On Handling Process Information: Results from Case Studies and a Survey*

Bernd Michelberger[1], Bela Mutschler[1], and Manfred Reichert[2]

[1] University of Applied Sciences Ravensburg-Weingarten, Germany
{bernd.michelberger,bela.mutschler}@hs-weingarten.de
[2] Institute of Databases and Information Systems, University of Ulm, Germany
manfred.reichert@uni-ulm.de

Abstract. An increasing data overload makes it difficult to provide the needed information to knowledge-workers and decision-makers in today's process-oriented enterprises. The main problem is to identify the information being relevant in a given process context. Moreover, there are new ways of collaboration in the context of distributed processes (e.g., automotive engineering, patient treatment). The goal is to provide the right process information, in the right format and granularity, at the right place, at the right point in time to the right people. Picking up this goal, enterprises crave for an intelligent and process-oriented information logistics. In this paper we investigate fundamental issues enabling such information logistics based on two exploratory case studies in the automotive and the clinical domain. Additionally, we present results of an online survey with 219 participants supporting our case study findings. Our research does not only reveal different types of process information, but also allows for the derivation of factors determining its relevance. Understanding these factors, in turn, is a fundamental prerequisite to realize effective process-oriented information logistics.

Key words: information logistics, process information, empirical study

1 Introduction

Market globalization has led to increasing competitive pressure for enterprises. Products and services must be developed in ever-shorter cycles. New forms of collaboration within enterprises and between organizations are continuously emerging. As examples consider distributed engineering processes in the automotive domain [1] or the treatment of patients in healthcare networks [2]. To cope with these challenges, effective *business process management* (BPM) [3] becomes more and more success-critical for enterprises.

So far, supporting business processes through information technology has focused on modeling, analyzing, and executing processes (e.g., using BPM technology) [4]. What has been neglected so far is the support of knowledge-workers

* This research was performed in the niPRO project. This project is funded by the German Federal Ministry of Education and Research (BMBF) under grant number 17102X10. More information can be found at http://www.nipro-project.org.

and decision-makers with personalized and contextualized process information. More attention to this challenge will become necessary, however, as an extensive amount of process information is exchanged within enterprises and between organizations using techniques and tools such as e-mail, Web 2.0 applications or enterprise information systems (e.g., enterprise resource planning (ERP) systems, customer relationship management (CRM) systems) [5].

More specifically, we apply the definition of information by Bocij et. al [6] as well as Rainer and Turban [7], and define the term *process information* as follows: *process information* refers to data that have been processed to support process users in the modeling, execution, monitoring, optimization, and design of processes, so that data has a meaning and a value with respect to the process users' activities. Examples of process information include textual process descriptions, working guidelines, graphical processes models, operational instructions, forms, checklists, lessons learned, and best practices (documented in text documents, spreadsheets, or e-mails).

Note that the mere availability of process information is not sufficient to adequately support knowledge-workers and decision-makers as requested above. Only when considering a user's process context it becomes possible to effectively provide personalized and contextualized process information. In practice, many problems arise in this context, e.g., revision control of process information, archiving of process information, inter-departmental exchange of process information, and handling of distributed process information. Further, inconsistencies (schematic and semantic) occur and an increased communication overhead can be observed due to the different structures of digital and paper-based process information. *Process-oriented information logistics* [8] can help to overcome these issues and to effectively manage and distribute process information.

Following these considerations, we investigate the handling of process information in enterprises based on three empirical studies. Thereby, our research has been guided by the following three research questions:

- **RQ1**: In what different forms is process information specified?
- **RQ2**: How can a process context be determined?
- **RQ3**: How can the relevance of process information be determined?

RQ1 and RQ2 are addressed by means of two exploratory case studies. Based on an online survey we answer RQ3 and further concretize RQ1.

The presented research is performed in the niPRO project. In this project we apply semantic technology to integrate both unstructured and structured process information within intelligent, user-adequate *process information portals*. The overall goal is to support knowledge-workers and decision-makers with the needed process information depending on their preferences and current process context. So far, both research and practice have not addressed how processes and related process information can be effectively merged. Currently, conventional methods of information retrieval or enterprise search engines are mainly used for this purpose. The niPRO process information portal, by contrast, aims at determining required information for knowledge-workers and decision-makers

dynamically and automatically. Key challenges include the role-oriented provision of process information, a flexible visualization of process information, and the design of innovative approaches for different levels of information granularity.

This paper is organized as follows. Section 2 presents the results of our empirical studies. Section 3 discusses related work. Section 4 concludes the paper with a summary and an outlook.

2 Empirical Studies

Our empirical studies comprise two parts (cf. Figure 1). First, we performed two qualitative exploratory case studies based on face-to-face interviews and questionnaires. Second, we conducted an online survey to collect further data.

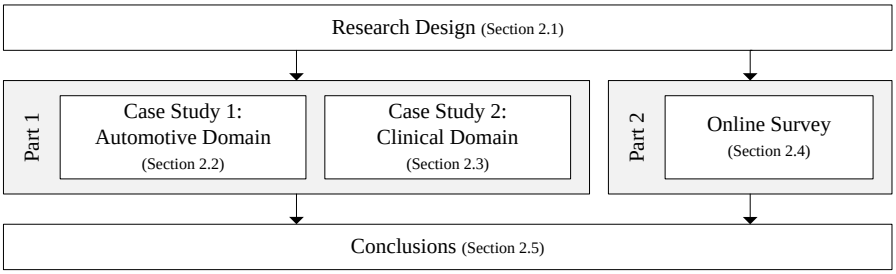

Fig. 1. Our empirical studies.

2.1 Research Design

Our case study research is of explorative nature. According to Yin [9] case studies are a research method to answer *why* and *how* research questions. Kitchenham et. al [10] add to this statement that case studies usually investigate what is happening in "typical" project settings, so it is *research-in-the-typical*. For *research-in-the-large*, i.e., to capture what is happening broadly over large groups, surveys are used. We therefore additionally conduct an online survey to collect further data that helps us to generalize our case study results.

Part 1: Case Studies Two organizations are involved: one from the automotive domain (cf. Section 2.2) and one from the clinical sector (cf. Section 2.3).

In the first case study eight persons have been interviewed, nine in the second one. The interviewees work in different areas of their respective organizations. Both knowledge-workers and decision-makers are involved. Participants were selected in consultation with contact persons from each organization. None of the participants was a member of the research team.

The interviews addressed three major topics: (a) the *processes* in which the interviewees participate, (b) the *types of process information* (RQ1) needed, and (c) the factors determining a *process context* (RQ2). The interviews were conducted in November 2010. Each of them lasted about 90 minutes.

In both case studies, data was gathered through face-to-face interviews following a semi-structured interview guideline. After each interview, an additional questionnaire had to be filled out by each interviewee to collect further data.

Part 2: Online Survey The online survey was conducted via a web questionnaire (cf. Section 2.4). The survey was accessible from mid-December 2010 to late January 2011. 219 users from different enterprises participated. The online survey was advertised via private contacts, business contacts, mailing lists, and groups in social platforms (e.g., LinkedIn). The questionnaire comprised 23 questions on (a) *demographic issues*, (b) *business process management in general*, and (c) *handling of process information* (in order to pick up RQ1 and RQ3).

2.2 Case Study 1: Automotive Domain

Our first case study was conducted in the automotive industry. The participants mainly stem from electric/electronic engineering departments, but also from the departments responsible for project management and safety planning. These departments were selected because of the knowledge-intense business processes they are involved in.

RQ1 Yin [9] states that research questions in case studies are usually too abstract and broad. Therefore, we divide our first research question *"In what different forms is process information specified?"* into three sub-research questions:

- **SRQ1**: Where is process information located?
- **SRQ2**: What are important file formats/systems during daily work?
- **SRQ3**: How is the quality of the available process information?

To answer the first sub-research question we consider the Information Technology (IT) application landscape of the involved departments. The IT application landscape in the automotive industry is extremely complex. There are numerous applications in use providing needed process information. In addition to standard software (e.g., Lotus Notes, RPlan, DOORS) there exists a large number of individual applications (e.g., process portals, Visual Basic for Applications macros etc.). Process information is also available on shared drives, local drives, and in the Internet. Finally, not all process information is available in digital form. Some information is only available in paper form (e.g., technical drawings or circuit diagrams).

Participants confirmed that most process information is available in databases, in applications, in the Internet, and on shared drives. Due to the extensive use of shared drives, a revision control system not officially supported (so far: Subversion, in future: MKS Integrity) is used. The file explorer and the Intranet are

the most common ways to access process information. Information access via applications is not always possible since not all employees typically have needed licenses. Hence, system discontinuities occur, as information is often printed, manually processed, re-entered in enterprise information systems, and further processed.

To answer the second sub-research question we examined file formats and existing information systems. All participants stated that they use Excel files, PowerPoint files, and PDF files. 7 out of 9 participants said that diagrams (e.g., circuit diagrams, technical drawings) are relevant as well. To establish an order of priority, we asked for the three most important file formats during daily work. The result: Excel files, PDF files, and PowerPoint files are most important.

To answer the last sub-research question we take a closer look at the quality of process information. Because the structure and quantity of process information affect its quality [11], we also want to investigate these factors. Most process information is available in unstructured form. However, as unstructured process information is difficult to handle, employees often try to store process information in a structured way (e.g., via templates, databases, applications). In seven of our interviews it was said that the existence of process information is more important than its quality in daily work. However, the interviews also showed that employees often have no overview on available process information due to its large amount; i.e., they often cannot say whether they have all necessary process information. This, in turn, leads to decreased process quality. Not surprisingly, the amount of process information is classified by most participants as too much (cf. Figure 2A). By contrast, the quality of process information was rated differently (cf. Figure 2B). Some process information is rated as being very good (e.g., databases, own documents, information about own tasks). Other information is rated as being very poor (e.g., process documentation, information on tasks).

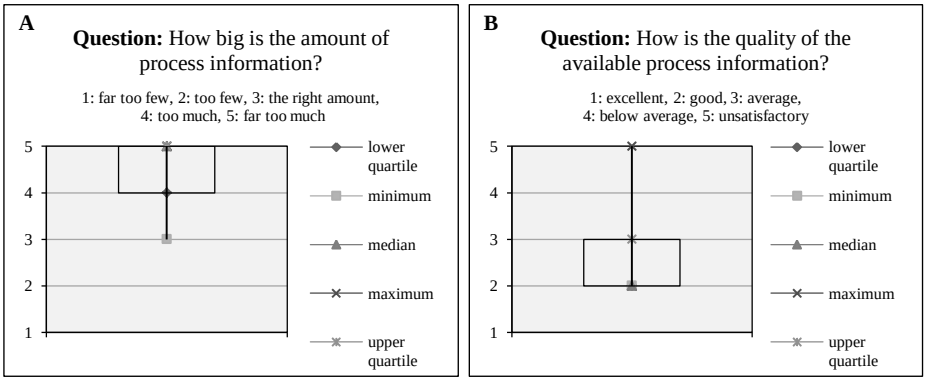

Fig. 2. A) Quality and B) amount of process information.

RQ2 To investigate the employees' process context we ask for factors that can be used to identify a specific process context.

The participants confirmed that the process context is determined based on the progress of a process; specifically by milestones or quality gates (specific milestones) for instance. Some interviewees said that some documents have metadata in which the relation to process steps is noted. Another possibility, also used, is the information progress (e.g., customer data available to 80%). A context, for example, can be determined by folder names because they are often labeled with the name of a respective milestone. Other useful information to determine a specific context is, for example, user names, roles, departments, project memberships, and the time. In summary, there exist various options to determine a context. The more factors are considered, the more accurately a context can be determined.

2.3 Case Study 2: Clinical Domain

In our second case study we considered a process of an unplanned, stationary hospitalization in a surgical clinic. It includes the patient admission, the medical indication in the anesthesia, the surgical intervention, the post-surgery stay on the ward, patient discharge and the financial accounting and management.

RQ1 Like in Case Study 1, we investigate our first research question based on the sub-research questions introduced in Section 2.2.

In the clinical sector, both standard software (e.g., SAP ERP) and individual applications are in use. Clinical staff interacts with them using fat-clients (e.g., DIACOS) and thin-clients (e.g., iMED, CIRS). Process information is available in shared drives, in local drives, in the Internet, in digital archives, and in paper form (e.g., patient files, medical reports, and patient checklists). Our study has revealed that a large amount of process information is not available in electronic form at all. Therefore, exchange of process information between departments is often done manually and only automated to a limited degree. In addition, much process information is available on the Internet, in the Intranet, and in clinical databases or applications (e.g., CIRS). Typically, the processes are not implemented but scattered over multiple more or less integrated systems (e.g., after computer-aided enrollment of patient data via SAP the data is printed and further processed manually by different departments).

To answer sub-research question SRQ2 we analyzed file formats and information systems. 6 out of 8 interviewees confirmed that they mainly use PDF files and Word files. None of the participants uses audio files and only one of them uses video files (medical tutorials). Like in Case Study 1 we asked for the three most important file formats the participants need during daily work. The formats most frequently used are paper-based documents, Word files, and SAP data records.

To answer sub-research question SRQ3, we address quality issues of process information. Like in Case Study 1 we also consider the structure and quantity

of process information. Analogous to Case Study 1 most process information is only available in unstructured form. Further Case Study 2 shows that daily problems are the poor quality (e.g., poorly maintained data about utilization of hospital beds) and the incompleteness (e.g., not all necessary information is available on the emergency protocol) of process information. Besides, process information is often outdated mostly due to the lack of responsibilities concerning information maintenance. The amount of process information is classified by most interviewees as too low (cf. Figure 3A). Reason is that process information is typically paper-based and only one person at a certain point in time can access this information (e.g., the patient file is needed for preliminary investigations, medical reporting, patient care, medical surgery, accounting, etc.). Quality of process information is rated different (cf. Figure 3B). Finally, self-made process information is ranked higher than third-party process information.

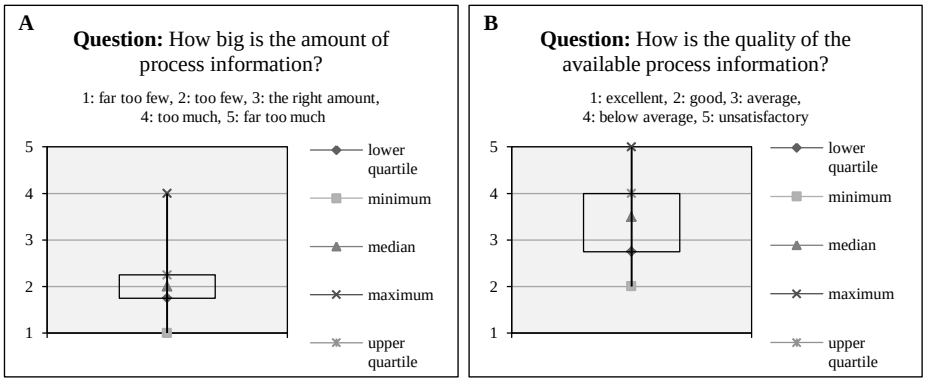

Fig. 3. A) Quality and B) amount of process information.

RQ2 Useful information to determine a context can be time, users or individual computers (because some computers are only used for certain tasks). Also, user location can be helpful, e.g., with the help of mobile devices. Based on the GPS-location of a user, it can be determined, for example, whether a doctor is currently on ward or in the operating theatre. However, 4 out of 8 interviewees believe that is very difficult to determine a process context in healthcare processes. In particular there are no fully pre-specified processes, instead they dynamically evolve and many tasks are performed manually without any IT support. Concerning tasks supported by information systems, the process context can be determined based on information progress (e.g., is a patient ready for accounting or is the patient already settled). Information state changes (e.g., State 1: "patient is in the operating room" or State 2: "patient is on ward")

in information systems also occur and can be used to determine the process context.

2.4 Online Survey

In our online survey, 219 employees from more than 100 enterprises participated. 26% of the participants were decision-makers and 57% were knowledge-workers.

In the first part of the online survey (cf. Section 2.1), we wanted to know whether or not business processes are documented. Obviously, most business processes are fully or partially documented (cf. Figure 4A). Only a small group of participants reported that their work does not take into account business processes or that business processes do only exist in their mind. No one from the production industry reported that processes are undocumented.

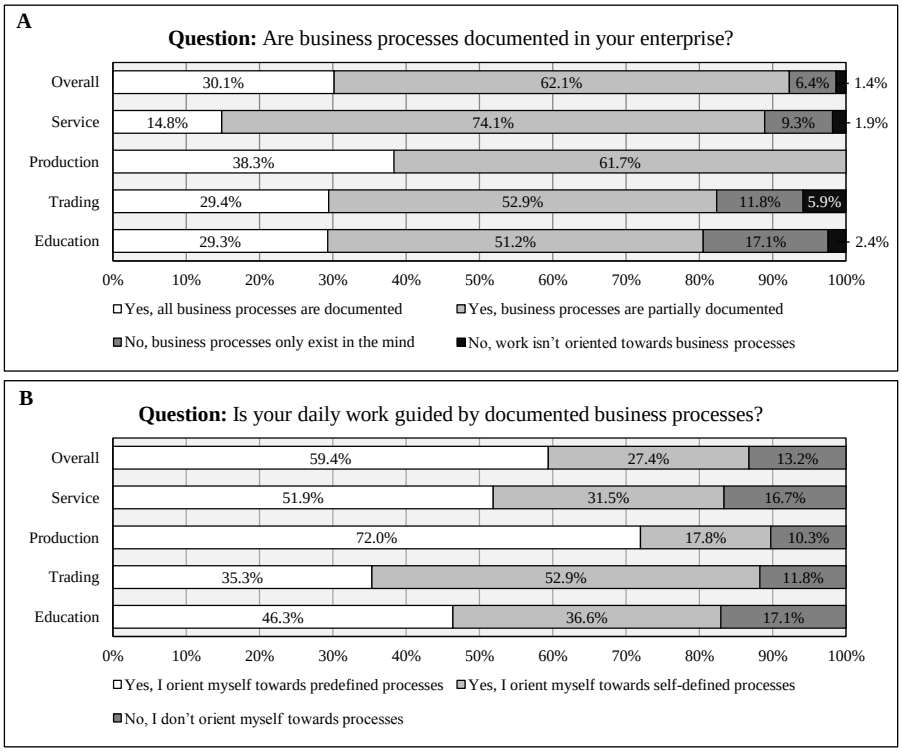

Fig. 4. Documentation of business processes.

We also wanted to know whether the employees' daily work is guided by documented business processes (cf. Figure 4B). More than a half of the respondents stated that they follow predefined business processes. 27.4% of the respondents

follow at least self-defined processes. Only 13.2% of respondents said that they perform their work without considering pre-specified business processes.

Interesting results were also given by means of individual statements of survey participants. Several participants confirmed that people are the most important information source since they can deal with difficult questions or explain other people's work processes. Participants also pointed out that inexperienced staff will benefit most from process information portals. Another participant said that if processes are undocumented, the identification of a process context gains importance.

The first questionnaire block concludes with the question whether an information portal providing needed information could help employees during daily work (cf. Figure 5A). Most of the respondents (85.8%) somewhat or totally agreed. Respondents also agreed that there exists a large number of diverse information sources from which they obtain their information (cf. Figure 5B).

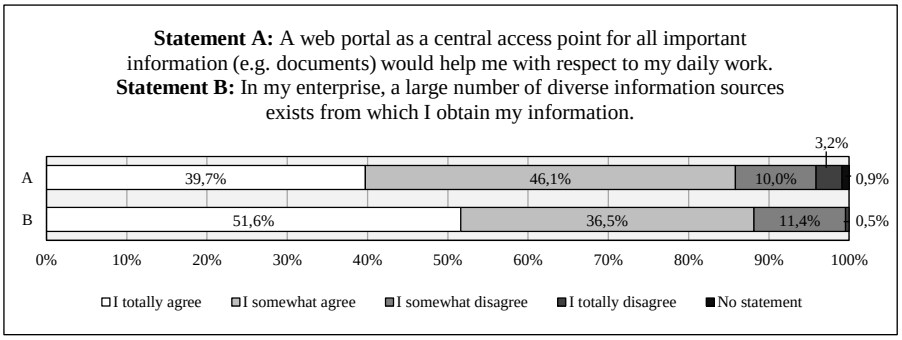

Fig. 5. Information portal and information sources.

RQ1 We asked where needed process information is located. Most participants referred to databases, applications, shared and local drives, and the Internet as the most important sources of process information. When comparing shared and local drives it becomes evident that the majority of process information is stored on shared drives (86%). Furthermore, several participants pointed out that people represent an important source of information as well (e.g., experts, colleagues or hotlines). The most important file formats are PDF, Excel, PowerPoint, and Word.

RQ3 We also investigated the relevance of process information. Many participants stated that self-made process information (e.g., own documents, e-mails) have a greater relevance than third-party information. Our survey results also show a direct relationship between the frequency a particularly information is

accessed and its relevance. The more frequent a particular process information is accessed the higher is its relevance. Hence, participants confirmed that standardized process information (e.g., forms) is more relevant than non-standardized one. In this context participants confirmed that the relevance of process information is significantly influenced by the reliability of the information source. Additionally relevance factors mentioned include the number of changes of an information, the date of the last access, the amount of metadata assigned to a process information, and the file size. As far as temporal consideration is concerned, recurring information and timeliness of process information influences its relevance as well. Most participants stated that only up-to-date and complete process information can be relevant. The accessibility to process information is denoted as a basic requirement.

We also analyzed the available amount of process information (cf. Figure 6). Obviously, decision-makers are confronted with too much information. 45.1% of the decision-makers confirm that they have too much or far too much process information (knowledge-workers: 24%). Knowledge-workers, by contrast, have the problem of being confronted with insufficient information. 48.1% of the knowledge-workers mentioned that they have too little or far too little process information (decision-makers: 27.5%).

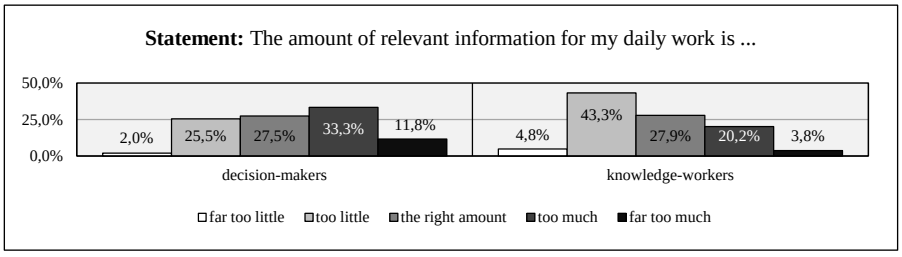

Fig. 6. Amount of process information.

2.5 Conclusions

Regarding research questions RQ1-RQ3 we can draw the following conclusions:

RQ1 *"In what different forms is process information specified?"* The majority of process information in enterprises is available in Word files, Excel files, PowerPoint files, PDF files, and in paper form. In addition, there are many enterprise-specific file formats. The most common information sources are the Internet, shared and local drives, and non-electronic information sources (e.g., documents in paper form). Significant problems in enterprises are the poor quality and timeliness of process information. Finally, access problems (e.g., lack of licenses) to process information sources are reported.

RQ2 *"How can a process context be determined?"* Our results show that a process context can be determined through various factors, e.g., by considering the progress of processes, information progress, data associated with processes (e.g., folder names, metadata), and specific computers for certain tasks. Other useful information to determine a specific context includes, for example, the user name, the role, the department, the project membership, and the time.

RQ3 *"How can the relevance of process information be determined?"* The results show that self-made process information has a much greater relevance than third-party process information. This relevance is affected by many factors: access frequency, standardization, reliability of information sources, number of changes, date of last access, available metadata, and size. Another important factor is the quality of process information, which can be determined based on characteristics such as periodicity, precision, and granularity. Process information must be up-to-date and complete to increase its relevance for employees' daily work.

3 Related Work

There are studies dealing with process-oriented information logistics in enterprises. Dinter and Winter [12] analyze current information logistics strategy practices by means of a survey. Their findings show that information logistics strategy is linked to company size and governance type. Bucher and Dinter [13] provide another empirical analysis to assess benefits, design factors, and realization approaches for process-oriented information logistics. The study of Lahrmann and Stroh [14] identifies typical scenario patterns in information logistics. A case study in a tourism setting is performed by Landqvist and Stenmark [15]. They investigate portal information integration and ownership misfits. Hristidis et. al. [16] conduct a survey about data management and analysis. They achieved the same results as in our study: Data are available in many different formats, have varying characteristics, and stem from different sources.

All these studies analyze process-oriented information logistics with different emphasis (e.g., strategy, design factors, scenarios, and misfits). However, the combination of different types of process information, process context, and process information relevance is addressed by none of them.

4 Summary and Outlook

This paper summarizes the results of two case studies and one online survey. We investigate different types of process information and factors determining the relevance of process information and process context. We further identify the most important characteristics of process information like source, location, and quality of the process information. In addition, we investigate how a specific process context can be determined.

Future research will include additional studies to investigate quality dimensions of process information such as periodicity, granularity, and completeness. These quality dimensions of process information need to be analyzed as they strongly influence overall quality of process information.

References

1. Müller, D., Herbst, J., Hammori, M., Reichert, M.: IT Support for Release Management Processes in the Automotive Industry. in: 4th Int'l. Conference on Business Process Management (BPM'06), pp. 368-377, Vienna (2006)
2. Lenz, R., Reichert, M.: IT Support for Healthcare Processes - Premises, Challenges, Perspectives. in: Data and Knowledge Engineering, 61(1), pp. 39-58 (2007)
3. Weske, M.: Business Process Management: Concepts, Languages, Architectures. Springer (2007)
4. Mutschler, B., Reichert, M., Bumiller, J.: Unleashing the Effectiveness of Process-oriented Information Systems: Problem Analysis, Critical Success Factors and Implications. IEEE Transactions on Systems, Man, and Cybernetics (Part C), 38 (3), pp. 280-291 (2008)
5. Laudon, K., Laudon, J.: Management Information Systems. Prentice Hall (2009)
6. Bocij, P., Chaffey, D., Greasley, A., Hickie, S.: Business Information Systems: Technology, Development and Management for the E-Business. Prentice Hall (2005)
7. Rainer, R.K., Turban, E.: Introduction to Information Systems: Supporting and Transforming Business. John Wiley & Sons (2008)
8. Meissen, U., Pfennigschmidt, S., Voisard, A., Wahnfried, T.: Context- and Situation-Awareness in Information Logistics. in: Current Trends in Database Technology - EDBT 2004 Workshops, pp. 335-344 (2005)
9. Yin, R.K.: Case Study Research: Design and Methods. (2008)
10. Kitchenham, B., Pickard, L., Pfleeger, S.L.: Case Studies for Method and Tool Evaluation. (1995)
11. Wang, R.Y., Strong, D.M.: Beyond Accuracy: What Data Quality Means to Data Consumers. in: J. of Management Information Systems, 12 (4), pp. 5-33 (1996)
12. Dinter, B., Winter, R.: Information Logistics Strategy - Analysis of Current Practices and Proposal of a Framework. in: Proc. of the 42nd Hawaii Int'l. Conference on System Sciences (HICSS-42), Hawaii (2009)
13. Bucher, T., Dinter, B.: Process Orientation of Information Logistics - An Empirical Analysis to Assess Benefits, Design Factors, and Realization Approaches. in: Proc. of the 41st Annual Hawaii Int'l. Conference on System Sciences (2008)
14. Lahrmann, G., Stroh, F.: Towards a Classification of Information Logistics Scenarios - An Exploratory Analysis. in: Proc. of the 42nd Hawaii Int'l. Conference on System Sciences (2009)
15. Landqvist, F., Stenmark, D.: Portal Information Integration and Ownership misfits: A Case Study in a Tourism Setting. in: Proc. HICSS '06 Proceedings of the 39th Annual Hawaii Int'l. Conference on System Sciences (2009)
16. Hristidis, V., Chen, S.C., Li, T., Luis, S., Deng, Y.: Survey of data management and analysis in disaster situations. in: J. of Systems and Software, 83 (10), pp. 1701-1714 (2010)