On the Fragmentation of Process Information: Challenges, Solutions, and Outlook

Han van der Aa1, Henrik Leopold1, Felix Mannhardt2,3, and Hajo A. Reijers1

1 Department of Computer Science, Faculty of Sciences, VU University Amsterdam, De Boelelaan 1081, 1081HV Amsterdam, The Netherlands
j.h.vander.aa@vu.nl

2 Department of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600MB Eindhoven, The Netherlands

3 Perceptive Software, Gooimeer 12, 1411DE Naarden, The Netherlands

Abstract. An organization’s knowledge on its business processes represents valuable corporate knowledge because it can be used to enhance the performance of these processes. In many organizations, documentation of process knowledge is scattered around various process information sources. Such information fragmentation poses considerable problems if, for example, stakeholders wish to develop a comprehensive understanding of their operations. The existence of efficient techniques to combine and integrate process information from different sources can therefore provide much value to an organization. In this work, we identify the general challenges that must be overcome to develop such techniques. This paper illustrates how these challenges should be and, to some extent, are being met in research. Based on these insights, we present three main frontiers that must be further expanded to successfully counter the fragmentation of process information in organizations.

1 Introduction

Corporate knowledge is recognized as a principle source of sustainable competitive advantage and value creation [3]. Process knowledge is a particular type of corporate knowledge that relates to processes, their context, and their execution. It is regarded as valuable corporate knowledge, because it can be used to enhance the performance of business processes and, hence, also of organizations [23]. In order to meet this value proposition, it is vital that process knowledge is available when needed. A significant threat to this required availability is that process knowledge is often fragmented throughout an organization. Some is only available in the form of tacit knowledge held by specialized process participants or domain experts [13]. Other knowledge is explicitly captured as process information, but scattered over a plethora of sources, such as documents, models, and systems. This paper focuses on the latter category of process knowledge.

© Springer International Publishing Switzerland 2015
DOI: 10.1007/978-3-319-19237-6_1
The problems resulting from the *fragmentation of process information* are considerable. First, information contained in different sources potentially contradicts each other. This can result, for instance, in situations where different stakeholders hold different expectations on what a process aims to establish and take mutually counter-productive measures. Second, even if sources do not contradict each other, insights from multiple sources must be combined to obtain a complete understanding of any given process. This may be quite tedious, depending on the types of sources that are attempted to be combined. Imagine, for example, the task of combining the insights for the same process from its work instruction, a listing of business rules, and a graphical model. So-called *process information defragmentation techniques* counter these problems. They automatically detect and resolve conflicts between various sources, and also integrate information from different sources in order to provide more comprehensive process insights.

At this point, we argue that the development of process information defragmentation techniques takes place in a haphazard way. A number of sources, such as *event logs*, are being studied and harvested intensively; others are virtually being ignored, such as *policies* and *spreadsheets*. Moreover, researchers that look into different information sources are hardly aware of each other’s work, while that would actually be highly beneficial to further the defragmentation process. A more thorough understanding of the challenges that are involved with the development of process information defragmentation techniques as well as the opportunities seem called for. Hence, it is the goal of this paper to present a systematic view on this field. We achieve this through the identification of the major challenges that process information defragmentation techniques must meet, a description of the state of the art, and a way forward for the advancement of further techniques.

The remainder of this paper is organized as follows. In Section 2, we first consider why information fragmentation exists and the problems that it causes for organizations. Section 3 then elaborates on the types of heterogeneity across information sources that defragmentation techniques must deal with. In Section 4, we illustrate existing approaches that consider these challenges in order to present insights into the current state of the field. Section 5 discusses the main shortcomings of the existing methods and considers future research directions. Finally, Section 6 concludes the paper.

### 2 Fragmented Information

The fragmentation of process-related information poses considerable problems for organizations, yet its causes are well-understood. In Section 2.1, we consider the factors that drive information fragmentation. Because it may be difficult or even undesirable to root out these factors, we argue that that the better approach is to repair the resulting defragmentation with efficient techniques. We identify two main streams of such techniques in Section 2.2.