Navneetand Akash (2020). *Ethnomedicinal Plant Use and Practice in Traditional Medicine (pp. 154-180).*
www.irma-international.org/chapter/ethnobotanical-phytochemical-and-pharmacological-aspects-of-hemidesmus-indicus/251620

Ethnobotanical Uses of Medicinal Plants for the Treatment of Herniated Disc in Turkish Traditional Medicine
Mustafa Eray Bozyel, Elif Merdamert Bozyeland Kerem Canli (2020). *Ethnomedicinal Plant Use and Practice in Traditional Medicine (pp. 221-234).*
www.irma-international.org/chapter/ethnobotanical-uses-of-medicinal-plants-for-the-treatment-of-herniated-disc-in-turkish-traditional-medicine/251624

Ethnic Use, Phytochemistry, and Pharmacology of Cyperus rotundus: A Medicinal Plant
Mohammed Rahmatullah, Khoshnur Jannat, Gerald R. Reeck, Rownak Jahan, Taufiq Rahman, Nasrin A. Shovaand Maidul Islam (2020). *Ethnomedicinal Plant Use and Practice in Traditional Medicine (pp. 82-104).*
www.irma-international.org/chapter/ethnic-use-phytochemistry-and-pharmacology-of-cyperus-rotundus/251617

Phytochemistry and Antigenotoxic Properties of Six Ethnobotanically Important Members From the Family Zingiberaceae
Anish Nag (2020). *Ethnomedicinal Plant Use and Practice in Traditional Medicine (pp. 131-153).*
www.irma-international.org/chapter/phytochemistry-and-antigenotoxic-properties-of-six-ethnobotanically-important-members-from-the-family-zingiberaceae/251619

Semantic Interation, Text Mining, Tools and Technologies
Chandrakant Ekkirala (2016). *Software Innovations in Clinical Drug Development and Safety (pp. 47-63).*
www.irma-international.org/chapter/semantic-interation-text-mining-tools-and-technologies/138979