BPM-in-the-Large – Towards a Higher Level of Abstraction in Business Process Management

Constantin Houy[1], Peter Fettke[1], Peter Loos[1],
Wil M.P. van der Aalst[2], and John Krogstie[3]

[1] Institute for Information Systems
at the German Research Center for Artificial Intelligence (DFKI)
Stuhlsatzenhausweg 3, Geb. D3 2
66123 Saarbrücken, Germany
`{Constantin.Houy,Peter.Fettke,Peter.Loos}@iwi.dfki.de`
[2] Department of Mathematics and Computer Science
Technische Universiteit Eindhoven (TUE)
Den Dolech 2
5612 AZ Eindhoven, Netherlands
`w.m.p.v.d.aalst@tue.nl`
[3] Department of Computer and Information Science
Norges Teknisk-Naturvitenskapelige Universitet (NTNU)
Hogskoleringen 1
7491 Trondheim, Norway
`krogstie@idi.ntnu.no`

Abstract. Business Process Management (BPM) has gained tremendous importance in recent years and BPM technologies and techniques are widely applied in practice. Furthermore there is a growing and very active research community looking at process modeling and analysis, reference models, workflow flexibility, process mining and process-centric Service-Oriented Architectures (SOA). However, it is clear that existing approaches have problems dealing with the enormous challenges real-life BPM projects are facing. Large organizations have hundreds of processes in place. These processes are often poorly documented and the relationships between them are not made explicit. Conventional BPM research seems to focus on situations with just a few isolated processes while in reality the real challenge is to cope with large collections of interconnected processes. Moreover, new technologies such as ubiquitous computing (sensor technologies, mobile devices, RFID tagging etc.) and pervasive networks (cf. "internet of things") generate enormous volumes of event data. Organizations have problems handling and using such data. Event data is scattered over various subsystems and not used well. This paper coins the term "BPM-in-the-Large" to describe the above situation and describes challenges and opportunities for BPM research.

Keywords: Business Process Management, BPM-in-the-Large, abstraction, complexity, Programming-in-the-Large.

1   Introduction

Business Process Management (BPM) has gained tremendous importance in the last decade and is increasingly used by organizations around the world [1]. The usage of information and communication technology (ICT) in order to manage business processes in companies and administrations has gained more and more importance in recent years [2]. BPM is characterized by a cross-fertilization of results from management science and computer science. Adequate techniques and software tools supporting the design, enactment, control as well as the analysis of operational business processes are already applied in businesses in order to facilitate an optimized value creation [3].

However, it is clear that existing approaches have problems dealing with the enormous challenges real-life BPM projects are facing. Large organizations have hundreds of processes in place. These processes are often poorly documented and the relationships between them are not made explicit. The amount of different process types which are characteristic for large enterprises are documented by well-established reference process models like Scheer's Reference Model for Industrial Enterprises [4] as well as the SAP reference model. As comprehensive and inter-corporate value chain networks keep gaining importance and as new possibilities of managing business processes have emerged on the basis of newer technologies – like mobile communication or ubiquitous computing technologies – the complexity of business processes has increased in recent years. Conventional BPM research seems to focus on situations with just a few isolated processes while in reality the real challenge is to cope with large collections of interconnected processes. In the future existing techniques and tools for BPM will not suffice to handle this ever increasing complexity.

In the following we argue for the need of new methods and technologies in BPM for the handling of the complexity mentioned ("BPM-in-the-Large"). A first overview of challenges and opportunities will be given based on conceptual consideration. Our contribution is structured as follows: In Section 2 the underlying understanding of BPM is introduced. Section 3 expatiates upon the reasons for the increasing complexity in BPM. Section 4 introduces BPM-in-the-Large as a new level of abstraction before Section 5 presents resulting challenges as well as the opportunities of BPM research. Section 6 focuses on the role of process mining in BPM-in-the-Large. Section 7 concludes the paper.

2   Business Process Management Framework

In general BPM comprises a management principle which companies apply in order to sustain their competitive advantage [5]. It focuses on business processes as a sequence of executions in a business context based on the purpose of creating goods and services [6]. There are two basic approaches to BPM and process improvement: (1) Business Process Reengineering as a radical redesign of business processes by a singular transformation [7] and (2) evolutionary improvement of business processes by continuous transformation. Today the latter is certainly of more importance for practical BPM efforts [8]. This continuous improvement approach is typically
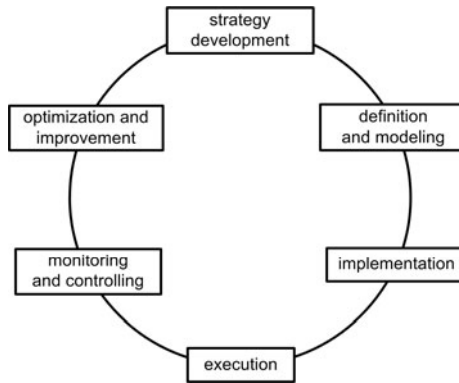
Fig. 1. BPM Life Cycle for continuous process improvement

conceptualized by a BPM Life Cycle [9] which will be used as a reference framework for the description of opportunities of BPM-in-the-Large.

3   Reasons for Increased Complexity

As emerging technologies permit new ICT system architectures, new and promising opportunities in the context of BPM arise. Based on globalization trends, new challenges pop up, particularly when multi-national companies need to coordinate their local business units in order to serve other multi-national companies in an integrated fashion. In an earlier reported case of a certification company [10, 11], there was a need to standardise the processes of the company's national branches in order to build a common image of the organization, and to support the certification of the cross-national processes of their multi-national customers, but at the same time adhere to national and cultural rules and expectations.

In such a context the integration of common technologies like mobile devices, techniques from the ubiquitous computing context, so called smart environments as well as the increasing use of sensor network technologies for the collection of process-relevant data and the application of service-oriented architectures (SOA) as well as Web 2.0 technologies can improve the flexibility of inter-corporate BPM [12] and thus increase the effectiveness and efficiency of business processes in inter-corporate value chain networks. Furthermore the options of action for human actors involved increase. Figure 2 illustrates a collaborative scenario in a value chain network applying the technologies mentioned. Scenarios like these are not only of importance in business, but also in the public administration area as described in the last EU Ministerial Declaration on eGovernment [13] which emphasizes the special need to develop and improve cross-border eGovernment services, making it easier for businesses and citizens to operate in and across any EU-member state.
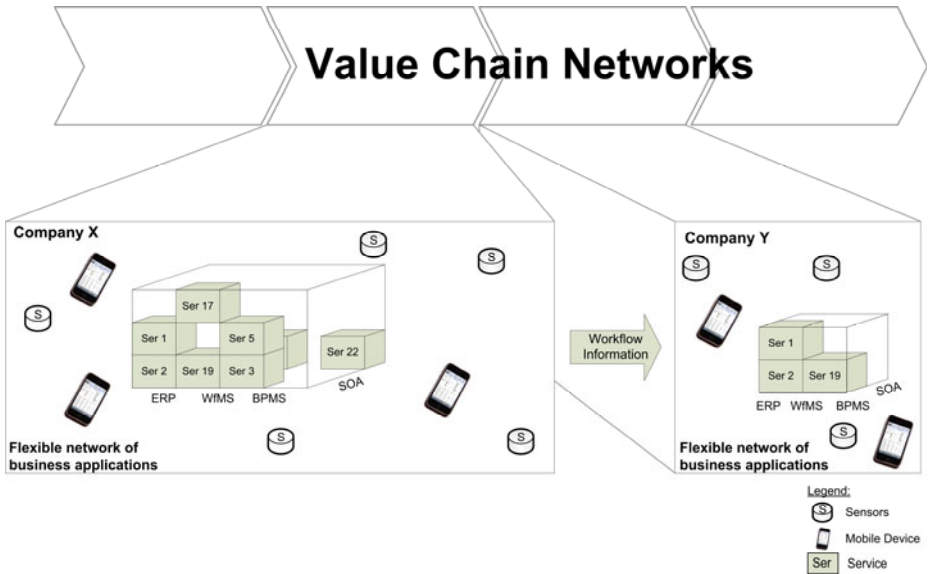
Fig. 2. Collaborative scenario in Value Chain Networks

The above scenario shows four important *trends*:

1. processes are *increasingly interconnected* and it makes no sense to look at a single process in isolation,
2. the *number of processes* an organization has to cope with *is rapidly increasing* (large organizations have hundreds of processes which need to be managed),
3. modern technology is generating unprecedented streams of event data representing the states of different processes (sensor data, RFID data, remote logging, remote services etc.) and
4. different devices are used to access the BPM System (BPMS) in different situations necessitating a flexible multi-channel support influencing which parts of the workflow are available in which manner depending on the context of use.

Based on these trends and the application of the mentioned technologies on the one hand the enterprises' agility and the handling of more and more dynamic business conditions can be improved. On the other hand Business Process Management becomes more and more complex. The *reasons for this complexity* are manifold:

1. the range of inter-corporate collaborative business processes,
2. the number of organizational units involved in a business process,
3. the need to manage and control mobile actors in business processes,
4. the need to control person-machine and machine-machine interactions,
5. the interdependencies in sensor networks, and
6. the need to manage services in a business process applying SOA etc.

The handling of this complexity which is generated by many thousands of process types, process instances and process events is a new challenge in BPM.

4   BPM-in-the-Large as a New Level of Abstraction

The study and development of accurate abstractions of relevant objects is a fundamental issue in computer science. So far business processes have been considered the highest level of abstraction in the context of information systems. This has been considered in the following figure illustrating a taxonomy of levels of abstraction within an information system based on Denning [14]. In Denning's original model business processes represent the highest level of abstraction which has to be managed by information system infrastructures. Still today business processes are being managed by distributed information system infrastructures applying different data processing units ("multimachine level"). Based on the presented facts a new level of abstraction needs to be introduced. The illustration has been expanded by a new level of management and which we entitle "BPM-in-the-Large" according to DeRemer/Kron's "Programming-In-The-Large" [15]. We think that the situation in BPM is quite similar to the situation in programming in 1975. BPM-in-the-Large (as opposed to "BPM-in-the-Small") can be understood by looking at the characteristics mentioned earlier: highly interconnected scenarios, hundreds of processes, large data streams and the usage of diverse devices within the BPMS.
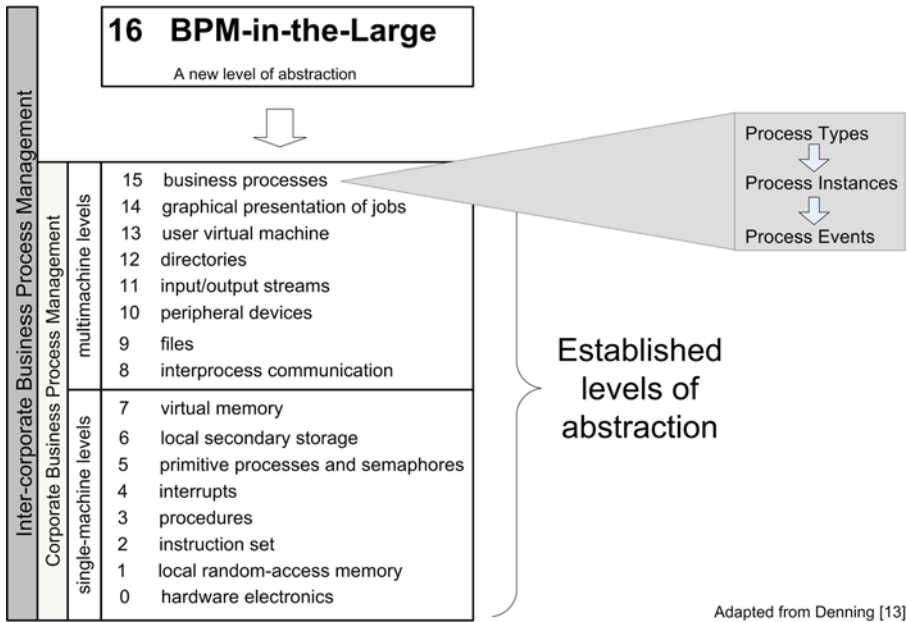

Fig. 3. Levels of abstraction within a business process-oriented information system

5   Challenges and Opportunities of BPM-in-the-Large

Based on the introduced phases of the BPM Life Cycle the various challenges and opportunities of BPM-in-the-Large are presented in the following.

Strategy development
On the level of strategy development BPM-in-the-Large can offer an improved overview and steering of comprehensive value creation networks by providing large scale process models and process-relevant data collected from Enterprise Resource Planning (ERP) Systems, Workflow Management Systems (WfMS) and lower scale BPMS. Important structures of collaborative processes can thus be controlled effectively, an improved development of the further process strategy is rendered possible and the full strategic potentials can be tapped.

The application of analytical concepts based on the data provided by different applications such as business process mining can make an important contribution to the strategy development as well. In connection with sensor networks and mobile devices the integration of analytical concepts like business process mining supports the identification of strategically important process structures from existing data by identifying patterns within this data [16].

Definition and Modeling
Managing business process models is one of the crucial points in BPM and in BPM-in-the-Large as well. Available experience in the application of business process models shows the considerable complexity of the definition and design of original business process models and the maintenance of existing ones [17]. Reasons for this complexity are:

1. the size and extensiveness of applied process models in practice,
2. the collaborative design and maintenance of process models and
3. the different usage areas of models by different stakeholders at different times [11], combined with the manifold of modifications of these models during their life cycle in the context of continuous improvement.

In order to manage the complexity mentioned, BPM-in-the-Large should integrate tools for improved collaborative business process modeling. Furthermore the support of automatic or semi-automatic discovery of complex process models and their modifications during the life cycle from event log files by advanced methods of business process mining [18] can make a substantial contribution to a successful management of business process models.

Another interesting aspect in this context is the retrieval of structural analogies in process models. Structural analogies express similarity of several model constructs and can thus be useful for the reduction of complexity by providing "common parts" of business process models [19]. Thus a once implemented module of a business process model can be used multiple times. Data provided by sensors and mobile devices offers important context information which supports not only the discovery of structural patterns, but also semantic subtleties in this context. Thus the quality of business process models and their management can effectively be enhanced. Capable

methods of assessing and improving the quality of important processes support the entire following BPM Life Cycle and using these methods in BPM-in-the-Large supports a comprehensive and effective handling of business process models.

Implementation

BPM-in-the-Large should enable an optimal implementation of comprehensive inter-corporate business processes. Thus an improved inter-corporate communication is of importance. In this context, Web 2.0 technologies like Wikis, Blogs and Social Bookmarking etc. as an extension of approaches of interactive process modeling [20] can support the improvement of business process-related communication of cooperative partners [12].

Furthermore, according to defined and modelled processes the quality of models and their implementation have to be assured. As a high quality of process models is generally considered to influence the success of a company the improvement of process model quality and the quality of their implementation is of high importance.

Although the quality of models in general can be analysed as the interplay of quality at different levels (physical, empirical, syntactic, semantic, pragmatic, social, organizational [21]), existing approaches for the measurement of process model quality have not been defined consistently [22]. For BPM-in-the-Large the measurement of process model quality is therefore of high importance. Adequate instruments for the measurement of model quality support a distinctive effectiveness of business process models and their implementations on business success. The following questions are of specific importance for process model quality: (1) Does the model describe what was intended (organizational quality)?, (2) Is the model understandable (pragmatic quality)?, (3) Is the model internally consistent (e.g. absence of deadlocks) (an aspect of semantic quality)? and (4) Does the model conform to reality (semantic quality)? The latter point can be addressed using conformance checking techniques that compare the behaviour observed and recorded in event logs with the model [23]. Furthermore, the implementation of process models can be facilitated by the adequate configuration of promising "common parts" extracted in the preceding phase. It has to be noted that the need for federated process models in order to address variants of models across different cultures often makes this a difficult problem. A good sales process in India for instance will be different than the sales process in Europe or US even if the goal of the process is the same [10].

Execution

In the execution context BPM-in-the-Large envisions a wide-ranging workflow functionality which has to be adapted for comprehensive cooperative business processes in order to support an optimal execution of process instances according to the collaboratively defined process models in value creation networks. In this context the interoperability of the workflow management components has to be assured. A realization can be assured by using SOA for the implementation of the workflow components.

Monitoring and Controlling

Also in the context of monitoring and controlling of business processes BPM-in-the-Large provides an enormous potential in comparison to established BPMS. In connection with further information provided by different distributed data sources which

collect process-relevant information, business process instances and anomalies during their execution can be controlled and identified more effectively. Relevant data is provided by the plethora of sensors and mobile devices. These data allow for a more comprehensive monitoring and analysis of business process instances. By utilizing this opportunity, the optimization and improvement phase which addresses process types can be better supported. Using process mining methods in this phase can furthermore support the deduction and construction of reference process models based on successfully executed process instances. Reference models play an important role in the context of successful business engineering [24].

Optimization and Improvement
The actually measured performance of business processes forms the basis for improvement initiatives and for the development of new process structures which can improve the performance of future process instances. In this context the developed quality and performance measures provide the basis for comprehensive and effective comparisons. Moreover, the identification of structural analogies provided by the application of business process mining techniques supports the improvement and the further development of process models applied in the organisation. Figure 4 summarizes these findings.
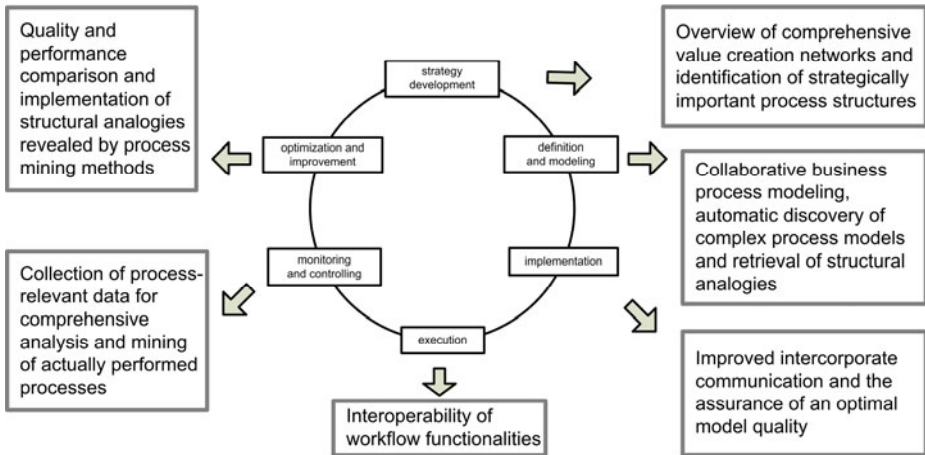

Fig. 4. BPM Life Cycle supported with BPM-in-the-Large

6   Process Mining as a Driver for BPM-in-the-Large

The importance of having short feedback-loops from actual process performance to future process support has been acknowledged for some time in areas such as emergent and interactive process support [21] and active knowledge modeling [25]. However, these techniques are appropriate mostly for knowledge intensive tasks with relatively few processes and few stakeholders involved, with primarily manual support for process knowledge management [26]. Therefore, we consider automated

analysis based on event data. In the previous section there were several references to the use of process mining technology to support the various phases. Therefore, we now describe the state-of-the-art in process mining and position this work in BPM-in-the-Large.

More and more information about (business) processes is recorded by information systems in the form of so-called "event logs". IT systems are becoming more and more intertwined with these processes, resulting in an "explosion" of available data that can be used for analysis purposes. Process mining techniques attempt to extract non-trivial and useful information from event logs. One aspect of process mining is *control-flow discovery*, i.e., automatically constructing a process model (e.g., a Petri net or BPMN model) describing the causal dependencies between activities. The basic idea of control-flow discovery is very simple: given an event log containing a set of traces, these traces are used to automatically construct a suitable process model describing the behavior seen in the log. Such discovered processes have proven to be very useful for the understanding, redesign, and continuous improvement of business processes [27].

Process mining is not limited to control-flow discovery. First of all, besides the *control-flow perspective* ("How?"), other perspectives such as the *organizational perspective* ("Who?") and the *case/data perspective* ("What?") may be considered.

Second, process mining is not restricted to *discovery*. There may be existing models that can be related to the event logs. For example, process mining can be used to analyze deviations from some normative model (e.g. *conformance checking* as discussed in [23]), thus enabling auditing functionality.
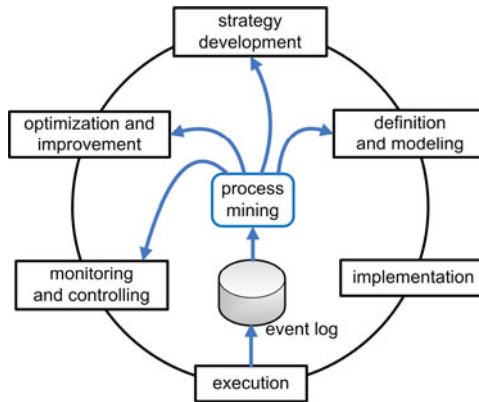

Fig. 5. Process mining positioned in the BPM Life Cycle

Figure 5 shows the central role of process mining in the BPM Life Cycle. It is used as input for four of the six phases. To conclude this section of process mining we focus on *two challenges* related to BPM-in-the-large and process mining.

The first challenge is related to difficulties in making process models which are understandable and correct in the presence of massive amounts of activities and (sub-) processes. Here we make the analogy to cartography. The first geographical maps date back to the 7th Millennium BC. Since then cartographers have improved their

skills and techniques to create maps thereby addressing problems such as clearly representing desired traits, eliminating irrelevant details, reducing complexity, and improving understandability. Today, geographic maps are digital and of high quality. This has fueled innovative applications of cartography as is illustrated by modern car navigation systems (e.g. TomTom, Garmin, etc.), Google maps, Google Street View, Mash ups using geo-tagging, etc. People can seamlessly zoom in and out using the interactive maps in such systems. Moreover, all kinds of information can be projected on these interactive maps (e.g., traffic jams, four-bedroom apartments for sale, etc.). Process models can be seen as the "maps" describing the operational processes of organizations. Unfortunately, *accurate and interactive process maps are typically missing when it comes to BPM*. The challenge is to reach the same level of maturity for "business process maps". This is one of the grand challenges in process mining; how to automatically generate models which are well-understood, correct, and showing dynamic information. Furthermore new ways of visualizing the process knowledge have to be found [28].

The second challenge is related to the input data itself. BPM-in-the-large is characterized by a large number of processes and massive streams of event data. This requires a more systematic approach towards the collection and storage of event data. We use the term *business process provenance* to refer to the systematic collection of the information needed to reconstruct what has actually happened. The term signifies that for process mining it is vital that "history cannot be rewritten or obscured". For example, from an auditing point of view the systematic, reliable, and trustworthy recording of events is essential. Therefore, we propose to collect (whenever possible) provenance data outside of the operational information. This means that events need to be collected and stored persistently. Note that semantics play an important role here, i.e. *events need to refer to a commonly agreed-upon ontology*. Here we propose to use the *Semantically Annotated Mining XML* (SA-MXML) format developed in the context of the SUPER project [29]. This reflects that we would like to treat event data as a "first class citizen" in BPM-in-the-Large.

7   Conclusion

BPM-in-the-Large is of high relevance for both research and practice. The opportunities which can be offered by the implementation of BPM-in-the-Large are of the utmost importance for enterprises which aim at utilizing innovative information technology components and at rationally organizing these components in order to tap their full potential. By the usage of SOA and interchangeable components future BPM-in-the-Large implementations can be highly scalable and can thus be applied optimally in small and medium enterprises (SME) as well as in large multinationals and across governments.

Furthermore BPM-related topics gain more and more importance in IS research. The study of integration possibilities of mentioned technologies in BPM is of a very high interest as the field of research grows steadily and BPM has become a considerable trend in management as well as in computer science. The successful introduction of a higher level of abstraction can be seen as a useful and valuable challenge in BPM research.

BPM technology has been successfully implemented in many different organizations and has shown to directly influence the operational processes. Therefore, BPM-in-the-Large can have an enormous technological and economic impact. The effectiveness and efficiency, and thus the success of business processes will be enhanced by the application of BPM-in-the-Large across various industries. Furthermore, the scientific impact of the presented methods, e.g. systematically measuring process model quality and innovative process mining techniques, can be considerable, as these methods can support further innovation in practical BPM as well as BPM-related research.

References

1. Fettke, P.: How Conceptual Modeling Is Used. Communications of the AIS 25, 571–592 (2009)
2. ter Hofstede, A.H.M., van der Aalst, W.M.P., Adams, M., Russell, N. (eds.): Modern Business Process Automation: YAWL and its Support Environment. Springer, Berlin (2010)
3. van der Aalst, W.M.P., ter Hofstede, A.H.M., Weske, M.: Business Process Management: A Survey. In: van der Aalst, W.M.P., ter Hofstede, A.H.M., Weske, M. (eds.) BPM 2003. LNCS, vol. 2678, pp. 1–12. Springer, Heidelberg (2003)
4. Scheer, A.-W.: Business Process Engineering - Reference Models for Industrial Companies. Springer, Berlin (1994)
5. Hung, R.Y.: Business Process Management as Competitive Advantage: a Review and Empirical Study. Total Quality Management 17, 21–40 (2006)
6. Scheer, A.-W.: ARIS - Business Process Frameworks. Springer, Berlin (1999)
7. Hammer, M., Champy, J.: Reengineering the Corporation: A Manifesto for Business Revolution. Harper Business, New York (1993)
8. Weske, M.: Business Process Management: Concepts, Languages, Architectures. Springer, Berlin (2007)
9. Houy, C., Fettke, P., Loos, P.: Empirical Research in Business Process Management - Analysis of an emerging field of research. Business Process Management Journal 16 (2010) (accepted)
10. Krogstie, J., Dalberg, V., Jensen, S.M.: Harmonising Business Processes of Collaborative Networked Organisations Using Process Modelling. In: Virtual Enterprises and Collaborative Networks. IFIP, vol. 149, pp. 81–88. Springer, Boston (2004)
11. Krogstie, J., Dalberg, V., Jensen, S.M.: Process modeling value framework. In: Manolopoulos, Y., Filipe, J., Constantopoulos, P., Cordeiro, J. (eds.) ICEIS 2006. LNBIP, vol. 3, pp. 309–321. Springer, Heidelberg (2008)
12. Vanderhaeghen, D., Fettke, P., Loos, P.: Organizational and Technological Options for Business Process Management from the Perspective of Web 2.0 - Results of a Design Oriented Research Approach with Particular Consideration of Self-Organization and Collective Intelligence. Business & Information Systems Engineering 2, 15–28 (2010)

13. Ministerial Declaration on eGovernment, Malmø, Sweden, 5th Ministerial eGovernment Conference (2009)
14. Denning, P.J.: Work is a closed-loop process. American Scientist 80, 314–317 (1992)
15. DeRemer, F., Kron, H.: Programming-in-the large versus programming-in-the-small. In: Proceedings of the International Conference on Reliable Software, pp. 114–121. ACM, Los Angeles (1975)
16. Tiwari, A., Turner, C.J., Majeed, B.: A review of business process mining: state-of-the-art and future trends. Business Process Management Journal 14, 5–22 (2008)
17. Cardoso, J., Mendling, J., Neumann, G., Reijers, H.A.: A discourse on complexity of process models. In: Dustdar, S., Fiadeiro, J.L., Sheth, A.P. (eds.) BPM 2006. LNCS, vol. 4102, pp. 117–128. Springer, Heidelberg (2006)
18. van der Aalst, W.M.P., Weijters, A.J.M.M., Maruster, L.: Workflow mining: Discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering 16, 1128–1142 (2004)
19. Fettke, P., Loos, P.: Identifying Structural Analogies in Data Models - A Procedure and its Application Exemplified by Scheer's Y-CIM Reference Model. Wirtschaftsinformatik 47, 89–100 (2005)
20. Krogstie, J., Jørgensen, H.: Interactive Models for Supporting Networked Organisations. In: Persson, A., Stirna, J. (eds.) CAiSE 2004. LNCS, vol. 3084, pp. 550–563. Springer, Heidelberg (2004)
21. Krogstie, J., Sindre, G., Jørgensen, H.: Process Models as Knowledge for Action: A Revised Quality Framework. European Journal of Information Systems 15, 91–102 (2006)
22. Moody, D.L.: Theoretical and Practical Issues in Evaluating the Quality of Conceptual Models: Current State and Future Directions. Data & Knowledge Engineering 55, 243–276 (2005)
23. Rozinat, A., van der Aalst, W.M.P.: Conformance Checking of Processes Based on Monitoring Real Behavior. Information Systems 33, 64–95 (2008)
24. Fettke, P., Loos, P. (eds.): Reference Modeling for Business Systems Analysis. Idea Group, Hershey (2007)
25. Lillehagen, F., Krogstie, J.: Active Knowledge Modeling of Enterprises. Springer, Berlin (2008)
26. Jørgensen, H.D.: Interactive Process Models. PhD-thesis NTNU, Trondheim, Norway (2004)
27. van der Aalst, W.M.P., Reijers, H.A., Weijters, A.J.M.M., van Dongen, B.F., Alves de Medeiros, A.K., Song, M., Verbeek, H.M.W.: Business Process Mining: An Industrial Application. Information Systems 32, 713–732 (2007)
28. Nossum, A., Krogstie, J.: Integrated Quality of Models and Quality of Maps. In: Halpin, T., Krogstie, J., Nurcan, S., Proper, E., Schmidt, R., Soffer, P., Ukor, R. (eds.) EMMSAD 2009. LNBIP, vol. 29, pp. 264–276. Springer, Berlin (2009)
29. Alves de Medeiros, A.K., van der Aalst, W.M.P.: Process Mining towards Semantics. In: Dillon, T.S., Chang, E., Meersman, R., Sycara, K. (eds.) Advances in Web Semantics I. LNCS, vol. 4891, pp. 35–80. Springer, Heidelberg (2008)