ARTICLE IN PRESS
JID: EOR [m5G;July 4, 2017;12:0]

European Journal of Operational Research 000 (2017) 1–16

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Developing a business analytics methodology: A case study in the foodbank sector

Giles A. Hindle[a,*], Richard Vidgen[b,c]

[a] Management Systems Group, Hull University Business School, University of Hull, Cottingham Road, Hull HU6 7RX, UK
[b] School of Information Systems, Technology and Management, University of New South Wales, Sydney NSW 2052, Australia
[c] School of Business, Economics and Informatics, Birkbeck University, London WC1E 7HX, UK

ARTICLE INFO

Article history:
Received 1 August 2016
Accepted 9 June 2017
Available online xxx

Keywords:
Analytics
OR for community development
Data mining
Problem structuring methods
Soft systems methodology

ABSTRACT

The current research seeks to address the following question: how can organizations align their business analytics development projects with their business goals? To pursue this research agenda we adopt an action research framework to develop and apply a business analytics methodology (BAM). The four-stage BAM (problem situation structuring, business model mapping, analytics leverage analysis, and analytics implementation) is not a prescription. Rather, it provides a logical structure and logical precedence of activities that can be used to guide the practice of analytics (i.e., a mental model). The client for the action research project is The Trussell Trust, which is a UK charity with the mission of empowering local communities to combat poverty and exclusion. As part of the action research project the research team created the UK's first dynamic visualization tool for crises related to food poverty. The prototype uses foodbank data to map geographical demand and aligns findings to 2011 Census data to predict where additional foodbanks may be needed. Research findings are that: (1) the analytics methodology provides an umbrella for, and applies equally to, data science and Operational Research (OR); (2) that the practice of business analytics is an entangled and emergent mix of top-down analysis and bottom-up action; and, (3) that, for the third sector in particular, analytics can be usefully approached as a collective and community endeavor.

1. Introduction

There is much excitement around business analytics and data science as commercial organizations explore how they can use their large volumes of data to create value in their business, and governments and communities seek to create value of a broader nature through exploitation of their data resources (Vidgen, Shaw, & Grant, 2017; Davenport, 2013; Davenport & Harris, 2007; McKinsey et al., 2011; Yui, 2012). A number of researchers have argued that the growing attention and prominence afforded to analytics presents an important challenge and opportunity for the OR (Operational Research) community (Liberatore & Luo, 2010; Mortenson, Doherty, & Robinson, 2015; Ranyard, Fildes, & Tun, 2015). Many in the OR community have sought to align themselves with analytics; for instance, INFORMS in the USA and The OR Society in the UK now offer analytics related events, training, certification and publications. However, the number of analytics-orientated studies in journals associated with OR is still comparatively low (Mortenson et al., 2015).

A popular view of analytics is encapsulated by Davenport and Harris' (2007) succinct and widely adopted definition: "By *analytics* we mean the extensive use of data, statistical and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions." (p. 7, emphasis in the original). Business analytics can also be viewed as sitting at the intersection of OR, artificial intelligence (machine learning) and information systems (Mortenson et al., 2015). It can be further characterized by descriptive (e.g., customer segmentation), predictive (e.g., customer churn modelling), and prescriptive (e.g., offer this loyal customer a discount) model building using data sources that may be heterogeneous (e.g., text, video) and 'big'. These models enable organizations to make quicker, better, and more intelligent decisions to create business value in the broadest sense – potentially the difference between survival and extinction in an increasingly competitive world. Thus, business analytics is concerned primarily with the context in which techniques from OR and data science are deployed.

Organizations are keen to jump on the analytics bandwagon but, as with previous phenomena, such as the growth of

* Corresponding author.
E-mail addresses: giles.hindle@hull.ac.uk (G.A. Hindle), r.vidgen@unsw.edu.au (R. Vidgen).

http://dx.doi.org/10.1016/j.ejor.2017.06.031

information technology in the 1990s and the dotcom bubble at the turn of the century, many are likely to waste money, resources and attention in their quest to become data-driven and to adopt evidence-based decision making. Consequently, how the application of analytics might unfold within organizations is a fertile area for research. George, Haas, and Pentland (2014), in a message from the editors of the Academy of Management Journal argue that "… management scholars will need to unpack how ubiquitous data can generate new sources of value, as well as the routes through which such value is manifest (mechanisms of value creation) and how this value is apportioned among the parties and data contributors …" (p. 324).

Thus, the current research seeks to address the following question: how can organizations align their business analytics development projects with their business goals and strategy? To pursue this research agenda we adopt an action research framework to develop and apply a business analytics methodology (BAM). Because the creation of business value is dependent upon an understanding of the nature of the 'business' in which analytics will be deployed, BAM adopts an approach based upon the emerging field of business modelling (Baden-Fuller & Haefliger, 2013; Zott, Amit, & Massa, 2011). Specifically, we draw on the business model canvas of Osterwalder and Peigneur (2010) in combination with problem structuring and modelling tools from the soft systems methodology (SSM) (Checkland, 1981; Checkland & Scholes, 1990; Wilson, 1984). BAM seeks to expose, define, and potentially innovate or reinvent an organization's business model and then use this analysis to systematically identify key leverage points for the deployment of analytics. Thus, our aim is to develop a BAM that will connect analytics with an organization's ongoing thinking regarding purpose, strategy and core activities and thus ultimately to help an organization to create business value.

The structure of the paper is as follows: in the next section we review the literature and develop the BAM framework. In the third section the research methodology and the action research setting are described. The results of the case intervention are described in section four and the contribution and implications of the work discussed in section five. A summary of the paper is given in the final section.

2. Theoretical development and approach

2.1. Business analytics methodologies

While methodologies are commonplace in information systems development, ranging from the software-focused (e.g., agile software development (Highsmith & Cockburn, 2001)) to the organizational (e.g., Multiview (Avison & Wood-Harper, 1990)) they appear to be less prevalent in business analytics and data science. Searching the literature resulted in remarkably little on business analytics methodologies and data science methodologies that addressed the organizational context. However, one exception is the area of data mining. A poll of 200 users of the KDNuggets Web site in 2014 (Piatetsky, 2014) asked "What main methodology are you using for your analytics, data mining, or data science projects" and reported that 43% (42%) use CRISP-DM, 27.5% (19%) use their own methodology, 8.5% (13%) use SAS's SEMMA (Sample, Explore, Modify, Model, Assess) and 7.5% (7.3%) use KDD (Knowledge Discovery in Databases). The equivalent 2007 percentages are shown in parentheses. The remaining responses (covering 13.5% of respondents) include categories such as in-house methodology, non-domain specific approaches, and no methodology.

The CRoss-Industry Standard Process for Data Mining (CRISP-DM) (Chapman et al., 2000) reference model (Fig. 1) consists of six phases. The arrows show the most important dependencies between stages (although this sequence is not fixed) and the outer cycle reflects the ongoing nature of data mining work. The business understanding phase is concerned with the project objectives and business requirements, which are then converted into a data mining problem definition and project plan. The data understanding phase is concerned with becoming familiar with the data, identifying data quality problems, discovering initial insights and finding interesting areas for making hypotheses. These two phases are reciprocally linked.

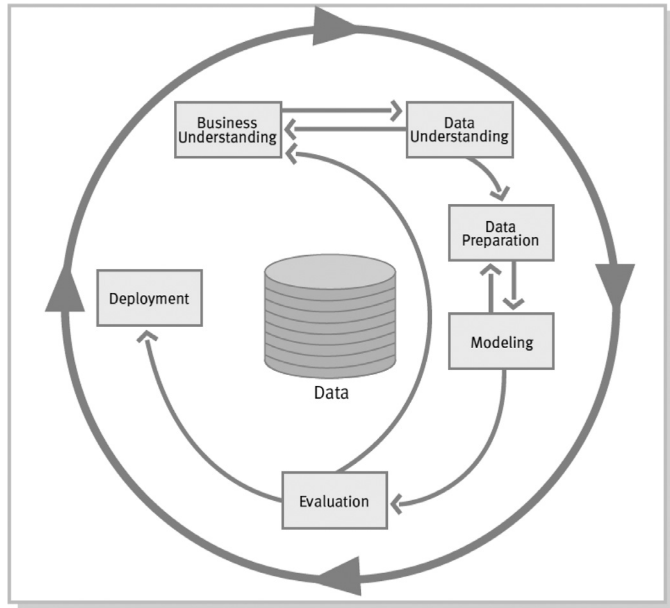

Fig. 1. Phases of the CRISP-DM reference model (Chapman et al., 2000).

The SEMMA process (Azevedo & Santos, 2008) was developed by the SAS Institute. The acronym SEMMA (Sample, Explore, Modify, Model, Assess) covers the steps involved in a data mining project. Similarly, the KDD (Knowledge Discovery in Databases) process, as presented in Fayyad et al. (1996), consists of five stages: Selection; Pre-processing; Transformation; Data Mining; Interpretation/Evaluation. The input to the KDD process is data and the output is knowledge.

The KDD and SEMMA approaches are primarily data-driven and neither gives substantial attention to business context and business objectives. The CRISP-DM process takes greater account of the business context, breaking the business understanding phase into four tasks: determine business objectives, assess situation, determine data mining goals, and produce project plan. The CRISP-DM process model suggests that business objectives are couched in terms of business goals (e.g., to retain customers) that can be couched as business questions (e.g., will lower transaction fees reduce the number of customers who leave?). CRISP-DM advises that the outcomes from a data mining project should be assessed in business terms, ranging from the relatively objective (e.g., reduction in customer churn) to the more subjective (e.g., to give rich insight into customer relationships).

It is clear from the CRISP-DM process that identifying business goals is viewed as an essential aspect of projects that might be labeled 'data mining'. This view is further supported by Khabaza (2010), who proposes nine laws of data mining. Rule 1 (Business Goals Law) argues:

> "… data mining is concerned with solving business problems and achieving business goals. Data mining is not primarily a technology; it is a process, which has one or more business objectives at its heart. Without a business objective (whether or not this is articulated), there is no data mining."