Incorporating External Data Into a BI Solution at a Public Waste Management Organization

Mattias Strand, University of Skövde, Skövde, Sweden
Anna Syberfeldt, University of Skövde, Skövde, Sweden

https://orcid.org/0000-0003-3973-3394

ABSTRACT

Organizations are showing an increasing interest in incorporating external data into their business intelligence solutions. Such data allows for advanced analytics and enables more comprehensive and inclusive decision-making. However, external data incorporation is relatively unexplored in the literature, and scientifically published details on up-and-running BI solutions are very sparse. In addition, published literature concerning the incorporation of external data into BI solutions is often rather synoptic or rather old (originating from data warehouse related literature). Therefore, the authors present the results of an action case study at a public waste management organization, illustrating detailed aspects of external data incorporation related to the back-end of the solution such as data selection, source characteristics, acquisition technologies and frequencies, and integration approaches. Given that the external origin of the data poses specific problems that must be overcome in order to allow for successful incorporation initiatives, special attention was paid to such problems.

KEYWORDS

Business Intelligence, Case Study, External Data, Waste Management

1. INTRODUCTION

Most organizations are operating in a business climate of intense competition, where management must pay close attention to the whereabouts of a myriad actors. Although this is a daunting task, there are no shortcuts to maintaining a competitive edge. Organizations that neglect to do this are at severe risk of being overwhelmed by competitive forces.

In order to maintain their edge, organizations need advanced systems that are able to present a holistic view of the organization and its business in a timely and accurate manner. Various (often overlapping) solutions have been proposed in the literature, involving concepts such as business intelligence, decision support systems (DSS), analytics, and big data. For this work, we have used the concept business intelligence (BI), since it has attracted enormous interest in recent decades and is still regarded as the gold standard by many chief information officers (Watson & Wixom, 2007, Kappelman et al., 2013).

DOI: 10.4018/IJBIR.2019070104
However, even the concept of BI is defined in different ways. Given that this work has a very specific focus on the data aspect of BI solutions, we have adopted Wixom and Watson’s definition of a BI solution as “a broad category of technologies, applications, and processes for gathering, storing, accessing, and analyzing data to help its users make better decisions” (Wixom and Watson, 2010, p. 14). BI could also be described very simply as getting data in and getting data out. The part of the system architecture supporting getting data in is referred to as the back-end.

BI projects have traditionally focused on incorporating data from systems that are internal to the organization. The benefits of using data acquired from outside the organization (external data) have seldom been considered (Devlin, 1997). For clarity, external data in this case refers to data that is incorporated from systems/suppliers external to the organization. Ram et al. (2016, p.221) claim that “currently, BI-solutions mainly focus on structured and internal data of an enterprise. As a result, a lot of valuable information embedded in unstructured and external data remains hidden, which could potentially lead to an incomplete view of the reality and resultantly biased business decision-making.” Yet external data may contribute to other insights or show other facets of a customer, a product, or a competitor, and so its importance has long been recognized (see, for example, Kelly, 1996; Alavi & Haley, 1997; Devlin, 1997; Chen & Frlickr, 2000; Huang et al., 2002; Inmon, et al., 1997; March & Hevner, 2007; Anderson-Lehman et al., 2008; Ponniah, 2010; Poletto et al., 2017; Jukic et al., 2017).

Although the literature strongly emphasizes the importance and benefits of external data in relation to BI (and to data warehousing, business analytics, and big data), much remains to be learned. The literature lacks detailed accounts of how external data have been incorporated in fully functional systems, aiding decision-makers in business contexts. In addition, the existing scientific literature is mostly very general or rather old (as seen above, some published more than two decades ago). Some of this literature was included in this work, partly to illustrate the need for more up-to-date knowledge but also because, despite our best efforts, we have not managed to find more recent literature. While white papers, practitioner journals and consultant resource web pages (e.g. Hendler, 2014, Page, 2015, Weil, 2015) may contribute with some details, they are often rather conceptual, software specific, and not externally validated.

In addition, most literature concerning external data incorporation has traditionally focused on data procured from companies specialized in collecting, aggregating, summarizing, and refining data. (These companies are sometimes referred to as syndicate data suppliers (SDSs) (Strand et al., 2003, Strand & Carlsson, 2008)). However, in recent years, there has been an increased interest in other types of external data sources, e.g. open data/public data (Baesens et al., 2016). They describe open data/public data, as data about topics such as weather, traffic, maps, environment and housing.

In order to alleviate the lack of information about the use of external data in general, and public data in particular, we present the results of an action case study conducted at a public waste management organization. The study started in 2010 and is still ongoing. The scope of this paper covers the activities of the external data incorporation process (EDIP) (Strand & Wangler, 2004) related to the back-end of the system, comprising identification, acquisition, and integration of external data from several sources, including open data/public data. We deal with such things as the selection of data sources, characteristics of the data sources, data acquisition approaches and frequencies, and data integration approaches. Thus, the study presented complements and extends previous studies investigating the application of external data in waste management organizations. Vitorino de Souza Melaré et al. (2017) show, as a result from a comprehensive literature review, that geographical information systems (GIS) is the main (and often only) data source used. In this study, we identify and use several additional external data sources for waste management and elaborate on their benefits.

At the same time, this work complements and updates the external data incorporation problems categorized and described by Strand et al. (2006). The problems they identified mainly originate from studies focused on external data procured from SDSs. By contrasting the most common problems presented by Strand et al. (2006) with problems experienced when incorporating external data (public data) from government sources, we also determine how generic the problems are, regardless of the data
Related Content

Lessons from the Private Sector: A Framework to Be Adopted in the Public Sector
www.igi-global.com/chapter/lessons-from-the-private-sector/142634?camid=4v1a

The Practice of Structural Equation Modeling
Ho Yin Wong, Jason Sit and Jia-Yi Hung (2014). Encyclopedia of Business Analytics and Optimization (pp. 2521-2530).
www.igi-global.com/chapter/the-practice-of-structural-equation-modeling/107433?camid=4v1a
Minimal Intelligence Agents in Double Auction Markets with Speculators
Senlin Wu and Siddhartha Bhattacharyya (2006). *Computational Economics: A Perspective from Computational Intelligence* (pp. 79-98).
www.igi-global.com/chapter/minimal-intelligence-agents-double-auction/6781?camid=4v1a

Social Spider Algorithm for Training Artificial Neural Networks
www.igi-global.com/article/social-spider-algorithm-for-training-artificial-neural-networks/238064?camid=4v1a