

プロセスマイニングを目的としたイベントログ抽出：系統的な文献批評

2019.11.6

ダキッチ・ドゥサンカ, ステファノヴィッチ・ダルコ, ロリック・テオドラ, ナランジッチ・ダジヤナ, シメウノヴィッチ・ネナド

## 抄録

プロセスマイニングは、プロセスモデルの自動発見、同一プロセスのイベントログと既存のプロセスモデルの比較、既存のプロセスモデルの改善を可能にすることで、プロセスモデル分析とデータマイニングなどのデータ指向の分析とのギャップを埋める。プロセスマイニングの前提条件は、現実のビジネスプロセスを支援・制御し、その結果、メッセージ、トランザクション、ログなどのイベントデータをイベントログとして何らかのデータベースに格納する情報システムであることである。イベントデータは、抽出され、フィルタリングされ、プロセスマイニングソフトウェアにロードされ、そこである種のプロセスマイニングを行うことができる。プロセスを意識した情報システム(PAIS)は、プロセスのイベントを相関させるためにケースの概念を明示的に仮定しており、そのようなログを直接提供する。しかし、ビジネスプロセスの実行をサポートする情報システムの多くは、明示的にプロセスを認識しておらず、また、イベントデータのソースが多様であるため、プロセスマイニングのこのフェーズは困難であり、最も時間がかかる。その結果、様々なイベントログ抽出技術、アプローチ、ツールが開発されている。本論文では、この問題に貢献するために、アプローチの汎用性、専門家以外による適用可能性、および開発された実現可能なツールに関する疑問に答えることを目的として実施された体系的な文献レビューを紹介する。

## 1. 導入

プロセス指向の組織観以来、いくつかの企業情報システムは、明示的なプロセス概念を採用し、「構造化されたビジネスプロセスのための汎用的なモデリングと実行機能」を提供している[1]。これらの情報システムは、何らかの形でプロセスの実行に関するデータを保存している。プロセスマイニングでは、このデータからイベントログを生成することが可能であることを前提としており、「イベントログの各イベントは、アクティビティ（プロセス内の明確に定義されたステップ）を参照し、特定のケース（プロセスインスタンス）に関連している」[2]。しかし、すべてのPAISがこのような方法でイベントデータを記録するわけではない。この伝統的なワークフロー管理システム(WMS)やビジネスプロセス管理(BPM)システムの他に、企業資源計画(ERP)や顧客関係管理(CRM)システムなどの他の情報システムは、どちらかというデータ中心でオブジェクト中心、つまり「それらのプロセスは複数の相互に関連したビジネスオブジェクト上で動作し、それぞれが独自のケース識別子、独自の動作、相互作用を持っている」とい

う意味である[10].

ビジネスオブジェクト (ERP のドキュメントなど) に関する情報は、様々なデータテーブル[14, 19]に散在している。したがって、様々なデータソースからイベントログを抽出することは、プロセスマイニングの最も困難なフェーズであり、特に非専門家、すなわちプログラミングの知識を持っていないビジネスアナリストにとっては、障害となる可能性がある。過去 10 年の間に、イベントログの抽出をサポートし、自動化するためのアプローチとそれに対応するツールが開発されてきた。しかし、研究者と実務者が合意した汎用的なアプローチはまだ存在せず、それぞれのアプローチにはまだ限界がある。現在の文献では、これらのアプローチの特徴や課題をまとめたものがなく、実務家がプロセスマイニングのこの初期段階を実施することは困難である。

本論文では、バーバラ・キッチナムさん[2]の「ソフトウェア工学におけるシステムマティクな文献レビュー実施のためのガイドライン」に基づき、過去 10 年間に実施したイベントログ抽出アプローチに関するシステムマティクな文献レビューの結果を、主に次の点に焦点を当てて紹介する。専門家以外でも適用可能なイベントログ抽出アプローチ、あらゆる情報システムからのデータ抽出に適用可能な汎用的なイベントログ抽出アプローチの可能性、イベントログ抽出をサポートする実現可能なツール。

本稿の残りの部分は次のように構成されている。第 2 節では、プロセスマイニングとイベントログの基本的な概念の紹介、第 3 節では、システムマティクな文献レビューの方法論について、第 4 節では、システムマティクな文献レビューのレビューフェーズの実施計画を示す。第 5 節では、システムマティクな文献レビューのレビューフェーズを実施し、一次研究を選択し、レビュー結果を提示した。第 6 節では、システムマティク・リテラシー・レビューの結果について詳細に論じ、第 7 節では論文を締めくくり、今後の研究を示唆している。

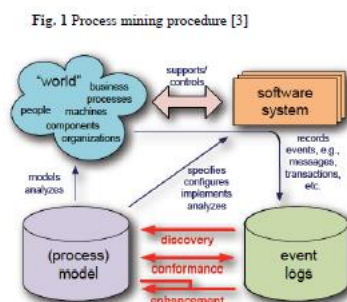
## 2. プロセスマイニングとイベントログの概念

プロセスマイニングは新しい研究分野であり、IEEE Task Force on Process Mining によって導入され、プロセスマイニング宣言[3]で定義されている。それらはプロセスマイニングを次のように定義している。

プロセスマイニングの考え方は、今日の (情報) システムで容易に利用可能なイベントログから知識を抽出することで、(想定されたプロセスではなく) 実際のプロセスを発見し、監視し、改善することである。プロセスマイニングには、(自動化された) プロセス発見 (すなわち、イベントログからプロセスモデルを抽出する)、適合性チェック (すなわち、モデルとログを比較して逸脱を監視する)、ソーシャルネットワーク/組織マイニング、シミュレーションモデルの自動構築、モデル拡張、モデル修復、ケース予

測, 履歴ベースの推奨などがある.

Fig. 1. presents prerequisites of process mining and a basic process mining procedure [3].



ソフトウェアシステムは, 現実のビジネスプロセス, 機械, コンポーネント, 組織, および人々をサポートし, 制御している. ソフトウェアシステムは, プロセスモデルによって構成され, 指定されることを前提とする. その結果, これらのソフトウェアシステムは, メッセージやトランザクションなどのイベントを記録し, 後に様々なデータソースから抽出されてイベントログとして構築される. イベントログは, 3つの基本的なタイプのプロセスマイニングを実行するための出発点であり, 発見, 適合性, 増進である. IEEE Task Force on Process Mining によると, 構築されたイベントログでは, [3]を想定している.

- ・ イベントとは, プロセスの中で明確に定義されたステップであり, 特定のケース, すなわちプロセスインスタンスに関連するプロセス活動やタスクを指す.
- ・ イベントは順番に行われる.
- ・ ケースまたはプロセスインスタンスは, ビジネスプロセスの特定の発生であり, アクティビティは, 実行されているケースの一部の操作である.
- ・ 各ケースには固有の識別子がある.
- ・ イベントログには, ケースやアクティビティの情報だけでなく, リソース (アクティビティの実行に使用された人やデバイス), イベントのタイムスタンプ (イベントが開始された瞬間や終了した瞬間), その他のデータ要素が記録されている.

2010年, IEEE Task Force on Process Mining は, プロセスマイニングイベントログの標準規格として XES (Extendable Event Stream)を採用した[3]. XES は XML ベースのフォーマットで, 各イベントログエントリはイベントタイプ, タイムスタンプ, その他の追加属性を持っている.

プロセスマイニング宣言[3]は, プロセスマイニングの指導原則と主要な課題を概説した. イベントログに関連する指導原則と課題, およびこの文献レビューの範囲について議論する. イベントログに関連する指導原則は以下の通りである.

- ・ GP1: イベントデータは第一級市民として扱われる.
- ・ GP2: ログの抽出は質問によって駆動される.

- ・ GP6: プロセスマイニングは連続的なプロセスである。

GP1 では、プロセスマイニングの品質はイベントログの品質、すなわちプロセスマイニングの入力に直接関係していると述べている。そのため、IEEE Task Force on Process Mining は、イベントログが満たすべき品質基準を提案している。「イベントは信頼できるものでなければならない、すなわち、記録されたイベントが実際に起こったと仮定しても安全であり、イベントの属性が正しいものでなければならない。イベントログは完全でなければならない、つまり、特定のスコープが与えられていて、イベントが欠落してはいけい。記録されたイベントはすべて、十分に定義された意味論を持つべきである。さらに、イベントデータは、イベントを記録する際にプライバシーやセキュリティの懸念に対応しているという意味で安全であるべきである」 [1].

GP2 では、イベントデータを抽出するデータソースには様々な情報が含まれているため、どの情報を抽出するかを決めるのは簡単なことではない。したがって、イベントログの抽出は、質問によって駆動される。

GP6 では、「生きたプロセスモデル」を生成するために、過去のイベントデータと現在のデータの両方をプロセスマイニングに使用できることを提案している。したがって、連続的なイベントログの抽出を可能にするアプローチも前提条件となる。

以下の課題は、イベントログに関するものである：

- ・ C1: イベントデータの検索、マージ、クリーニング。
- ・ C2: 多様な特性を持つ複雑なイベントログへの対応。
- ・ C10: 専門家ではない方に対して使いやすくする。
- ・ C11: 専門家ではない方に対して分かりやすくする。

イベントデータの検索、マージ、およびクリーニングに関する主な C1 の課題は、「イベントデータは「プロセス中心」ではなく「オブジェクト中心」であることが多く、イベントデータは不完全、である[3].

課題 C10 と C11 は、現在のプロセスマイニング技術が専門家ではない人にとって実行と理解が困難である可能性があることを強調している。プロセスマイニング技術の性能にはイベントログ抽出の段階が含まれており、この課題はイベントログ抽出のアプローチにも適用されている。

### 3. 方法論

この系統的な文献批評を実施するために、Barbara Kitchenham[2]によって開発された系統的な文献批評の手順に従った。Kitchenhamによると、系統的な文献批評はレビューの計画、レビューの実施、レビューの報告の3つの主要なフェーズに要約できる。

#### 3.1 レビューの計画

レビューの計画段階では、系統的な文献レビューの必要性を決定し、精緻化し、レビュープロトコルを開発することを目標としている。系統的な文献レビューの必要性は、Kitchenham[2]が提案したガイドラインに沿って、その特定の研究領域における既存の文献レビューを批評することによって確立することができる。レビューの必要性が確立されると、レビュープロトコルを作成することができる。

レビュープロトコルは、調査の背景と事前に決定した調査の根拠を簡潔に提示しなければならない。レビュープロトコルの第二の構成要素は、調査の質問を策定することである。Kitchenham [2]によると、研究の質問は、実務家や研究者にとって意味のある重要なものであり、現在のソフトウェアエンジニアリングの実践に変化をもたらすか、現在の実践の価値に対する信頼を高め、一般的に抱かれている信念と現実との矛盾を明らかにするものでなければならない。さらに質問の構成は、レビューの影響を受ける人々のグループ、特定の問題に対処するソフトウェア技術と結果、すなわち実務家にとって重要な要素の3つの異なる視点を考慮しなければならない。

研究の質問が定まった後、初期のスコーピングに基づいた一次研究の検索戦略により、どの検索用語とデータベースを使用するかが決まる。研究の選択基準は、一次研究の初期分析のための包含基準と除外基準として使用し、結果的に文献批評から研究を除外するために確立する必要がある。

レビュープロトコルの次のステップでは、文献レビューに含まれる一次研究に適用される品質評価チェックリストを決定する。Kitchenhamが述べているように、「品質」には独自の定義はない。しかし、CRDガイドライン[4]やコクランレビューアハンドブック[5]を書いた他の著者は、「質とは、研究のバイアスを最小化し、内部的妥当性と外部的妥当性を最大化する程度に関係している」と示唆している。またバイアスによって、「真の」結果から系統的に乖離した結果を生み出す傾向が考慮される。内部妥当性とは、研究の設計と実施が系統的な誤りを防ぐ可能性があるかどうかを示す程度のことである。外部妥当性、すなわち一般化可能性、適用可能性とは、研究で観察された効果が研究の外でも適用可能であるかどうかの程度である。最後のステップは、データ抽出のた

めの戦略の定義, 抽出されたデータの合成, およびプロジェクトのタイムテーブルである.

### 3.2 レビューフェーズの実施と報告

プロトコルが開発されると, レビューを開始することができる. これには以下の活動が含まれる [2].

1. 研究の識別.
2. 研究の選択.
3. 学習の質の評価.
4. データの抽出と進捗状況のモニタリング.
5. データの合成.

レビューの報告は, 系統的な文献レビューの最終段階であり, 系統的批評の結果を効果的に伝えることが重要である. 通常, 系統的批評は少なくとも 2 つの形式で報告される [2].

- 技術報告書や博士論文の一節.
- 雑誌や学会の論文.

### 4. レビューの計画

レビューの計画段階の最初の活動は, 対象の既存の文献レビューを批評することによって, 系統的な文献レビューの必要性を精緻化することである [2]. しかし, イベントログ抽出をテーマとした明示的な系統的な文献レビューはなく, プロセスマイニング分野の一般的なレビュー, その適用可能性, 技術の現状と課題についての批評である [6, 7, 8, 9].

イベントログ抽出アプローチに関する他の情報源としては, テーマの新規性を提示している研究の関連作業章がある. しかし, これらの研究は, 関心のある特定のトピックに関連する過去のイベントログ抽出アプローチについてのみ言及している. この系統的な文献レビューでは, 現在のイベントログ抽出アプローチ, 非専門家による使い勝手, 汎用性, 開発されたツールについて, 研究課題に答える形で情報をまとめている.

Kitchenham [2]のガイドラインに基づき, 以下のような研究課題を設定した.

- Q1: 専門家ではない人でも適用できるイベントログ抽出のアプローチはあるか.
- Q2: あらゆるタイプの情報システムからイベントデータを抽出するための汎用的なイベントログ抽出アプローチはあるか.

Q3: イベントログの抽出をサポートする実現可能なツールはあるか。

この文献レビューの目的で、以下のデータベースを検索した。

- SCOPUS,
- Web of Science and
- Google scholar.

エルゼビアの Scopus は、査読付き文献の最大の抄録・引用データベースである。Web of Science は、単一のインターフェース内で、複数のソースからのリンクされたコンテンツの引用メトリクスを介して接続された信頼性の高い統合された学際的な研究へのアクセスを提供する。

これらのデータベースを検索するために定義された検索用語を以下に示す。

“Process mining” AND “event log” AND extraction AND PUBYEAR > 2008

このレビューのために定義された包含基準は以下の通りである。

1. 本論文では、プロセスマイニングを目的としたイベントログ抽出のためのアプローチを提示しなければならない。
2. アプローチには、開発された手順に関する詳細な情報が含まれていなければならない。
3. この研究では、実生活のシナリオに適用可能な実現可能なアプローチを提示しなければならない。

レビューのために定義された除外基準は以下の通りである。

4. 異なるデータベースで発見された重複論文は削除しなければならない。
5. 1人の著者が同じアプローチに関する複数の論文を持っていた場合は、1つの論文のみをレビューに含める必要がある。

この文献レビューの目的のために、データ抽出戦略を開発した。各一次研究について、研究上の質問に答えるために、以下の特徴を抽出する。

1. 出版年と出典の種類。
2. 開発したツール。

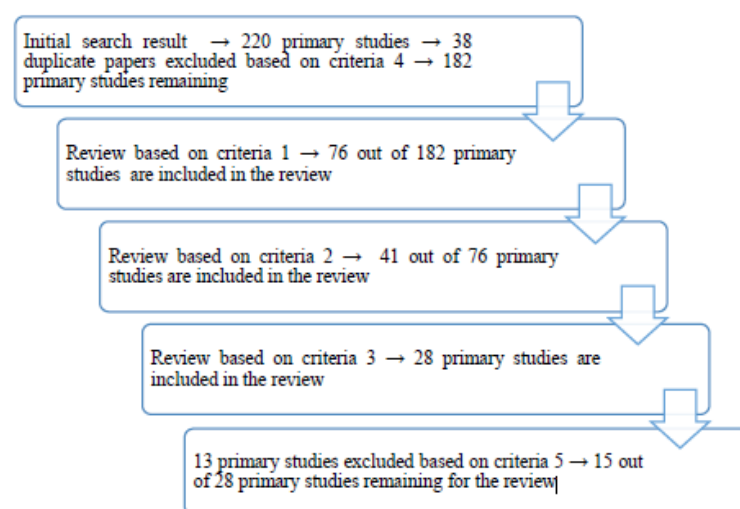
3. アプローチの実現性と適用性.
4. アプローチや情報システムの汎用性.
5. 専門家以外の方への適用可能性.

## 5. レビューの実施

### 5.1 一次研究の特定と選択

レビューフェーズを実施するための最初の活動は、研究の特定であり、系統的な文献レビューに含まれる一次研究である。一次研究は、レビュープロトコルで定義された包含基準と除外基準に従って同定される。Fig. 2は、適切なフロー図に、包含・除外プロセスの定量的エビデンスをまとめたものである。

Fig. 2 Flow diagram of the exclusion and inclusion process



あらかじめ定義された検索語を用いてデータソースを検索した結果、220件の一次研究が特定された。除外基準4（異なるデータベースで発見された重複論文は削除しなければならない）に基づき、38件の重複論文が発見され、それ以降の研究から除外された。また、基準1（プロセスマイニングを目的としたイベントログ抽出のためのアプローチを提示しなければならない）に基づいて、論文のタイトルとアブストラクトを対象にレビューを行い、76件の一次研究を含めることができた。基準2（開発されたアプローチに関する詳細な情報を論文に記載しなければならない）に基づくレビューでは、文献レビューに41件の一次研究が含まれていた。41件の一次研究のうち、28件が基準3（実生活に適用可能な実現可能なアプローチを提示していること）を満たしていた。最後のステップでは、基準5（1人の著者が同じアプローチに関する複数の論文を持って



いた場合は、1つの論文のみをレビューに含める)に基づいて 13 本の論文を除外した。最後に、15 の一次研究を文献レビューに含める。

## 5.2 データの抽出とまとめ

本章では、一次研究の主な特徴をまとめ、データ抽出戦略に基づいて抽出されたデータを表にして提示する。

Fig. 3 は、このアプローチを提示した主要研究が発表された年別のイベントログ抽出アプローチを示したものである。折れ線図から、プロセスマイニング宣言[3]が発表された 2 年後の 2012 年と 2017 年の 2 つのピークがあることがわかる。

Fig. 3 Summary of event log extraction approaches by years

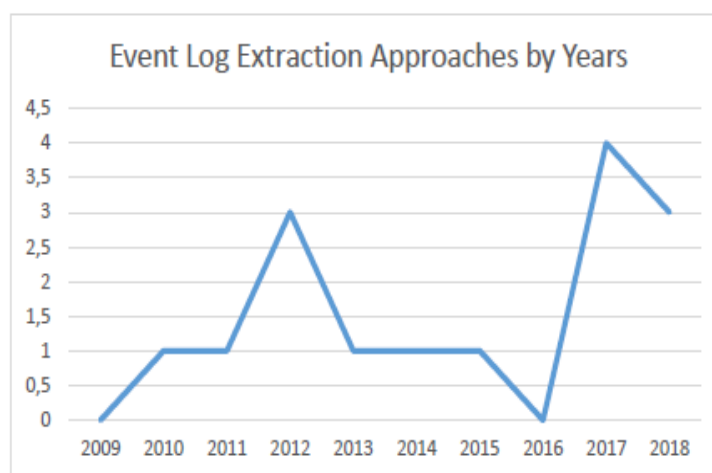


Table 1. は一次研究の出典の種類を示しており、アプローチの 40% が雑誌論文として発表されており、次いで会議論文と修士論文が 27%、イベントログ抽出の手順が 1 つ発表されていることを示している。発表された 4 件の修士論文のうち 3 件はアイントホーフェン工科大学で実施されたものである。

Table 1 Source types of primary studies

Source type	Primary studies	%
Journal Article	10, 15, 16, 17, 19, 20	40%
Conference paper	12, 18, 22, 23	27%
Master's thesis	11, 13, 14, 24	27%
Procedure	21	7%

レビュープロトコルでは、3つの研究質問が定式化されている。Table 2. は、最初の質問、Q1: 専門家でなくても適用できるイベントログ抽出アプローチはあるか、に答える抽出データを示している。

Table 2 Primary studies and applicability of their approaches by non-experts

Applicable for non-experts?	Primary studies	%
yes	14, 16, 21	20%
no	10, 11, 12, 13, 15, 17, 18, 19, 20, 22, 23, 24	80%

Table 2 は、レビューされたアプローチの 20%が非専門家でも適用可能であることを示しているが、80%のアプローチは広範なプログラミング知識を必要としている。

Table 3 は、第 2 の質問に答える抽出データを示しています。

Q2: あらゆるタイプの情報システムからデータを抽出するための汎用的なイベントログ抽出アプローチはあるか。

Table 3 Primary studies and information system types that their approaches deal with

Information system types	Primary studies	%
ERP	10, 12, 13, 19, 21, 22, 23, 24	53%
Potentially generic	14, 16, 17	20%
Cloud systems	11, 20	13%
Non-process aware information systems	15	7%
Unstructured business processes	18	7%

データ抽出の段階では、レビューされたイベントログの抽出アプローチは、異なる情報システムの種類によってグループ化できると結論づけられた。最も多く使用されているのは、オブジェクトとデータ中心のシステムである ERP システムである。3つのアプローチ[14, 16, 17]は、汎用的である可能性を示した。さらに、クラウドシステムからのイベントログ抽出を扱った2つのアプローチ[11, 20]、完全にプロセスを意識していない情報システムからのイベントログ抽出のためのソリューションを提示したアプローチ[15]、そして最後に[18]では、非構造化ビジネスプロセスをサポートする情報システムからのイベントログ抽出のためのアプローチを開発した。

Table 4 は、イベントログ抽出のアプローチを自動化するために開発されたツールやプラグインの情報を、開発された年が高い順に、イベントログ抽出処理の結果として生成されるイベントログのフォーマットについての追加情報とともに示している。第三の質問(Q3: イベントログ抽出をサポートする実現可能なツールはあるか)では、これらのツールの実現可能性の問題を取り上げている。開発された各ツールまたはプラグインについて、対応する一次研究では、アプローチの適用可能性を示すために、少なくとも1つのケーススタディを提示した。したがって、すべてのツールやアプローチには一定の制限があるものの、それらは実現可能であると考えられる。

Table 4 Developed tools and plug-ins with event log format

Tools/plugin-ins	Primary studies	Year	Event log format
XES Mapper	14	2010	XES and MXML
Prototype for event log extraction from SAP ECC 6.0	13	2011	CSV
Xtract	12	2012	XES
Event-traces Injector (ETI)	15	2012	MXML
Eventifier	17	2012	XES
MANA	18	2013	XES
Xtract V2	10	2015	XES
Xtract V3	11	2017	XES
Ontology-driven extraction plug-in for ProM	23	2017	XES
Process mining tool and ERP connector	24	2017	XES
OpenSlex and PADAS	16	2018	new format, convertible to XES
XOC Log Generator Plugin in ProM 6	22	2018	eXtensible Object-Centric event logs (XOC)

## 6. 議論

本節では、第5節で紹介した系統的な文献批評を実施した結果について述べる。

レビューされたアプローチのうち、最も多い数が2012年と2017年に出版されたものである。前述のプロセスマイニング宣言は2012年にIEEE Task Force on Process Miningによって発表されたもので、その年に発表されたアプローチの数が多いことがわかる。また2017年までに、これらの最初のイベントログ抽出アプローチは、業界で活用され、限界と課題、および改善の理由を示した。2017年以降に開発されたアプローチは、この課題を解決する傾向がある。ほとんどのアプローチは雑誌論文で発表されている。しかし、かなりの数のアプローチは、プロセスマイニングのアイデアが生まれたアイントホーフェン工科大学の修士論文を通して開発された。

研究質問Q1に答える目的で、非専門家によるイベントログ抽出の使いやすさを向上させる傾向のあるアプローチを特定した。J.C.A.M. Buijsは修士論文[14]の中で、ビジネスアナリストがデータソースからプロセスマイニングに適したイベントログ形式に変換する際に、プログラミングを必要とせずにデータを変換するためのガイドとなるツールがないと述べており、XES Mapper (後にXESame1と呼ばれる)というツールを作成している。[16]では、プロセスとデータの視点を統合し、SQLの知識がなくても多視点のイベントログ構築と分析を可能にするメタモデルの必要性を認識していた。[21]では、プロセスマイニングの知識が乏しいプロセスアナリストをイベントログ抽出プロセスに誘導し、プロセスインスタンス、アクティビティ、属性の選択に関連した決定とその結果に対するアナリストの理解を高め、イベントログ抽出アプローチを実施するための実践例や十分な背景情報をプロセスアナリストに提供する手順を開発した。

質問 Q2 は、あらゆるタイプの情報システムに適用できる可能性のある 3 つのアプローチ [14, 16, 17] があることから、肯定的に回答されている。これらのアプローチを報告した主要な研究は、異なる情報システムのデータソースでイベントログ抽出が行われているケーススタディを提示しているか、少なくともイベントログ抽出アプローチやツールの一般的な適用のためのガイドラインを提供している。これらの研究のうち 2 件 [14, 16] は、非専門家にも適用可能であることがわかった。

一方、ERP システムからのイベントログ抽出のためのアプローチは、ほとんどが開発されている [10, 12, 13, 19, 21, 22, 23, 24]。ERP システムは、ビジネス文書に沿ったオブジェクト指向であり、ビジネスプロセスに沿ったものではない。静的データの背後にあるプロセスデータを発見するためには、業務文書と関連するアクティビティをケース、すなわち同じプロセスの実行を一致させる必要がある。そのため、複雑なケースにも対応できる適切なケース識別方法を見つけなければならない。また ERP システムからのイベントログ抽出を実行している間に、収束と発散という 2 つの主要な問題が特定された [10, 11, 12, 13, 14, 22, 24]。[14] では、著者は以下のように定義している。収束は複数のプロセスインスタンスに対して一度に同じアクティビティを実行した場合に発生し、発散は 1 つのプロセスインスタンスに対して同じアクティビティを複数回実行した場合に発生する。発散と収束は未解決の問題として残っているが、アーティファクト中心のアプローチを開発したいくつかの主要研究は、特定の状況でこれらの問題を回避する方法を示唆している [10, 11, 14, 24]。

Q3 を考慮して、Table 4 に示すように、イベントログの抽出処理を自動化するために、様々なツールやプラグインが開発されている。

## 7. 結論

今日の高度に自動化されたビジネス環境に伴い、記録されるデータ量は急速に増加している。これに対応して、プロセスマイニングなど、このデータを活用した新しい研究分野が登場してきている。プロセスマイニングは比較的新しい研究分野であり、課題に直面している。本論文では、プロセスマイニングの最も困難なフェーズであるイベントログの抽出について、体系的な文献レビューを行った。既存のアプローチを特定し、以下の研究質問に答える。Q1: 専門家でなくても適用できるイベントログ抽出アプローチはあるか。Q2: あらゆるタイプの情報システムからイベントデータを抽出するための汎用的なイベントログ抽出アプローチはあるか。Q3: イベントログ抽出をサポートする実現可能なツールはあるか。

いくつかのアプローチが、専門家以外の方がより高い理解度とユーザビリティを提供

していると結論づけられているが、これはまだ少数のアプローチであると言える。さらに、いくつかのアプローチは、汎用的、すなわちあらゆる情報システムのタイプに適用できる可能性を持っている。実現可能なツールが検討されているため、研究者がプロセスマイニングのイベントログ抽出フェーズを自動化するために使用できる多くのプロトタイプやプラグインが開発されている。

今後は、収束と発散、連続的なイベントログの抽出、専門家以外のユーザビリティの向上、可能な解決策の提示など、検出された課題に焦点を当てていく必要がある。

## 認知度

この記事は、セルビア共和国の教育・科学・技術開発省の支援を受け、2011年から2019年までの期間、研究プロジェクトの一環として制作されました：no.47028「EU加盟プロセスにおけるセルビアの競争力の向上」。

## 8. 参考文献

1. Van der Aalst, W. M. & Weijters, A. J. M. M. (2004) Process mining: a research agenda, *Comput. Ind.*, doi: 10.1016/j.compind.2003.10.001
2. Kitchenham B. (2004), Procedures for Undertaking Systematic Reviews, In: Joint Technical Report, Computer Science Department, Keele University (TR/SE-0401) and National ICT Australia Ltd. (0400011T.1)
3. Van der Aalst, W. M. et al (2012), Process Mining Manifesto, In: BPM 2011 Business Process Management Workshops, Lecture Notes in Business Information Processing, Vol 99., Daniel F., Barkaoui K., Dustdar S. (eds), pp. 169-194, doi: 10.1007/978-3-642-28108-2\_19
4. Khan, Khalid, S., ter Riet, Gerben., Glanville, Julia., Sowden, Amanda, J. and Kleijnen, Jo. (eds) Undertaking Systematic Review of Research on Effectiveness, CRD's Guidance for those Carrying Out or Commissioning Reviews, CRD Report Number 4 (2nd Edition), NHS Centre for Reviews and Dissemination, University of York, ISBN 1 900640 20 1, March 2001
5. Alderson P., Green S, Higgins JPT. (eds) Cochrane Reviewers' Handbook 4.2.2 [updated March 2004]. In: The Cochrane Library, Issue 1, 2004. Chichester, UK: John Wiley & Sons, Ltd
6. R'bigui, H. & Cho, C. (2017) The state-of-the-art of business process mining challenges, *Int. J. of Bus. Process. Integr. and Man.*, doi: 10.1504/IJBPI.2017.088819
7. Dakic, D., Stefanovic, D., Cosic, I., Lolic, T. & Medojevic, M. (2018) Business process mining application: a literature review, In: Proceedings of the 29th DAAAM International Symposium, pp.0866-0875, B. Katalinic (Ed.), Published by DAAAM International, ISBN 978-3-902734-20-4, ISSN 1726-

9679, Vienna, Austria doi: 10.2507/29th.daaam.proceedings.125

**8.** Tiwari, A., Turner, C. J. & Majeed, B. (2008) A review of business process mining: state-of-the-art and future trends, *Bus. Process. Man. J.*, doi: 10.1108/14637150810849373

**9.** Van Der Aalst, W.M. (2012) Process mining: Overview and opportunities, *ACM Trans. on Man. Inf. Syst.*, doi:10.1145/2229156.2229157

**10.** Lu, Xixi & Nagelkerke, Marijn & van de Wiel, Dennis & Fahland, Dirk. (2015) Discovering Interacting Artifacts from ERP Systems, *IEEE Trans. on Serv. Comput.*, doi: 10.1109/TSC.2015.2474358

**11.** Santana Calvo, H.A. (2017) Artifact-Centric Log Extraction for Cloud Systems, Master's thesis, Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherland

**12.** H. J. Nooijen, Erik & F. van Dongen, Boudewijn & Fahland, Dirk (2012) Automatic Discovery of Data-Centric and Artifact-Centric Processes, In: *Lecture Notes in Business Information Processing*, doi: 10.1007/978-3-642-36285-9\_36

**13.** Piessens, D.A.M. (2011) Event Log Extraction from SAP ECC 6.0, Master's thesis, Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherland

**14.** J.C.A.M. Buijs (2010) Mapping Data Sources to XES in a Generic Way, Master's thesis, Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherland

**15.** Ricardo Pérez-Castillo, Barbara Weber, Jakob Pinggera, Stefan Zugal, Ignacio García-Rodríguez de Guzmán & Mario Piattini (2011) Generating event logs from non-process-aware systems enabling business process mining, *Enterpr. Inf. Syst.*, doi: 10.1080/17517575.2011.587545

**16.** González López de Murillas, E., Reijers, H.A. & van der Aalst, W.M.P. (2019) Connecting databases with process mining: a meta model and toolset, *Soft. Syst. Model.*, doi: 10.1007/s10270-018-0664-7

**17.** Rodríguez C, Engel R, Kostoska G, Daniel F, Casati F, Aimar M (2012) Eventifier: extracting process execution logs from operational databases. In: *Proceedings of the demonstration track of BPM 2012*, vol. 940, pp. 17–22

**18.** Esposito, P. M., Vaz, M. A. A., Rodrigues, S. A., & De Souza, J. M. (2013). MANA: Identifying and mining unstructured business processes. In: *Lecture Notes in Business Information Processing*, doi: 10.1007/978-3-642-36285-9-20

**19.** Fliegner, W. (2014) Extracting process-related information from ERP systems for process discovery, *Res. in Logist. & Prod.*, Vol. 4, No. 4, pp. 315-329

**20.** Bernardi, M. L., Cimitile, M., & Mercaldo, F. (2018) Cross-Organisational Process Mining in Cloud Environments, *J. of Inf. & Knowl. Manag.*, doi: 10.1142/s0219649218500144

**21.** Jans, M. (2017) From relational database to valuable event logs for process mining purposes : a procedure, doi: 10.13140/RG.2.2.11343.69289

**22.** Li, Guangming & González López de Murillas, Eduardo & M. de Carvalho, Renata & Van der Aalst, Wil M. P. (2018) Extracting Object-Centric Event Logs to Support Process Mining on Databases, In:

Information Systems in the Big Data Era, doi: 10.1007/978-3-319-92901-9\_16

**23.** Calvanese D., Kalayci T.E., Montali M., Tinella S. (2017) Ontology-Based Data Access for Extracting Event Logs from Legacy Data: The onprom Tool and Methodology, In: Lecture Notes in Business Information Processing, doi: 10.1007/978-3-319-59336-4\_16

**24.** Henny Selig (2017) Continuous Event Log Extraction for Process Mining, Degree project in information and communication technology, KTH Royal Institute of Technology, School of Information and Communication Technology, Stockholm, Sweden