Decision Support Systems 97 (2017) 12–22

Contents lists available at ScienceDirect

Decision Support Systems

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

Bridging the gap between decision-making and emerging big data sources: An application of a model-based framework to disaster management in Brazil

Flávio E.A. Horita[a,*], João Porto de Albuquerque[a,b,e], Victor Marchezini[d], Eduardo M. Mendiondo[c,d]

[a] Institute of Mathematics and Computer Science, University of São Paulo, ICMC-USP, São Carlos, Brazil
[b] Centre for Interdisciplinary Methodologies, University of Warwick, Coventry, United Kingdom
[c] São Carlos School of Engineering, University of São Paulo, EESC-USP, São Carlos, Brazil
[d] Brazilian National Center of Monitoring and Early Warning of Natural Disasters, Ministry of Science, Techn., Innovation, & Communication, Cemaden-MCTIC, São José dos Campos, SP, Brazil
[e] GI Science Group, Department of Geography, Heidelberg University, Germany

ARTICLE INFO

Article history:
Received 11 August 2016
Received in revised form 22 February 2017
Accepted 6 March 2017
Available online 9 March 2017

Keywords:
Business process model and notation
Decision model and notation
Big data sources
Design science research
Case study
Disaster management

ABSTRACT

With the emergence of big data and new data sources, a challenge posed to today's organizations consists of identifying how to align their decision-making and organizational processes to data that could help them make better-informed decisions. This paper presents a study in the context of disaster management in Brazil that applies oDMN+, a framework that connects decision-making with data sources through an extended modeling notation and a modeling process. The study results revealed that the framework is an effective approach for improving the understanding of how to leverage big data in the organization's decision-making.

1. Introduction

With the widespread adoption of smartphones, social media platforms, and wearable technologies, there has been not only an increase in the amount of available data but also a proliferation of new data sources. All of these "big data" have the potential to transform the entire business process [10,35], as well as to provide greater support for decision-making in different contexts, such as business management and marketing [16,36]. However, a remaining challenge lies on how to align decision-making within the organization with the data sources, e.g., how to determine, which data sources could provide useful information for assessing market trends? The reason for this challenge is that despite the fact that the available data could be of great value in supporting decision-making, they often fail to reach the decision-makers in a suitable way [34]. As a result, decision-makers are supplied with useless information that still requires extended knowledge or experience for further data processing. This also makes it difficult to predict the impact that a change of data availability may have on specific tasks, since it is virtually impossible to find out if and where there is a lack of information.

In light of this challenge, this paper proposes the following research question: "How can the decision-makers' tasks be connected to emerging big data sources?". In order to investigate this question in a practical scenario, we perform a study based on the context of disaster management in the Brazilian National Center for Monitoring and Early Warning of Natural Disasters (Cemaden) (http://www.cemaden.gov.br). Cemaden has the mandate of monitoring disasters across the entirety of Brazil's continental territory with 8.5 million km^2. This is thus a notable scenario of decision-making within a "big data" context, since Cemaden must cope with datasets characterized by volume, velocity, variety and veracity. It must process a considerable *volume* of data, since it monitors around 1000 municipalities with recurrent disaster problems using more than 4750 rainfall gauges, about of 550 humidity and rainfall sensors, 9 weather radars, and almost 300 hydrological stations. Data from

* Corresponding author.
E-mail address: horita@icmc.usp.br (F.E. Horita).

http://dx.doi.org/10.1016/j.dss.2017.03.001

these sources have different *velocities*, such as near real-time (in-situ sensors with a reading in each 5 min), every 12 h(weather radars), and unsystematic (semi-automated rainfall gauges). Furthermore, they are provided in a *variety* of formats such as numerical observations, satellite images, and unstructured text. Furthermore, this different sources are disconnected and prone to errors of different types (e.g., sensor errors) and inconsistencies, requiring the analysis of their *veracity*. Additionally, the monitoring natural hazards is a scenario characterized by frequent changes in the information requirements (e.g., new sensors may become available, or existing sources are rendered ineffective), organizational complexity (e.g., it involves several organizational roles, decisions and activities) and tight time pressure — decisions made may have significant consequences to mitigate material and human losses. In summary, these features amount to a highly relevant case for the study of the challenges and opportunities for integrating big data sources into decision-making.

The existing standards and notations from business modeling (i.e., BPMN and DMN) could potentially be used in this context. However, these fail to clearly represent the connection between the information required to make a decision and the data sources. To address this need, we have supplemented the existing standards and notations (i.e., BPMN and DMN) by conducting a design science research to develop and evaluate the oDMN$^+$ Framework for connecting decision-making to useful data sources. Based on the lessons learned from the case study within Cemaden, this framework extends our previous work [13] – which has presented the observation-aware Decision Model and Notation (oDMN) – with the aim of improving the knowledge basis and thus supporting the generalization.

Therefore, this article makes the following key contributions:

- *Case study*: Lessons were learned from a case study carried out in the Cemaden, which is a notable scenario of decision-making within a "big data" context.
- *Extended model and notation*: The oDMN of Horita et al. [13] is extended here by simplifying the notation and using a layer-based structure that makes it easier to understand all the elements. The extended model is now called oDMN$^+$ which is an improved generalization based on the lessons learned from the case study.
- *Modeling process*: The modeling process is proposed to systematically employ the oDMN$^+$ for the case study. This fills a gap both in the modeling practice and in literature, which is lacking clear procedures and guidelines about how to model the conceptual elements of decision-making (e.g., activities) within a specific application context.

This paper is structured as follows. Section 2 outlines the related literature while Section 3 describes the research methodology employed in this research that forms the basis of the oDMN$^+$ Framework set out in Section 4. Section 5 describes the case study used for evaluating the framework and Section 6 discusses the lessons learned from this case study and contributions achieved in this research. Finally, Section 7 summarizes our conclusions and makes recommendations for future work.

2. Literature review

2.1. Connection of decision-making with data sources

Although the use of data for supporting decision-making might be regarded as an established area in the literature, it has attracted further attention in the past few years due to the growing interest in data science, and thus more research has been carried out in this area [7]. Another group of studies seeks to analyze the use of different data sources (e.g., the external data server and sensors) to support decision-making in different scenarios [10,32,35]. Within this group, Vera-Baquero et al. [33] outlined an architecture for integrating data from different organizations, which included a data server for improving the analysis of business performance management. Social media messages have also been employed to support organizational tasks like marketing trends [18,19], or disaster management [34]. For instance, Mandviwalla and Watson [20] described an organization as a mix of capitals (human, economic, social, symbolic, and organizational), which is generated through a social media strategy. Moreover, Kleindienst et al. [15] integrated social media analytics and the business goals of an organization by breaking down these goals into critical success factors that make it possible to find out the information requirements. Although these works addressed an important issue, they failed to provide a method for establishing a connection with the information needed for the decision-making, as well as for providing a representative model that describes this connection.

2.2. Modeling the connection of decision-making with data sources

Business processes can be defined as "a chain of functionally connected activities using information and communication technologies, which lead to a closed outcome providing a measurable benefit for a customer" [22]. The Business Process Model and Notation (BPMN) is a standard model and notation that defines a set of conceptual elements for modeling these processes. BPMN has been applied in different areas, such as customer services [28] and business management [8]. Sackmann et al. [27] proposed an extension to BPMN for including elements to represent place-related information, such as water hydrants or ambulances. Although process modeling notations are valuable to represent organizational activities that involve decision-making, they do not include an explicit consideration of the decisions involved.

Decision Model and Notation (DMN) overcomes this gap by providing conceptual features (e.g., business rules and required information) for modeling decisions, and thus establishing a relationship between business processes and decision-making [23]. Some first examples of studies that have employed DMN can be found in the literature [2,14]. DMN and related studies are certainly an important step for providing a further understanding of decision-making in organizations. However, it does not take into account the kind of data sources that could provide the modeled information requirements to support decision-making. This problem can be partially solved by resorting to the Observations and Measurements (O&M) standard. In short, O&M provides an abstract view of observations that originate from various data sources, as well as being able to integrate the data sources to the requirements of the information [24]. The use of BPMN, DMN and O&M presents a relevant alternative for connecting the business process of an organization to data sources, as showed in our previous work [13]. However, they do not describe a process for obtaining information about conceptual elements from decision-makers of the application context.

2.3. Systematic modeling of the connection of decision-making with data sources

The process used for eliciting the conceptual elements from decision-makers and generating model diagrams of their business processes play an essential role in business process management. These modeling processes are particularly important since many organizations lack a clear understanding of the details (activities, sequence, decisions) of their work practices.

Traditionally, the business process analyst adopted an interview-based approach for extracting data from decision-makers and modeling their business processes [1,3]. Other works proposed the use of