❐ 4626

Process Mining in Supply Chains: A Systematic Literature Review

Bambang Jokonowo[1], Jan Claes[2], Riyanarto Sarno[3], Siti Rochimah[4]
[1]Information System Department, Universitas Mercu Buana, Jakarta, Indonesia
[1,3,4]Informatics Department, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2]Department of Business Informatics and Operations Management, Ghent University, Belgium

Article Info	ABSTRACT
	Performance analysis and continuous process improvement efforts are often supported by the construction of process models representing the interactions of the partners in the supply chain. This study was conducted to determine the state of the art in the process mining field, specifically in the context of cross-organizational process. The Systematic Literature Review (SLR) method is used to review a collection of twenty-one papers that are classified according to the Artifact framework of Hevner, et al. and within the Process Mining framework of Van der Aalst. In the reviewed papers, the authors conducted a variety of techniques to establish the event log, which is then used to perform the process mining analysis. Eight of the reviewed papers focus on the definition of concepts or measures. Five of the papers describe models and other abstractions that are used as a theoretical basis for process mining in the context of supply chains. The majority twenty of papers describe some kind of informal method or formal algorithm to perform process mining analysis. Nine of the papers that propose a formal algorithm also present an accompanying software implementation. Eight papers discuss the data preparation challenges and twelve papers discuss process discovery techniques.

Corresponding Author:

Bambang Jokonowo,
Information system department,
Universitas Mercu Buana,
Jalan Meruya Selatan No 1, Kembangan, Jakarta Barat, 11650, Indonesia.
Email: bambang.jokonowo@mercubuana.ac.id

1. INTRODUCTION

With the advancement of technology, collaborations between organizations have become more natural to realize. The limitations of physical distance decrease and companies expand their scope to global proportions [1,2]. At the same time, the number of transactions between companies increases as they are more closely working together. By synchronizing their processes [3,4], they are forced to become more flexible and more transparent. Hence, for closely collaborating partners, access to accurate, detailed, and complete information about the supply chain wide processes has become indispensable.

To facilitate the communication about and the synchronization of their processes, the majority of collaborating partners construct business process models [5]. These models graphically specify and represent the flow of activities within the supply chain, such that the current collaborative process can be analyzed or improved more effectively and efficiently [6]. The supply chain business process models can also be used to represent the relations between the public and the private process views of each partner in the supply chain [5] or to show the interactions between different partners in the supply chain. The construction of supply chain wide processes poses a real challenge because often the knowledge about the overall process is

distributed over the involved parties and no single party has an overview on the complete process and all its details.

Therefore, in the context of supply chain process modeling, process mining may be used as a solution to construct the overall process model. Process mining techniques include a wide variety of (semi-)automated techniques that study processes based on historical process data extracted from the supporting information systems into structured event logs. The most known and most applied technique type is process discovery [8]. It is a type of technique to automatically construct a business process model that captures the real process by analyzing the event log [9,10]. Process discovery is thus proposed to produce more objective, more complete and more up-to-date business process models [11]. It is currently not clear, however, how these techniques can be applied in the context of cross-organizational processes [12].

Therefore, we conducted a Systematic Literature Review to collect, analyze, structure, and integrate the current academic knowledge about cross-organizational process mining. Except for the collection of metadata, such as the number of published papers over time and the evolution of geographical spread of the authors, the analysis was mainly driven by two frameworks. These frameworks are selected to be suitable to get insights into the addressed research topics, the proposed contributions, and the applied research methods. The first framework describes the types of research outcomes for each paper, whereas the second is applied to classify the types of practical solutions targeted by each paper. This paper proceeds as follows. Section 2 describes how we have implemented the Systematic Literature Review method. In Section 3, the results of the analysis are presented. Section 4 provides a discussion and conclusion.

2.   RESEARCH METHOD

To reveal the current knowledge and to get insights into potentially missing knowledge about process mining of cross-organizational processes, the Systematic Literature Review (SLR) methodology was implemented. This method is assessed as reliable, profound and controllable [13]. We adopted the practical guidelines from [13–15]. Based on a search phrase, derived from the research question, a selection of databases is automatically searched to find relevant papers [14]. The resulting paper set is reduced by fine-graining the search with the manual application of inclusion and exclusion criteria [13]. The final paper set is then studied to get insights into the current state of the art of the research domain and to identify research opportunities (as in [15]). The elements that lead to the paper selection are discussed in more detail.

2.1. Research question

The research question is based on the general research goal to get an overview of current and missing academic knowledge about cross-organizational process mining. Such an overview is now lacking, whereas researchers in the past have discussed the need for it [12,16]. Therefore, the research question addressed in this paper is:

RQ1. Which knowledge about cross-organizational process mining exists in academic literature? By addressing this research question, an overview of current academic knowledge is created. This overview is useful for practitioners, who are now reporting the difficulty of finding suitable information for their cross-organizational process mining projects . On the other hand, also researchers will benefit from the overview. For example, the lack of knowledge about cross-organizational process mining was explicitly mentioned as a research challenge in the process mining community Manifesto [16].

2.2. Search and selection process

Search phrase. Based on the research question, a search string was composed to be used in an automatic search process in multiple databases to find the relevant literature for the overview. The search phrase relates to the two key concepts, which are "cross-organizational process" and "process mining". For the former concept, we consider two synonyms, i.e., "supply chain process" and "inter-organizational process". Further, the latter concept was split up in "process mining" and "workflow mining". Finally, because the early papers in these fields did not always use the more modern term "mining", we also included descriptions of these techniques that use on the one hand the words "process" or "workflow", and on the other hand one of these terms: "event log", "log file", or "audit trail". This way, the final search phrase is as follows:

("supply chain" OR "cross-organization" OR inter-organization) AND ("process mining" OR "workflow mining" OR ((process OR workflow) AND ("event log" OR "log file" OR "audit trail"))). Databases. The search phrase was used to find articles in a set of academic databases. There is no standard set of databases. Inspired by the guidelines and the examples of [14,15], we selected the five databases presented in Table 1.

Table 1. Table of Academic Databases

Code	Publisher	Database	Link
Spr	Springer	SpringerLink	www.springer.com
Sci	Elsevier	Science Direct	www.sciencedirect.com
Acm	ACM	ACM Digital Library	dl.acm.org/advsearch.cfm
Wos	Thomson Reuters	Web of Science	apps.webofknowledge.com/Search
Ieee	IEEE	IEEE Explore	ieeexplore.ieee.org/Xplore/home.jsp

This approach of selecting multiple databases is proposed to improve the completeness of the study. Note that the selected databases are academic databases, to be aligned to the research question. Inclusion and exclusion criteria. Because the automated search process includes too many papers that are not relevant, the search process is followed by a manual selection process that aims to eliminate these unrelated works from the paper set. This elimination happens according to predefined inclusion and exclusion criteria. For practical reasons, and according to the guidelines of [13], this process is performed in two phases. First, the inclusion and exclusion criteria are assessed based on only the title, abstract and keywords. In case of doubt, the paper is not discarded from the paper set to be processed further on the next step. Secondly, the criteria for the remaining papers are assessed back based on the full text.

The applied inclusion criteria (IC) and exclusion criteria (EC) are:

IC 1. Cross-organizational process model. The study needs to discuss research about process models, which describe processes that are crossing the boundaries of a single organization, spanning over two or more organizations within a supply chain.

IC 2. Process mining. The study needs to discuss research about techniques that aim to automatically construct, complete or analyze process models from historical process execution data. The techniques should be data-driven: for example, but not limited to techniques that start from event logs.

EC 1. Other models. Studies of other types of models than business process models are excluded. For example, we exclude studies about other types of process models (such as software process models) and general conceptual models (such as data models, business models, and value models).

EC 2. Management. Studies that discuss other aspects, tools or techniques than modeling, are excluded. For example, we explicitly exclude studies about business process management and supply chain management.

EC 3. Technology. Studies that discuss general technical aspects of collaborating partners are excluded. For example, we exclude studies that discuss supply chain software or technologies for data exchange between partners, if they do not relate their findings to a cross-organizational process (model).

Snowballing. To maximize the completeness of the paper set, as proposed by [13], we applied a technique called snowballing. Moreover, we applied backward snowballing that all the papers that are referenced by the papers in the set so far are also considered. By implementing the same inclusion and exclusion criteria, the paper set is extended in two steps (first considering only title, abstract and keywords, and later also the full text of the referenced papers).
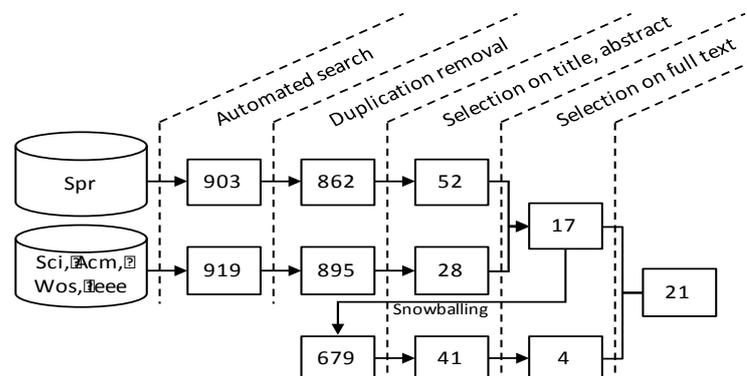

Figure 1. Overview of the search and selection process

Overview of the applied research method Figure 1 shows an overview of the search and selection process. Because SpringerLink does not support to export directly to a reference manager (but only to CSV

file), the results of this database were first analyzed separately. The automatic search with the search string resulted in an initial paper set of 903 papers from SpringerLink and 919 papers from the other databases. After removal of duplicates, respectively 41 and 24 papers were excluded. Next, based on the application of the selection criteria on the title, abstract, keywords and conclusion, the paper set was further reduced to 52 and 28 papers respectively. After downloading the full papers from Springer and after assessing the selection criteria on the full texts, the resulting paper set contained 17 unique papers. The application of the snowballing technique added 679 papers to the set, which are reduced to 41 after assessing the title and finally to 4 additional papers when the full text is being evaluated. This way, the final paper set contains 21 unique articles about cross-organizational process discovery. An overview of these papers is presented in Table 2.

Table 2. Final Paper Set

Ref.	Author & year	Title
[17]	Van der Aalst, 2000	Loosely coupled inter-organizational workflows: modeling and analyzing workflows crossing organization boundaries
[18]	Chiu, et al., 2002	Workflow view based E-contracts in a cross-organizational E-services environment
[19]	Maruster, et al., 2003	Discovering distributed processes in supply chains
[20]	Che, et al., 2007	A method for inter-organizational business process management
[21]	Gerke, et al., 2009	Process mining of RFID-based supply chains
[22]	Lau, et al., 2009	Development of a process mining system for supporting knowledge discovery in a supply chain network
[23]	Khan, et al., 2010	Applying process mining in SOA environments
[24]	Li, 2010	An automatic virtual organization structure modeling method in supply chain management
[25]	Sun, et al., 2011	Process-mining-based workflow model fragmentation for distributed execution
[26]	Van der Aalst, 2011	Intra- and inter-organizational process mining: Discovering processes within and between organizations
[27]	Buijs, et al., 2012	Towards cross-organizational process mining in collections of process models and their executions
[28]	Engel, et al., 2012	Mining inter-organizational business process models from EDI messages: A case study from the automotive sector
[29]	Rozsnyai, et al., 2012	Business process insight: An approach and platform for the discovery and analysis of end-to-end business processes
[30]	Azzini, et al., 2013	Consistent process mining over big data triple stores
[31]	Comuzzi, et al., 2013	Optimized cross-organizational business process monitoring: Design and enactment
[32]	Zeng, et al., 2013	Cross-organizational collaborative workflow mining from a multi-source log
[9]	Bernardi, et al., 2014	Discovering cross-organizational business rules from the cloud
[33]	Claes, et al., 2014	Merging event logs for process mining: A rule-based merging method and rule suggestion algorithm
[34]	Irshad, et al., 2015	Preserving privacy in collaborative business process composition
[35]	Engel, et al., 2016	Towards comprehensive support for privacy preservation cross-organization business process mining
[7]	Liu, et al., 2016	Analyzing inter-organizational business processes

3.   RESULTS AND ANALYSIS

This section describes the results of our analysis on the final paper set. First, an overview of the number of papers and the geographical spread of the first authors is presented to provide a context for further analysis. Then, the papers are classified and discussed based on two frameworks (i.e., theoretical contribution types and practical contribution types). Lastly, we provide a less systematic overview of the field and of the technologies used to distract the necessary data.

3.1. Analysis of the meta-data

Figure 2 shows the number of papers that discuss process mining techniques in the context of supply chains, according to the selected paper set. The research into supply chain process mining seems not to be abundant. The research appears to have accelerated since 2009.
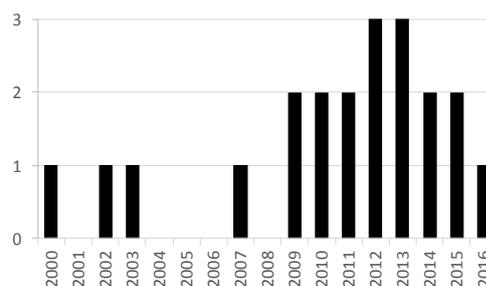

Figure 2. Number of papers per year

Further, the primary affiliation countries of the first author are presented in Figure 3. From this image, it can be concluded that supply chain process mining research is dominated by two countries: China and the Netherlands. They jointly count for 12 of the 21 papers (57%) in the literature set.
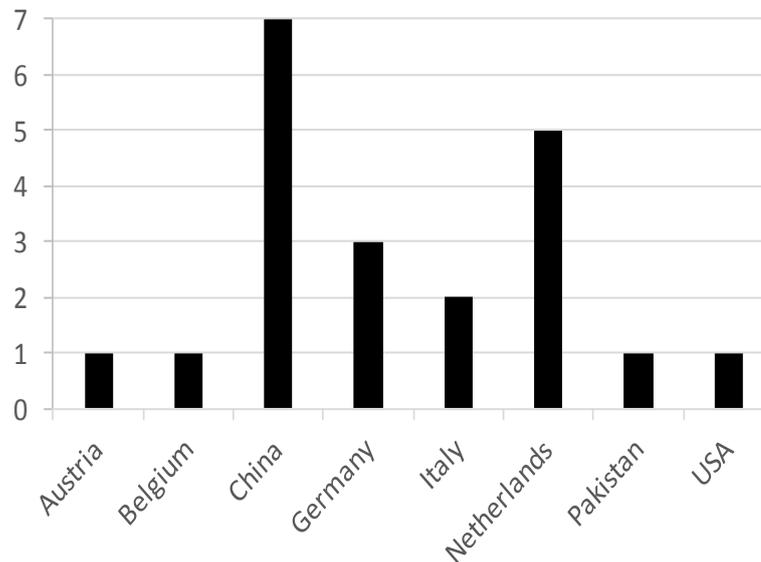

Figure 3. The number of papers per country

3.2. Classification of the Artifact framework

Hevner, et al. define four kinds of artifacts that can be developed and investigated by design science research [36]. We refer to this classification as the Artifact framework, presented in Table 3. According to the framework, products of design science research can be *constructs* (languages, terminology, definitions, and measures), *models* (abstractions and representations), *methods* (approaches and algorithms), or *instantiations* (prototype and implemented systems) [36].

Table 3. Artifact Framework by Hevner et al. [36] (p. 78ff.)

Code	Design Science Artifact	Description
A1	Constructs	*"Vocabulary and symbols. Constructs provide the language in which problems and solutions are defined and communicated."*
A2	Models	*"Abstractions and representations. Models use constructs to represent a real-world situation: the design problem and its solution space."*
A3	Methods	*"Algorithms and practices. Methods define processes; they provide guidance on how to solve problems, that is, how to search the solution space."*
A4	Instantiations	*"Implemented and prototype systems. Instantiations show that constructs, models, or methods can be implemented in a working system."*

For each paper of the paper set, it was determined which artifacts and contributions are proposed, and for each artifact, the type was derived from Table 3. A difference was made between newly proposed artifacts that can be considered the contributions proposed in the paper (presented in Table 4) and potential existing artifacts that were used for the research described in the paper (not represented in Table 4).

As can be noted in Table 4, the early contributions mainly focused on terminology and informal approaches to represent and analyze supply chain processes. Only later, from 2009 on, also concrete algorithms and techniques were developed for (semi-)automated analysis based on historical process data (=process mining techniques). The papers proposing an algorithm have an underlined x in the column labeled A3. It can be seen that 13 of the 21 papers (62%) propose a process mining (support) algorithm, which is 13 of the 17 papers (76%) after 2009 (included). Exactly 9 of these 13 algorithm-proposing papers (69%) also propose an implementation of the algorithm.

Table 4. Artifacts and Contributions of the Selected Papers

Ref.	Author & year	A1	A2	A3	A4	Contributions
[17]	Van der Aalst, 2000	x		x		An *approach* to model and analyze new and existing inter-organizational processes, a *definition* for local and global soundness
[18]	Chiu, et al., 2002	x	x	x	x	*Terminology* and *representation* for cross-organizational services and a supporting architecture *approach* and software environment *implementation*
[19]	Maruster, et al., 2003			x		An *approach* to discover supply chain processes with existing discovery techniques by imposing to use a standard identifier across the involved parties
[20]	Che, et al., 2007			x		An *approach* to combine the use of UML models and XML Nets for respectively intra- and inter-organizational business process management
[21]	Gerke, et al., 2009			x̲	x	An *algorithm* to build event logs from dispersed data sources, based on correlating product codes from RFID data, with a *prototype* implementation
[22]	Lau, et al., 2009			x̲	x	An *algorithm* to reveal association rules representing inter-organizational dependencies, with a *prototype* implementation
[23]	Khan, et al., 2010		x	x̲	x	A *model* describing process data, an informal *approach* to identify and extract process data, an *algorithm* and *implementation* to extract process data from SAP
[24]	Li, 2010			x̲		An *algorithm* to discover a social network in a supply chain based on handover of work (called a virtual organization structure model)
[25]	Sun, et al., 2011	x	x	x̲	x	The *definition* and *representation* of fragmented process information, an *approach* to deal with the management of fragmented processes and various *implemented algorithms* related to this
[26]	Van der Aalst, 2011	x	x			The *definition* and *representation* of collaboration configurations, and of horizontal and vertical partitioning dimensions
[27]	Buijs, et al., 2012			x		An *approach* for cross-organizational process analysis proposing certain metrics to cross-correlate process models and event data in different organizations
[28]	Engel, et al., 2012	x		x̲	x	An *approach* to discover an inter-organizational process model, and a correlation *algorithm* to match EDI messages to an instance to build an event log, and a software *implementation*
[29]	Rozsnyai, et al., 2012			x̲	x	An *approach* and an *algorithm* to discover correlations between distributed process instance data and a software *implementation* linking the data correlation with process mining techniques
[30]	Azzini, et al., 2013			x̲		An *approach* and *algorithm* for semantic lifting of dispersed process data (aggregating events) using semantic data mismatch detection and map reduction techniques
[31]	Comuzzi, et al., 2013	x		x̲	x	An *approach*, based on formal *definitions* and an *algorithm*, to monitor cross-organizational process infrastructures, with a software *implementation*
[32]	Zeng, et al., 2013	x		x̲		An *approach* to discover cross-organizational process models supported by a formal *algorithm* and formal *definitions* to discover coordination patterns used for integrating individual models
[9]	Bernardi, et al., 2014			x		An *approach* to use process-related data from cloud systems in combination with existing live declarative process discovery techniques to detect business rules describing the process
[33]	Claes, et al., 2014			x̲	x	An *approach*, supported by an *algorithm*, to merge event logs of inter-organizational process partners into a single log file for standard process mining, and a software *implementation*
[34]	Irshad, et al., 2015	x		x̲		An *approach*, based on formal *definitions* and a privacy-aware trace extraction *algorithm*, to mine and generate business process models in a supply chain environment
[35]	Engel, et al., 2016		x	x	x	A detailed *approach* and *representation* to use EDI messages for analyzing and discovering inter-organizational process models, with a supporting software environment *implementation*
[7]	Liu, et al., 2016			x̲		An *approach* to combine individual public models into a supply chain wide process model, supported by *algorithms* for combining and matching the public and private process models
		8	5	20	9	

The majority of the algorithms appears to focus on (support of) the integration of decentralized process data in a single event log to enable the execution of traditional process mining techniques on supply chain process data. The type of proposed process mining techniques (e.g., data preparation, discovery, conformance checking) is investigated further in Section 0.

3.3. Classification in the Process Mining framework

The second framework was the Process Mining framework proposed by Van der Aalst [37] as shown in Table 5. It describes the different types of techniques in the process mining field. The activities can be grouped into data preparation (F0), process specification in the form of models (F1, F2, F3), process auditing (F4, F5, F6, F7), and process navigation (F8, F9, F10).

Table 5. Process Mining framework by Van der Aalst [37] REF (p. 242ff.)

Code	Process Mining technique	Description
F0	Provenance	Construction of event logs from historical process data
F1	Discover	Construction of process models from event logs
F2	Enhance	Annotating process models with additional data from event logs
F3	Diagnose	Investigating behavioral syntax errors in produced process models
F4	Detect	Detect deviations of a running process instance from a given process model
F5	Check	Detect all deviations from a given process model based on event logs
F6	Compare	Detect differences between as-is and to-be process models
F7	Promote	Promote differences between as-is and to-be models to the to-be model
F8	Explore	Visualize running process instances on as-is or to-be process models
F9	Predict	Predict final properties of running process instances based on event logs
F10	Recommended	Recommend next actions of running process instances based on event logs

For each paper in the set, it was determined which activities are supported by the proposed contributions. A distinction was made between direct support being concepts about, models of, methods for, and instantiations for these process mining activities as shown 'D' in the columns of Table 6, and indirect support being preparatory artifacts as shown 'I' in Table 6. Further, Table 6 also presents whether the proposed artifacts were evaluated and how. When the value of the contributions was shown with an artificial or simplified example or analysis, this was called demonstration. A more in-depth analysis of a real or at least realistic example was called case study. The term 'empirical' was added when non-trivial statistical techniques were used. Expert interview evaluation means that also perception data was used in the evaluation.

Table 6. Process mining techniques proposed directly (D) or indirectly (I) by the selected papers Note that also proposed algorithms without implementation are regarded as direct contributions (e.g., [20]) Note that papers may additionally present analysis techniques that are not included in this framework (e.g., [25])

Ref.	Author & year	F0	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	Evaluation (N/A = not available or not applicable)
[17]	Van der Aalst, 2000				I								N/A
[18]	Chiu, et al., 2002		I										N/A
[19]	Maruster, et al., 2003	I											N/A
[20]	Che, et al., 2007						D						N/A
[21]	Gerke, et al., 2009	D											Demonstration
[22]	Lau, et al., 2009		D										Case study
[23]	Khan, et al., 2010	D											Case study
[24]	Li, 2010		D										Demonstration
[25]	Sun, et al., 2011		D										Demonstration
[26]	Van der Aalst, 2011		I	I		I	I						N/A
[27]	Buijs, et al., 2012						I	I	I			I	Case study
[28]	Engel, et al., 2012	D	I										Case study
[29]	Rozsnyai, et al., 2012	D	D					D			D		Demonstration
[30]	Azzini, et al., 2013	D											Demonstration
[31]	Comuzzi, et al., 2013											I	Empirical case study
[32]	Zeng, et al., 2013		D										Demonstration & Case study
[9]	Bernardi, et al., 2014		D										Case study
[33]	Claes, et al., 2014	D											Multi-case study & Expert interview
[34]	Irshad, et al., 2015		D										Empirical testing
[35]	Engel, et al., 2016	D	D										Case study
[7]	Liu, et al., 2016		D										Case study
		8	12	1	1	1	3	2	1	0	1	2	

It can be noted that the majority of the papers (17 of 21 papers, 81%) focus on data preparation (8 of 21 papers, 38%) and process discovery (12 of 21 papers, 57%). In most cases, they (first) attempt to combine the data of different collaborating partners [19-22,25,28,29,32]. Indeed, when the data of the collaborating partners can be prepared in such a way that they can be combined in a single event log-grouping event data for the same process instance in a single trace-the existing process mining techniques can still be used. This way, no dedicated process mining algorithms or implementations for supply chain process models need to be created, which increases reusability of the mature and robust existing techniques. This method also means

that a high number of the papers aims to indirectly contribute to all other types of process mining (i.e., F1-F10), which was not indicated in the table to avoid overload.

Further, it appears that (relatively limited) demonstration and (extended) case study are the preferred form of evaluation. More comprehensive empirical evaluations, such as multiple-case studies, multiple technique comparisons or including user perception discussions are hardly applied in this field. This evaluation may have to do with the sophisticated setting where multiple organizations are involved by definition, which is difficult for researchers to access the appropriate data for evaluation purposes (both regarding quantity and quality of data).

3.4. Diving deeper

Except for classifying the papers, we also analyzed their contents less systematically to reveal conventional strategies and approaches. Table 7 presents an overview of the investigated topics in the supply chain process mining field, according to the paper set.

Table 7. Primary Research Focuses on Supply Chain Process Mining Literature

Research focus	Ref.
Merging event logs for process mining	[33]
Privacy-preservation in process mining	[7], [34]
Process mining in cloud computing	[9], [27]
Process mining on big data	[30]
Process mining on EDI or RFID data	[21], [28], [35]
Process mining on SOA environment data	[23]
Process mining for knowledge discovery	[22]
Process mining for monitoring purposes	[31], [20], [18]
Process mining for predictive analytics	[29]
The concept of a virtual organization	[24]

One common viewpoint on data-driven process analysis (=process mining) in supply chains, is that organizations have data that they want to remain private and other data that can be made public (e.g., [7,9,17,18,32,34]). Similarly, these authors typically distinguish between a private view on an organization's part of the supply chain wide process, and a public view on the process. They consider an approach in which the public data is shared (with each other or with a trusted third party) to construct an overall process model and then each organization can link its private data or model to this public process model to complement it with the details of their internal business processes. For example, Liu et al. [7] propose a method, which includes three steps: (1) each organization discovers its private and public business process models from its event logs, (2) a trusted third-party middleware takes the public process models as input and generates cooperative public process model fragments of each organization, and (3) each organization combines its private business process model with the for them relevant public fragments to obtain the organization-specific cross-organization cooperative business process model.

Another interesting angle we were triggered by Table 7 to investigate further, is the technological aspect of the papers. Where does the historical process data that is used to construct event logs come? Table 8 provides an overview. Many papers seem to focus on transactional data used for the physical or virtual exchange of goods (e.g., RFID), services (e.g., SAAS), or information (e.g., EDI).

Table 8. Technological base of the presented techniques

Technology	Ref.
Electronic Data Interchange (EDI)	[35], [28]
Other web service-based systems	[34], [21], [31]
Radio Frequency Identification (RFID)	[21]
Software-as-a-Service (SaaS) and Cloud Computing	[9], [27]
Software Oriented Architecture (SOA)	[23]
Supply Chain Management System (SCMS)	[24]

4. CONCLUSION

In this paper, the contribution is to provide a structured overview of the current academic literature about supply chain process mining. The practical approach appears to be to focus on merging the data of the different partners in the chain into a single event log, such that existing process mining techniques can be

utilized. Furthermore, in the context of privacy concerns, a distinction is made between the public and the private data of the partners. It is the public data, which is used by for example a trusted third party to produce a supply chain wide process model, after which each organization can map its private data on this public model.

The studied paper set with 21 papers lasted to 2009 was observed until considerable attention was spent on supply chains in the process mining field. China and the Netherlands dominate research contributions regarding the affiliation country of the first author. Less than 20 of the 21 papers discuss some formal or informal process mining *approach*; 13 papers propose a particular process mining *algorithm*, and nine papers also present an *implementation* of the algorithm available for download. The majority of papers focus on the *data preparation* (8 papers) and *process discovery* (12 papers) and most papers use a (limited) *demonstration* (6 papers) or an (extended) *case study* (10 papers) to evaluate their contribution.

Although this Systematic Literature Review shows that the research into supply chain process mining appears to be limited (only 21 papers were found), we believe that the results are useful. The research in this paper addresses the need for an overview of the state of the art expressed by both practitioners   and by researchers [16]. Furthermore, it can drive future research. Whereas this study is limited to reveal the current academic literature, future work may focus on missing academic knowledge, by investigating whether the literature gaps that can be found in this paper are in fact also research gaps. Indeed, from Table 4, Table 6, Table 7, and Table 8, it can be derived which aspects are understudied, but further research is needed to investigate whether this is a problem or not. Consequently, the discovered research gaps can be addressed appropriately in order to advance both the knowledge and the practice of process mining in supply chains.

ACKNOWLEDGEMENTS

REFERENCES

[1]   A. Singh, J.T.C. Teng, Enhancing Supply Chain Outcomes Through Information Technology and Trust, Comput. Hum. Behav. 54 (2016) 290–300. doi:10.1016/j.chb.2015.07.051.

[2]   E.A. Kadir, S.M. Shamsuddin, E. Supriyanto, W. Sutopo, S.L. Rosa, Food Traceability in Supply Chain Based on EPCIS Standard and RFID Technology, Indones. J. Electr. Eng. Comput. Sci. 13 (2015) 187–194. doi:10.11591/telkomnika.v13i1.6919.

[3]   S. Yongchareon, C. liu, J. Yu, X. Zhao, A view framework for modeling and change validation of artifact-centric inter-organizational business processes, Inf. Syst. 47 (2015) 51–81. doi:http://dx.doi.org/10.1016/j.is.2014.07.004.

[4]   W. Jiang, J. Zhang, J. Li, A multi-agent supply chain information coordination mode based on cloud computing, Telkomnika. 11 (2013) 6427–6433. doi:http://dx.doi.org/10.11591/telkomnika.v11i11.3453.

[5]   Q. Zeng, F. Lu, C. Liu, H. Duan, C. Zhou, Modeling and Verification for Cross-Department Collaborative Business Processes Using Extended Petri Nets, IEEE Trans. Syst. Man, Cybern. Syst. 45 (2015) 349–362. doi:10.1109/TSMC.2014.2334276.

[6]   J. De Weerdt, A. Schupp, A. Vanderloock, B. Baesens, J. De Weerdt, A. Schupp, A. Vanderloock, B. Baesens, Process Mining for the multi-faceted analysis of business processes—A case study in a financial services organization, Comput. Ind. 64 (2013) 57–67. doi:http://dx.doi.org/10.1016/j.compind.2012.09.010.

[7]   C. Liu, H. Duan, Q. ZENG, M. Zhou, F. Lu, J. Cheng, Towards Comprehensive Support for Privacy Preservation Cross-organization Business Process Mining, IEEE Trans. Serv. Comput. PP (2016) 1. doi:10.1109/TSC.2016.2617331.

[8]   J. Claes, G. Poels, Process Mining and the ProM Framework: A Survey, Proc. BPM '12 Work. 28 (2012) 187–198. doi:10.1007/978-3-642-36285-9_19.

[9]   M.L. Bernardi, M. Cimitile, F.M. Maggi, Discovering cross-organizational business rules from the cloud, Comput. Intell. Data Min. (CIDM), 2014 IEEE Symp. (2014) 389–396. doi:10.1109/CIDM.2014.7008694.

[10]   R. Sarno, Y.A. Effendi, Hierarchy process mining from multi-source logs, Telkomnika (Telecommunication Comput. Electron. Control. 15 (2017) 1960–1975. doi:10.12928/TELKOMNIKA.v15i4.6326.

[11]   B. Vazquez-Barreiros, M. Mucientes, M. Lama, ProDiGen: Mining complete, precise and minimal structure process models with a genetic algorithm, Inf. Sci. (Ny). 294 (2015) 315–333. doi:10.1016/j.ins.2014.09.057.

[12]   B.N. Yahya, M. Song, H. Bae, S. ook Sul, J.-Z.Z. Wu, Domain-driven actionable process model discovery, Comput. Ind. Eng. 99 (2016) 382–400. doi:10.1016/j.cie.2016.05.010.

[13]   B. Kitchenham, O.P. Brereton, D. Budgen, Mark Turner, J. Bailey, S. Linkman, Systematic literature reviews in software engineering-A systematic literature review, Inf. Softw. Technol. 51 (2009) 7–15. doi:10.1016/j.infsof.2008.09.009.

[14]   I. Moreno-Montes De Oca, M. Snoeck, H.A. Reijers, A. Rodríguez-Morffi, A systematic literature review of studies on business process modeling quality, Inf. Softw. Technol. 58 (2015) 187–205. doi:10.1016/j.infsof.2014.07.011.

[15] A. Tarhan, O. Turetken, H.A. Reijers, Business process maturity models: A systematic literature review, Inf. Softw. Technol. 75 (2016) 122–134. doi:10.1016/j.infsof.2016.01.010.

[16] W. van der Aalst, A. Adriansyah, A.K.A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, J.C. Bose, P. van den Brand, R. Brandtjen, J. Buijs, A. Burattin, J. Carmona, M. Castellanos, J. Claes, J. Cook, N. Costantini, F. Curbera, E. Damiani, M. de Leoni, P. Delias, B.F. van Dongen, M. Dumas, S. Dustdar, D. Fahland, D.R. Ferreira, W. Gaaloul, F. van Geffen, S. Goel, C. Günther, A. Guzzo, P. Harmon, A. ter Hofstede, J. Hoogland, J.E. Ingvaldsen, K. Kato, R. Kuhn, A. Kumar, M. La Rosa, F. Maggi, D. Malerba, R.S. Mans, A. Manuel, M. McCreesh, P. Mello, J. Mendling, M. Montali, H.R. Motahari-Nezhad, M. zur Muehlen, J. Munoz-Gama, L. Pontieri, J. Ribeiro, A. Rozinat, H. Seguel Pérez, R. Seguel Pérez, M. Sepúlveda, J. Sinur, P. Soffer, M. Song, A. Sperduti, G. Stilo, C. Stoel, K. Swenson, M. Talamo, W. Tan, C. Turner, J. Vanthienen, G. Varvaressos, E. Verbeek, M. Verdonk, R. Vigo, J. Wang, B. Weber, M. Weidlich, T. Weijters, L. Wen, M. Westergaard, M. Wynn, Process Mining Manifesto, (2012) 169–194. doi:10.1007/978-3-642-28108-2_19.

[17] W. van der Aalst, W. van der Aalst, Loosely coupled interorganizational workflows: modeling and analyzing workflows crossing organizational boundaries, Inf. Manag. 37 (2000) 67–75. doi:10.1016/S0378-7206(99)00038-5.

[18] D.K.W. Chiu, K. Karlapalem, Q. Li, E. Kafeza, Workflow view based E-contracts in a cross-organizational E-services environment, Distrib. Parallel Databases. 12 (2002) 193–216. doi:10.1023/A:1016503218569.

[19] L. Maruster, J.C. Wortmann, A. Weijters, W.M.P. van der Aalst, Discovering distributed processes in supply chains, in: Jagdev, HS and Wortmann, JC and Pels, HJ (Ed.), Collab. Syst. Prod. Manag., 2003: pp. 219–230.

[20] H. Che, M. Mevius, Y. Ju, W. Stucky, R. Trunko, A method for inter-organizational business process management, Proc. IEEE Int. Conf. Autom. Logist. ICAL 2007. (2007) 354–358. doi:10.1109/ICAL.2007.4338586.

[21] K. Gerke, A. Claus, J. Mendling, Process Mining of RFID-Based Supply Chains, in: 2009 IEEE Conf. Commer. Enterp. Comput., IEEE Computer Society, Washington, DC, USA, 2009: pp. 285–292. doi:10.1109/CEC.2009.72.

[22] H.C.W.C.W. Lau, G.T.S.T.S. Ho, Y. Zhao, N.S.H.S.H. Chung, Development of a process mining system for supporting knowledge discovery in a supply chain network, Int. J. Prod. Econ. 122 (2009) 176–187. doi:http://dx.doi.org/10.1016/j.ijpe.2009.05.014.

[23] A. Khan, A. Lodhi, V. Köppen, G. Kassem, G. Saake, Applying Process Mining in SOA Environments, in: 2010: pp. 293–302. doi:10.1007/978-3-642-16132-2_28.

[24] Y. Li, An Automatic Virtual Organization Structure Modeling Method in Supply Chain Management, Manag. Serv. Sci. (MASS), 2010 Int. Conf. (2010) 1–4. doi:10.1109/ICMSS.2010.5575841.

[25] S.X. Sun, Q. Zeng, H. Wang, Process-Mining-Based Workflow Model Fragmentation for Distributed Execution, IEEE Trans. Syst. Man, Cybern. - Part A Syst. Humans. 41 (2011) 294–310. doi:10.1109/TSMCA.2010.2069092.

[26] W.M.P. Van Der Aalst, Intra- and inter-organizational process mining: Discovering processes within and between organizations, Lect. Notes Bus. Inf. Process. 92 LNBIP (2011) 1–11. doi:10.1007/978-3-642-24849-8_1.

[27] W.M.P. Buijs, J.C.A.M., Dongen, B.F. and Aalst, Towards Cross-Organizational ProcessMining in Collections of ProcessModels and Their Executions, (2012) 2–13.

[28] R. Engel, W.M.P. Van Der Aalst, M. Zapletal, C. Pichler, H. Werthner, Mining inter-organizational business process models from EDI messages: A case study from the automotive sector, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 7328 LNCS (2012) 222–237. doi:10.1007/978-3-642-31095-9_15.

[29] S. Rozsnyai, G.T. Lakshmanan, V. Muthusamy, R. Khalaf, M.J. Duftler, Business Process Insight: An Approach and Platform for the Discovery and Analysis of End-to-End Business Processes, 2012 Annu. SRII Glob. Conf. (2012) 80–89. doi:10.1109/SRII.2012.20.

[30] A. Azzini, P. Ceravolo, Consistent Process Mining Over Big Data Triple Stores, in: 2013 IEEE Int. Congr. BIG DATA, 2013: pp. 54–61. doi:10.1109/BigData.Congress.2013.17.

[31] M. Comuzzi, I. Vanderfeesten, T. Wang, Optimized cross-organizational business process monitoring: Design and enactment, Inf. Sci. (Ny). 244 (2013) 107–118. doi:10.1016/j.ins.2013.04.036.

[32] Q. Zeng, S.X. Sun, H. Duan, C. Liu, H. Wang, Cross-organizational collaborative workflow mining from a multi-source log, Decis. Support Syst. 54 (2013) 1280–1301. doi:10.1016/j.dss.2012.12.001.

[33] J. Claes, G. Poels, Merging event logs for process mining: A rule based merging method and rule suggestion algorithm, Expert Syst. Appl. 41 (2014) 7291–7306. doi:http://dx.doi.org/10.1016/j.eswa.2014.06.012.

[34] H. Irshad, B. Shafiq, J. Vaidya, M.A. Bashir, S. Shamail, N.R. Adam, Preserving privacy in collaborative Business Process Composition, 2015 12th Int. Jt. Conf. E-Bus. Telecommun. 4 (2015) 112–123. doi:10.5220/0005567801120123.

[35] R. Engel, W. Krathu, M. Zapletal, C. Pichler, R.P.J.C. Bose, W. van der Aalst, H. Werthner, C. Huemer, W. Aalst, H. Werthner, C. Huemer, Analyzing inter-organizational business processes, Inf. Syst. E-Bus. Manag. 14 (2016) 577–612. doi:10.1007/s10257-015-0295-2.

[36] A.R. Hevner, S.T. March, J. Park, S. Ram, Design Science in Information Systems Research, MIS Q. 28 (2004) 75–105. doi:10.2307/25148625.

[37] W.M.P. der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, 1st ed., Springer Publishing Company, Incorporated ©2011, 2011.

BIBLIOGRAPHY OF AUTHORS

Bambang Jokonowo is a researcher at Informatics Engineering, Sepuluh Nopember Institute of Technology, Indonesia. His research focus Process mining and he is also a member of the research team in Lab-Based Education (LBE) Enterprise Intelligent System at (https://lbeifits.wordpress.com/about-us/). He has been a guest researcher process mining community (http://www.janclaes.info/news.php). He is teaching Business Process Analysis, Process Modeling course at Mercu Buana University, Jakarta.

Jan Claes is a research scientist at Business Informatics Ghent University, Belgium. His research focuses in a process modeling, process mining, and problem solving (Types of cognitive load, Cognitive Load Theory, and Cognitive Fit Theory). His research articles and activities can be found at http://www.janclaes.info/research.php

Riyanarto Sarno is a Professor in Informatics Engineering, Sepuluh Nopember Institute of Technology, Indonesia. He earned Ir in Electronical Engineering at Institut Teknologi Bandung, DrsEc in Economic at Pajajaran University. He received MSc (1988) and PhD (1992) in Computer Science from the University of New Brunswick UNB Canada. His research interests include Process Mining, semantic web service, Semantic ERP. He is a leader of Lab-Based Education (LBE) Enterprise Intelligent System, and He has published journal and conference paper at https://lbeifits.wordpress.com/

Siti Rochimah is a lecturer in Informatics Engineering, Sepuluh Nopember Institute of Technology, Indonesia. She receives a master degree at Institut Teknologi Bandung and doctoral degree at University Technology Malaysia. Her research interests are in software quality assurance and software evolution.