Chapter 1
Data Mining and Business Intelligence:
A Bibliometric Analysis

Ana Azevedo

https://orcid.org/0000-0003-0882-3426

CEOS.PP, ISCAP, Polytechnic of Porto, Portugal

ABSTRACT

From the middle of this second decade of the 21st century, analytics has become commonly associated with the topics business intelligence and data mining. Data mining (DM) is being applied with success in business intelligence (BI) environments and several examples of applications can be found. BI and DM have different roots and, as a consequence, have significantly different characteristics. DM came up from scientific environments; thus, it is not business oriented. DM tools still demand heavy work in order to obtain the intended results. On the contrary, BI is rooted in industry and business. As a result, BI tools are user-friendly. This chapter reflects on these differences from a historical perspective. Starting with a separated historical perspective of each one, analytics, BI, and DM, the author then discusses how they converged when DM is used and integrated in BI environments with success.

INTRODUCTION

Analytics represents a combination of computational technologies, scientific management techniques, and statistics to solve real-world problems, while considering that organizations have to analyze their data to understand what is happening, what will happen, and how to take the best option (Sharda, Delen & Turban, 2018). INFORMS defines analytics as the scientific process of transforming data into insights with the purpose of making better decisions. Usually, there can be considered three types of analytics, namely Descriptive, Predictive, and Prescriptive Analytics. Descriptive Analytics refers to knowing what happened and what is happening in organizations and perceiving some underlying trends and the causes of such occurrences. Predictive analytics aims to determine what is likely to happen in the future. Prescriptive Analytics aims to recognize what is happening as well as the most likely predictions

and make decisions to achieve the best performance. These three levels of Analytics are not exclusive, overlapping each other many times. From the middle of this second decade of the XXI century, Descriptive Analytics has become commonly associated with the topic Business Intelligence and prescriptive Analytics with the topic Data Mining.

BI is one area of the Decision Support Systems (DSS) discipline and refers to information systems aimed at integrating structured and unstructured data in order to convert it into useful information and knowledge, upon which business managers can make more informed and consequently better decisions. The term Business Intelligence (BI) was made popular by Gartner in 1989 (Power, 2007) (Zeller, 2007), but the first reference was made by Luhn in 1958 (Lunh, 1958), not necessarily with the same meaning. Being rooted in the DSS discipline, BI has suffered a considerable evolution over the last years and is, nowadays, an area of DSS that attracts a great deal of interest from both the industry and researchers (Azevedo & Santos, 2012).

The term Knowledge Discovery in Databases (KDD) was coined in 1989 (Wixon & Watson, 2010) to refer to the broad process of finding knowledge in data, and to emphasize the "high-level" application of particular data mining (DM) methods (Fayyad, Piatetski-Shapiro, & Smyth, 1996). The DM phase concerns, mainly, to the means by which patterns are extracted and enumerated from data. In recent years, DM has been applied with success to several diversified fields, such as bioinformatics, ecology and sustainability, finance, industry, marketing, scientific research, telecommunications, and several other applications, including BI.

DM is being applied with success in BI and several examples of applications can be found (Linoff, 2008) (Vercellis, 2009) (Hu & Cercone, 2004) (Cheung & Li, 2012) (Phan & Vogel, 2010). BI and DM have different roots and, as a consequence, have significantly different characteristics. DM came up from scientific environements, thus, in its roots, it is not business oriented (Kriegel, Borgwardt, Kröger, Pryakhin, Schubert, & Zimek, 2007) (Piatetsky-Shapiro, 2007). DM tools still demand heavy work in order to obtain the intended results, hence needing the knowledge of DM specialists to explore its full potencial value (Azevedo A., 2012). The main focus for DM researchers still is the improvement of algorithms and/or finding new algorithms that behaves better than others in some particular application, as can be recognized by a search in the main conferences and journal in the area of DM. On the contrary, BI is rooted in industry and business (Yermish, Miori, Yi, Malhotra, & Klimberg, 2010), thus it is business oriented. As a result, BI tools are user-friendly and can easily be accessed and manipulated by business users. The main focus for BI researchers is how to better use BI in organizations in order to improve decision making (Wixon & Watson, 2010).

Nevertheless, since the beginning of the second decade of the XXI century, DM and BI tend to converge. The gap existing between BI and DM is being filled (Azevedo & Santos, 2012) (Wang & Wang, 2008) (Hang & Fong, 2009). Researchers efforts are shifting towards the integration of DM in BI systens, and focusing on how DM can be used to improve decision making, and creating the possibility of DM tools being accessed and manipulated by business users at the same level as the other BI tools, thus making DM relevant to business. Analytics comes many times as a synonim of BI and both terms are used indistinguishably. At the beginning of the third decade of the XX1 century, we can see that the focus is now Big Data.

This papers presents DM and BI from a bibliometric comparative historical perspective. As far as our knowledge, there is no similar approach in the literature. The main contribution of this chapter is to provide a comparison between these different and yet convergent areas, from an historical perspective.

Related Content

ERP On-Premise or On-Demand
Fan Zhaoand Elias T. Kirche (2018). *International Journal of Business Analytics (pp. 1-16).*
www.irma-international.org/article/erp-on-premise-or-on-demand/201450

The Analytic Network Process – Dependence and Feedback in Decision-Making: Theory and
Validation Examples
Thomas L. Saaty (2006). *Business Applications and Computational Intelligence (pp. 360-387).*
www.irma-international.org/chapter/analytic-network-process-dependence-feedback/6033

Adaptive Intrusion Detection Systems: The Next Generation of IDSs
Hassina Bensefiaand Nassira Ghoualmi-Zine (2016). *Business Intelligence: Concepts, Methodologies,
Tools, and Applications (pp. 2189-2219).*
www.irma-international.org/chapter/adaptive-intrusion-detection-systems/142723

The State of Artificial Intelligence in Marketing With Directions for Future Research
Jing Chen, Jose Humberto Ablanedo-Rosas, Gary L. Frankwickand Fernando R. Jiménez Arévalo (2021).
International Journal of Business Intelligence Research (pp. 1-26).
www.irma-international.org/article/state-artificial-intelligence-marketing-directions/297062

Scalable Textile Business Intelligence With Operational Data Collection and Analytical
Processing
Kamalendu Pal (2024). *Data-Driven Business Intelligence Systems for Socio-Technical Organizations (pp.
85-104).*
www.irma-international.org/chapter/scalable-textile-business-intelligence-with-operational-data-collection-and-analytical-processing/344147