

BACHELORARBEIT

Technische Hochschule Deggendorf

Hochschule für angewandte Wissenschaften

Fakultät Angewandte Naturwissenschaften und Wirtschaftsingenieurwesen

Studiengang Wirtschaftsingenieurwesen

Herausforderungen und Potentiale von Process Mining

(Challenges and potentials of process mining)

Bachelorarbeit zur Erlangung des akademischen Grades:

Bachelor of Engineering (B.Eng.)

vorgelegt von: Mathias Kirsch, Deggendorf,
Matrikelnummer 646315

Prüfer: Prof. Dr. Stephan Scheuerer

Deggendorf, den 28.08.2019

Erklärung

Name d. Studierenden: Mathias Kirsch

Name d. Betreuers: Prof. Dr. Stephan Scheuerer

Thema der Bachelorarbeit: Herausforderungen und Potentiale von Process Mining

1. Ich erkläre hiermit, dass ich die Bachelorarbeit selbstständig verfasst, noch nicht anderweitig für Prüfungszwecke vorgelegt, keine anderen als die angegebenen Quellen oder Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate als solche gekennzeichnet habe.

Deggendorf, den _____ Unterschrift d. Studierenden: _____

2. Ich bin damit einverstanden, dass die von mir angefertigte Bachelorarbeit über die Bibliothek der Hochschule einer breiteren Öffentlichkeit zugänglich gemacht wird.

☐ Ja

☐ Nein

Deggendorf, den _____ Unterschrift d. Studierenden: _____

Ich erkläre und stehe dafür ein, dass ich alleiniger Inhaber aller Rechte an der Bachelorarbeit, einschließlich des Verfügungsrechts über Vorlagen an beigefügten Abbildungen, Plänen o. ä., bin und durch deren öffentliche Zugänglichmachung weder Rechte und Ansprüche Dritter noch gesetzliche Bestimmungen verletzt werden.

Bei Einverständnis des Verfassers mit einer Zugänglichmachung der Bachelorarbeit vom Betreuer auszufüllen:

3. Eine Aufnahme eines Exemplars der Bachelorarbeit in den Bestand der Bibliothek und die Ausleihe des Exemplars wird

☐ befürwortet

☐ nicht befürwortet

Deggendorf, den _____ Unterschrift d. Betreuenden: _____

Inhaltsverzeichnis

Abbildungsverzeichnis	i
Tabellenverzeichnis	ii
Abkürzungsverzeichnis	iii
1 Einführung	1
2 Theoretische Fundierung des Process Mining	4
2.1 Einordnung von Process Mining in das Wissenschaftsspektrum	4
2.2 Arten und Perspektiven des Process Mining	6
2.3 Begrifflichkeiten beim Process Mining	7
2.4 Einsatzgebiete des Process Mining	9
2.5 Praxisbeispiel aus der Jahresabschlussprüfung	11
2.6 Process Mining-Algorithmen	11
2.6.1 Qualitätskriterien von Process Mining-Algorithmen	13
2.6.2 Inductive Miner-Algorithmus	14
3 Methodisches Vorgehen beim Process Mining	19
3.1 Planung und Modellierung	19
3.2 Datenextraktion aus Quellsystem	20
3.3 Datenverarbeitung und Event Log-Generierung	20
3.4 Mining und Analyse	21
3.5 Evaluierung	22
3.6 Umsetzung von Prozessverbesserungen	22
4 Herausforderungen bei der Anwendung von Process Mining	23
4.1 Schaffung der Voraussetzungen für Process Mining	23
4.2 Zielfestlegung	24
4.3 Problematiken der einzelnen Prozessschritte	25
4.3.1 Planung und Modellierung	25
4.3.2 Datenextraktion aus Quellsystem	27
4.3.3 Datenverarbeitung und Event Log-Generierung	27
4.3.4 Mining und Analyse	27
4.3.5 Evaluierung	28
4.3.6 Umsetzung von Prozessverbesserungen	28
4.4 Ausgereiftheit der Algorithmen	29
4.5 Einbindung von transaktionaler Datenanalyse	31
4.6 Einhaltung von Datenschutzbestimmungen	33
5 Potentiale von Process Mining	35
5.1 Vorstufe zu innovativen Technologien	35

5.2	Anwendung in Echtzeit zur operativen Unterstützung	37
5.3	Machine-Learning-Algorithmen und Künstliche Intelligenz (KI)	38
5.4	Organisationsübergreifendes Process Mining	38
6	Anbieter von Process Mining-Tools.....	39
6.1	Process Mining Markt	39
6.2	Führende Produkte	40
6.2.1	Celonis	40
6.2.2	Fluxicon.....	41
6.2.3	Lana Labs.....	43
6.2.4	Process Analytics Factory	44
6.2.5	Process Mining Group	45
6.3	Vergleich.....	47
7	Zusammenfassung und Fazit	50
	Anhang	iv
	Literaturverzeichnis.....	ix

Abbildungsverzeichnis

Abbildung 1: Prognose zum Volumen der jährlich generierten digitalen Datenmenge in den Jahren 2018 und 2025 (in Zettabyte) [IDC18]	1
Abbildung 2: Process Mining als Bindeglied zwischen Data Science und Process Science [Aal16, S.18].....	5
Abbildung 3: Arten von Process Mining [Aal16, S. 32]	6
Abbildung 4: Auszug eines Event Logs (eigene Darstellung)	8
Abbildung 5: Einführung von Process Mining-Typen zwischen 2017 und 2020 [Ker19].....	10
Abbildung 6: Veränderung des Einsatzzweckes von Process Mining zwischen 2018 und 2020 [Ker19].....	10
Abbildung 7: Dasselbe Event Log angewendet auf unterschiedliche Algorithmen [Roz07]...	12
Abbildung 8: Die vier Qualitätskriterien der Prozessentdeckung [Aal16, S. 189]	13
Abbildung 9: Prozessbaumdarstellung gemäß Beispiel (eigene Darstellung gemäß ProM-Plattform)	17
Abbildung 10: Prozessbaumdarstellung mit exclusive-choice-cut zwischen Aktivität e und tau-Aktivität (eigene Darstellung gemäß ProM-Plattform).....	18
Abbildung 11: Methodisches Vorgehen beim Process Mining [Eck15]	19
Abbildung 12: Phasen der Datenverarbeitung, Mining und Analyse [Eck15]	21
Abbildung 13: Einzustellende Parameter zum Heuristics Miner in ProM	30
Abbildung 14: Ergebnisse von dab-Prüfschritt "Invoice date before Purchase Order date" ..	32
Abbildung 15: Process Mining Anbieter [Rob18]	39
Abbildung 16: Reiter "Overview" von Celonis Snap	41
Abbildung 17: Oberfläche des Disco-Tools von Fluxicon	42
Abbildung 18: Reiter "Model" von LANA Process Mining	44
Abbildung 19: PAFnow in Microsoft Power BI	45
Abbildung 20: Reiter "Actions" von ProM	46
Abbildung 21: Process Mining in BPM [Sch18a]	iv
Abbildung 22: Relevante SAP-Tabellen für P2P-Prozess [dab:GmbH]	v
Abbildung 23: Zukunftsthemen im IT-Umfeld [Cap19].....	vi
Abbildung 24: Reiter "Statistics" des Disco-Tools	vi
Abbildung 25: Reiter "Cases" des Disco-Tools.....	vii
Abbildung 26: Reiter „Rework Rate“ von Celonis Process Mining	vii
Abbildung 27: Reiter "Statistics" von LANA Process Mining.....	vii
Abbildung 28: Reiter "View" von ProM	viii
Abbildung 29: Inductive visual Miner in ProM.....	viii

Tabellenverzeichnis

Tabelle 1: Vereinfachte Darstellung der relevanten Tabellen aus SAP-MM für P2P-Prozess	
[dab:GmbH]	25
Tabelle 2: Vereinfachte Darstellung der relevanten Tabellen aus SAP-FI für P2P-Prozess	
[dab:GmbH]	26

Abkürzungsverzeichnis

BI	Business Intelligence
BPM	Business Process Management
BPMN	Business Process Management Notation
CRM	Customer-Relationship-Management
CSV	Comma-separated values
DSGVO	Datenschutz-Grundverordnung
dt.	deutsch
DTO	Digital Twin of an Organization
ERP	Enterprise-Ressource-Planning
GRC	Governance, Risk & Compliance
ID	Identifikator
IEEE	Institute of Electrical and Electronics Engineers
KI	Künstliche Intelligenz
KPIs	Key Performance Indikatoren
MXML	Extensible Markup Language-basierte Markup-Sprache
P2P	Purchase-to-Pay
PAF	Process Analytics Factory
ROI	Return on Investment
RPA	Robotic Process Automation
SAP-FI	SAP-Modul Finanzen
SAP-MM	SAP-Modul Materialmanagement
XES	Extensible Event Stream

1 Einführung

In der heutigen Zeit der rasch fortschreitenden Technologien, Digitalisierung und Vernetzung gewinnen Daten als wertvolle Rohstoffe zunehmend an Bedeutung. Daten repräsentieren „elektronisch gespeicherte Zeichen, Angaben, Informationen“ [Dud19]. Organisationen, wie Unternehmen und öffentliche Institutionen, haben die Zeichen der Zeit erkannt und gehen mit Daten wertschätzender um als noch vor einigen Jahren, als Daten noch als Nebenprodukt abgetan wurden und der Fokus zunächst rein auf das Sammeln von Massendaten gelegt wurde, ohne Kontext zur weiteren Verarbeitung.

Durch die digitale Transformation werden Daten nahezu überall und zu jeder Zeit erhoben. Auch im Unternehmen fallen bei jedem Arbeitsschritt Daten an, welche durch Informationssysteme aufgezeichnet und gespeichert werden. Diese Protokollierung in Informationssystemen dient zur Unterstützung operativer Vorgänge und führt dazu, dass sich eine enorme Zunahme der weltweiten Datenmenge beobachten lässt. Jährlich steigt diese um etwa 30 Prozent an. Wie in Abbildung 1 ersichtlich, werden bis zum Jahr 2025 Größenordnungen von 175 Zettabytes (Zettabyte ist eine Zahl mit 21 Nullen) prognostiziert. Das entspricht einer Verzehnfachung im Betrachtungszeitraum zwischen 2016 und 2025 [IDC18], [Joe18].

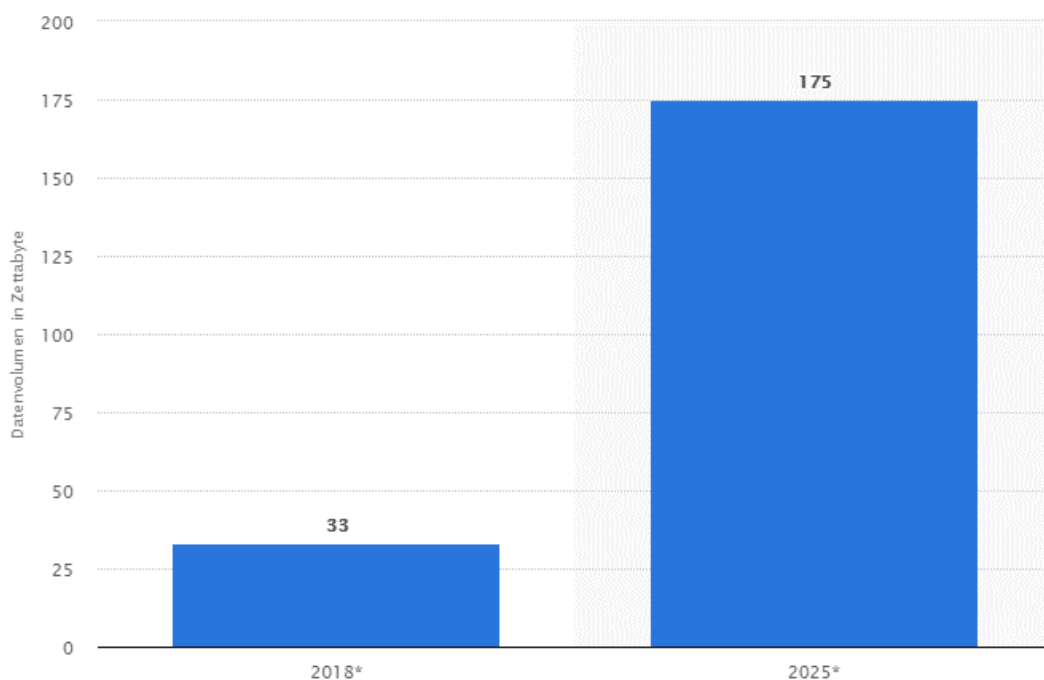


Abbildung 1: Prognose zum Volumen der jährlich generierten digitalen Datenmenge in den Jahren 2018 und 2025 (in Zettabyte) [IDC18]

Die Nutzung von Informationstechnologie erzielt heute keine entscheidenden Wettbewerbsvorteile mehr. Wird jedoch keine Informationstechnologie genutzt, so entsteht ein enormer

Nachteil. Auf den Zug der Digitalisierung nicht aufzuspringen ist längst zum Risikofaktor geworden. Einen großen Beitrag zur bereits erwähnten Zunahme der digitalen Datenmenge leisten Organisationen, die sich als Vorreiter in Sachen Digitalisierung positionieren. Denn die digitale Transformation hat auch Einfluss auf Geschäftsprozesse. Mittlerweile gibt es in Organisationen eine hohe Anzahl an datenintensiven Vorgängen. Aus diesem Grund ist der Umgang mit Daten in vielen Unternehmen und Betrieben der öffentlichen Verwaltung ein wichtiger Ansatzpunkt zur Ableitung von Handlungsempfehlungen geworden. In kompetitiven Globalisierungsmärkten werden rationales Handeln und Kostenüberlegungen zunehmend wichtiger. Mit einem zielgerichteten und effizienten Prozessmanagement lassen sich Wettbewerbsvorteile generieren. In der heutigen, sich schnell wandelnden Unternehmenswelt lässt sich aufgrund der Dynamisierung folgerichtig eine Hinwendung zum agilen Geschäftsprozess- und Changemanagement beobachten. Eine effiziente Gestaltung der Prozesse ist für viele, in einem globalen Wettbewerb stehende Unternehmen von enormer Bedeutung geworden [Beh16, S. 108].

Beide genannten Punkte, die Datenzunahme sowie das Prozessmanagement, werden durch eine zukunftsweisende toolbasierte Methode vereint, das Process Mining. Hierbei können die in Informationssystemen vorhandenen Ereignisprotokolldaten für die Gewinnung von Wissen nutzbar gemacht werden. Vor allem voll integrierte Informationssysteme, in denen alle Prozessschritte abgebildet werden, nehmen heutzutage eine zentrale Rolle im Tagesgeschäft von Organisationen ein. Sie bilden die eigenen Geschäftsprozesse systemseitig ab und sammeln und speichern alle Unternehmensdaten. Dies begünstigt die Anwendung von Process Mining. Jedoch bedingt das reine Sammeln von Daten noch keine Prozessverbesserungen. Vielmehr stellt es den initialen Startpunkt eines Process Mining-Vorhabens dar [Aal15].

Die Vorstellung, die Realität widerspiegelnde Protokolldaten zur Erstellung, Prüfung, Verbesserung und wirksamen Analyse von bestehenden Prozessmodellen zu nutzen, macht Process Mining zu einem mächtigen Tool, dessen Akzeptanz in den Führungsetagen der Unternehmen weiter steigt. Einerseits müssen die Prozessmodelle, die in der Praxis Anwendung finden, verifiziert werden. Zum anderen ist eine ganzheitliche Betrachtung der einzelnen Prozessschritte und deren Abhängigkeiten untereinander im Fokus des Prozessmanagements anzusiedeln. Dabei ist Prozessoptimierung nicht mehr ein einmaliges Projekt, sondern wird als kontinuierliches Vorhaben verstanden, um den sich agil und rasch verändernden Anforderungen begegnen zu können.

Aufgrund der hier kurz dargestellten Entwicklungen widmet sich diese Arbeit der Aufarbeitung der wesentlichen Herausforderungen und Potentiale von Process Mining. Die Aufarbeitung der Thematik erfordert zunächst eine theoretische Fundierung des zugrundeliegenden Tools. Zu

Beginn wird eruiert, wie sich Process Mining im Wissenschaftsspektrum positioniert und welche Fachbegriffe beim Process Mining verwendet werden. Mögliche Einsatzgebiete werden erläutert, wobei eines anhand des Beispiels der Jahresabschlussprüfung konkretisiert wird. Daraufgehend wird näher auf die für die Erstellung von Prozessmodellen unabdingbaren Algorithmen eingegangen sowie deren zu beachtenden Qualitätskriterien. Exemplarisch wird der Process Mining-Algorithmus Inductive Miner vorgestellt. Um größtmöglichen Nutzen aus einem Process Mining-Vorhaben zu erlangen, muss weiterhin die methodische Vorgehensweise bei der Durchführung von Process Mining eingehend betrachtet werden. Natürlich gibt es auch bei Anwendung dieses Tools einige Herausforderungen sowohl für die Software-Anbieter als auch für die Endanwender, auf die ausführlich eingegangen wird. Zukunftsträchtige Ansätze von Potentialen des Process Mining werden ebenso thematisiert. Die aktuelle Marktsituation wird eingehend betrachtet, da sich der Bedarf an Process Mining-Software ausweitete. Folgerichtig ist eine wachsende Anzahl an Anbietern festzustellen, welche Lösungen dieser toolbasierten Methode entwickeln. Aus diesem Grund folgt daraufhin die Vorstellung und der Vergleich von Anbietern anhand eines exemplarischen Beispieldatensatzes. Eine Zusammenfassung und ein Fazit runden die Arbeit ab.

2 Theoretische Fundierung des Process Mining

Um einer Aufarbeitung der Thematik gerecht zu werden, welche dieser Ausarbeitung zu Grunde liegt, müssen zunächst Definitionen und Erläuterungen der zentralen Begrifflichkeiten erfolgen sowie theoretische Aspekte im Zusammenhang mit Process Mining aufgearbeitet werden. Process Mining umfasst „Techniken, Werkzeuge und Methoden, um reale Prozesse (d.h. nicht bloß vermutete Prozesse) zu entdecken, zu beobachten und zu verbessern, in dem Wissen von Ereignislogs, die in heutigen (Informations-)Systemen verfügbar sind [...]“ [IEE12]. Process Mining zeichnet sich durch die Darstellung von Prozessen mittels Prozessgraph und dessen Analyse aus. Auf Grundlage realer Protokolldaten können so Engpässe und Schwachstellen in den eigenen Prozessen entdeckt und im Anschluss verbessert werden. Ebenso können die Neugestaltung von Prozessmodellen sowie die Steuerung operativer Geschäftsprozesse durch Process Mining unterstützt werden [Ker19].

2.1 Einordnung von Process Mining in das Wissenschaftsspektrum

Process Mining wird in der Literatur mit vielen verschiedenen Begriffen in Verbindung gebracht. Wil van der Aalst sieht Process Mining als das verbindende Element zwischen Data Science und Process Science an (Abbildung 2), bei dem datenorientierte Ansätze mit prozessorientierten Ansätzen vereint werden [Aal16, S. 15ff.]. Dieser Positionierung muss eine genauere Betrachtung der beiden genannten Wissenschaftsgebiete folgen.

Data Science umfasst ein weites Feld und bezieht mehrere Wissenschaftsdisziplinen ein, mit dem Anspruch, aus vorliegenden Daten Wissen zu generieren. Dabei ist es unerheblich, in welcher Form die Daten vorliegen und auf welche Art das Wissen realisiert wird [Aal16, S. 10]. Die Aktualität und Wichtigkeit von Data Science spiegelt sich ferner durch die Etablierung eines neuen Berufsfeldes wider. „Der Data Scientist zählt zu den attraktivsten Berufen des 21. Jahrhunderts“, so die Fraunhofer-Gesellschaft, und wird dank der explosiven Entwicklung von Datenmengen und der damit einhergehenden steigenden Komplexität stark nachgefragt [Fra18]. Der Bezug von Process Mining zu Data Mining wird häufig thematisiert. Der grundlegende Unterschied ist die Sichtweise der beiden Technologien. Anders als beim Data Mining legt das Process Mining den Fokus auf die Muster in den durch die Unternehmensdaten repräsentierten Prozesse sowie daraus ableitbare Möglichkeiten zur Prozessverbesserung. Das Data Mining fokussiert sich hingegen auf Muster in den Daten an sich. Das Ziel von Data Science ist die Erlangung von Wissen aus Daten. Die Unterscheidung von Daten, Information und Wissen ist dabei wesentlich und wird unter anderem in [Bod06, S. 1f.] und [Lau16, S. 639ff.] erläutert. Daten bestehen aus Zeichen mit entsprechender Syntax. Daten angereichert mit einer Bedeutung beziehungsweise einem Kontext werden als Informationen bezeichnet.

Informationen beantworten die Fragen „Wer?“, „Was?“, „Wo?“, „Wann?“ und „Wie viele?“. Werden Informationen mit dem Erfahrungsschatz von Menschen verknüpft, so entsteht Wissen. Erst mithilfe von Wissen lassen sich Fragen nach dem „Wie?“ und nach dem „Warum?“ beantworten und zielgerichtet zur Aufgabenlösung einsetzen.

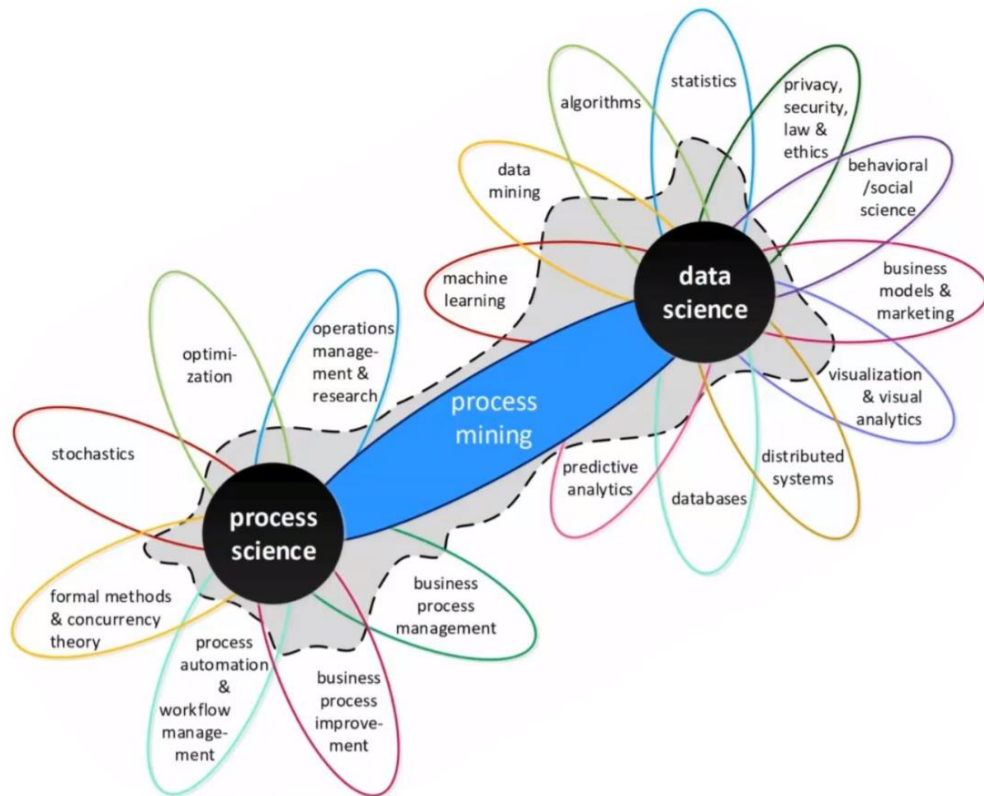


Abbildung 2: Process Mining als Bindeglied zwischen Data Science und Process Science [Aal16, S.18]

Die Prozessperspektive wird vom klassischen Data Science nur unzureichend betrachtet. An dieser Stelle setzt Process Science an. Process Science impliziert die wissenschaftliche Verzahnung von Betriebswirtschaftslehre und Informationstechnologie, welche sich mit der Durchführung und Verbesserung von organisationalen, modellgetriebenen End-to-End-Prozessen befasst [Aal16, S. 15] und detaillierte Einblicke zur Prozessausführung liefert [Beh16, S. 127f.]. Als Bindeglied beider Ansätze stellt Process Mining eine wichtige, gerechtfertigte und eigenständige Wissenschaft dar.

Innerhalb des Business Process Management (BPM)-Lebenszyklus kann Process Mining vor allem in den datengetriebenen Phasen Anwendung finden. Verglichen mit der früheren, rein modellgetriebenen Sichtweise [Aal16, S. 44f.], bei der es häufig systembedingte Lücken bei der Verbindung zwischen den Daten und dem Prozessdesign gab, bietet das Process Mining einen umfassenderen Einsatz. Zudem gewährt es die Chance, den BPM-Lebenszyklus endgültig zu schließen [Aal16, S. 31]. Prof Dr. Scheer beschreibt in [Sch18a] den Einsatz von Process Mining als idealen Ansatz zur taktischen Vervollständigung des eher strategischen

BPM-Kreises (siehe Anhang Abbildung 21). So wird der bekannte Top-Down-Ansatz des BPM mit einem sogenannten Bottom-up-Ansatz des Process Mining ideal kombiniert¹. In der einschlägigen Literatur wird die Verbindung von Process Mining zu weiteren Fachbegriffen behandelt, unter anderem in [Aal11] und [Aal16]. In dieser Arbeit werden lediglich die Gebiete erwähnt, die am häufigsten mit Process Mining in Verbindung gebracht werden.

2.2 Arten und Perspektiven des Process Mining

Process Mining-Verfahren lassen sich generell in drei Hauptarten einteilen, bei denen jeweils die Eingabe- und Ausgabeformen sowie die Beziehungen eines Event Logs zu einem Prozessmodell variieren. Dies sind die Prozesserkennung, die Konformitätsprüfung und die Prozessenerweiterung [Aal16, S. 33ff.], [IEE12]. Abbildung 3 stellt die Zusammenhänge visualisiert dar.

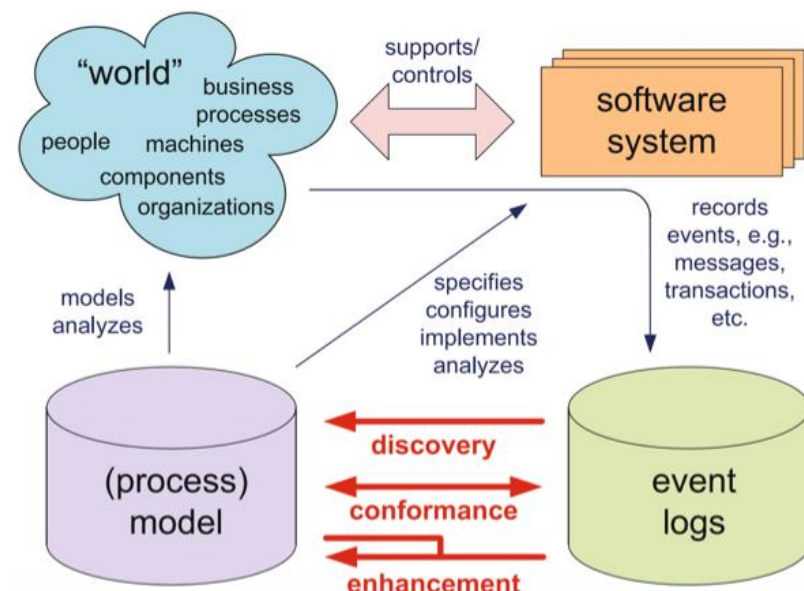


Abbildung 3: Arten von Process Mining [Aal16, S. 32]

- **Prozesserkennung** (Process Discovery): Es handelt sich um die bekannteste Process Mining-Methode, bei der Ereignislogdaten als Eingabe dienen und daraus ein reales Prozessmodell erzeugt und abgeleitet wird. Dies geschieht, ohne auf vorherige Informationen zurückzugreifen. So kann ein Prozessmodell allein aufgrund realer Event Log-Daten erstellt werden.
- **Konformitätsprüfung** (Conformance Check): Einem bereits bestehenden Prozessmodell werden reale Ereignislogdaten desselben Prozesstypen gegenübergestellt.

¹ <https://fluxicon.com/technology/>

Dies ermöglicht eine Prüfung auf Übereinstimmung zwischen der in den Logdaten dokumentierten Realität und dem theoretischen Modell, den Richtlinien oder Arbeitsanweisungen. Das bereits vorliegende theoretische Modell kann entweder durch vormalige Prozesserkennung oder Handfertigung vorliegen.

- **Prozesserweiterung** (Process Enhancement): Ein bestehendes Prozessmodell wird mit den gewonnenen Informationen aus den Ereignislogdaten angepasst. Anpassen impliziert sowohl eine Verbesserung als auch eine Erweiterung eines bestehenden Prozessmodells. Diese Art wird auch als Leistungsanalyse beschrieben [Aal15].

Neben diesen Arten wird die Anwendung von Process Mining anhand verschiedener Perspektiven eingestuft [Aal16, S. 34], [IEE12]. Die entsprechenden Perspektiven und Sichten auf die Daten werden durch die Entwicklung verschiedener Process Mining-Algorithmen (Kapitel 2.6) realisiert. In der Praxis überschneiden sich die Perspektiven.

- **Kontrollflussperspektive**: Alle Ausführungspfade und deren Aktivitätenabfolge werden ermittelt. Die Frage „Wie laufen die Prozesse?“ wird beantwortet. Process Mining wird zumeist mit dieser Perspektive in Verbindung gebracht oder gar darauf reduziert.
- **Organisationsperspektive**: Die Ressourceninformationen und die damit einhergehende Zuordnung von Akteuren zu den zugehörigen Rollen und Organisationseinheiten werden betrachtet. Die daraus resultierenden Beziehungen werden somit sichtbar gemacht. Die Frage „Wer ist wann und wie an den Prozessen beteiligt?“ wird beantwortet.
- **Fallperspektive**: Eigenschaften einzelner Fälle werden näher betrachtet, um bestimmte Muster und Auffälligkeiten deuten zu können. Die Frage, welche beantwortet werden soll, lautet: „Was passiert innerhalb eines Cases?“.
- **Zeitperspektive**: Diese Perspektive liefert Erkenntnisse zu Auftretungshäufigkeiten und Zeitrahmen von Ereignissen. „Wie häufig und zu welchen Zeitpunkten werden Prozesse durchlaufen?“ ist eine Frage, die damit beantwortet werden kann.

2.3 Begrifflichkeiten beim Process Mining

Zum besseren Verständnis des Process Mining ist es unerlässlich, die spezifischen Begrifflichkeiten zu kennen und zu verstehen. Im Folgenden sind die Fachbegriffe, welche beim Process Mining Gebrauch finden, erläutert [Aal16, S. 128ff.].

- **Event** (deutsch (dt.) Ereignis): Ein Event beschreibt eine atomare Aktivität mit den Mindestanforderungen Case Identifikator (ID), Aktivität und Zeitstempel und stellt die kleinste Dateneinheit im Process Mining dar. Die Case ID ist die eindeutige Vorgangsnummer für einen Case. Dies kann eine Auftragsnummer, Kundennummer, Bestellnummer oder eine Personalnummer sein. Eine Aktivität stellt eine konkret ausgeführte

Prozesstätigkeit beziehungsweise einen Prozessschritt dar. Als Beispiel seien „Create Purchase Requisition“ oder „Create Purchase Order Line Item“ genannt. Ein Zeitstempel enthält Datum und Uhrzeit, wann eine Aktivität begonnen oder/und beendet wurde. Mindestens ein Zeitstempel ist erforderlich.

- **Case** (dt. Fall/Prozessinstanz): Dieser besteht aus mehreren Events und stellt einen Prozessablauf dar. Alle Prozessschritte innerhalb eines Cases weisen stets dieselbe Case ID auf.
- **Trace** (dt. Spur): Ein Trace bedingt eine Reihe von Events innerhalb einer Kontrollflussrelation. Dabei ist der chronologische Ausführungsablauf der Events einer Prozessinstanz bereits berücksichtigt, zum Beispiel <a,b,c,d,e>.
- **Event Log** (dt. Ereignisprotokoll): Mehrere prozessverwandte Traces werden in einem Log zusammengefasst. Event Logs bilden die Grundlage für die Anwendung von Process Mining.

Abbildung 4 zeigt einen Auszug von einem Ereignisprotokoll.

Attribute							
Case ID	Aktivität	Start Zeitstempel	Ende Zeitstempel	Resource	Role		
Instanz	1	Create Purchase Requisition	2011/01/01 00:00:00.000	2011/01/01 00:37:00.000	Kim Passa	Requester	→
	1	Create Request for Quotation Requester	2011/01/01 05:37:00.000	2011/01/01 05:45:00.000	Kim Passa	Requester	→
	1	Analyze Request for Quotation	2011/01/01 06:41:00.000	2011/01/01 06:55:00.000	Karel de Groot	Purchasing Agent	→
	1	Send Request for Quotation to Supplier	2011/01/01 11:43:00.000	2011/01/01 12:09:00.000	Karel de Groot	Purchasing Agent	→
	1	Create Quotation comparison Map	2011/01/01 12:32:00.000	2011/01/01 16:03:00.000	Magdalena Predutta	Purchasing Agen	
	1	Analyze Quotation comparison Map	2011/01/01 22:44:00.000	2011/01/01 23:13:00.000	Immanuel Karagianni	Requester	...
	1	Choose best option	2011/01/01 23:13:00.000	2011/01/01 23:13:00.000	Tesca Lobes	Requester	
	1	Settle conditions with supplier	2011/01/02 01:22:00.000	2011/01/02 09:20:00.000	Francois de Perrier	Purchasing Agent	
	1	Create Purchase Order	2011/01/02 09:58:00.000	2011/01/02 10:10:00.000	Karel de Groot	Purchasing Agent	
	1	Confirm Purchase Order	2011/01/02 14:09:00.000	2011/01/02 14:43:00.000	Sean Manney	Supplier	
	1	Deliver Goods Services	2011/01/02 20:49:00.000	2011/01/03 03:37:00.000	Sean Manney	Supplier	
	1	Release Purchase Order	2011/01/03 11:20:00.000	2011/01/03 11:21:00.000	Elvira Lores	Requester	
	1	Approve Purchase Order for payment	2011/01/03 19:09:00.000	2011/01/03 19:10:00.000	Karel de Groot	Purchasing Agent	
	1	Send invoice	2011/01/04 00:54:00.000	2011/01/04 00:54:00.000	Kiu Kan	Supplier	
	1	Release Supplier's Invoice	2011/01/04 15:08:00.000	2011/01/04 15:13:00.000	Karalda Nimwada	Financial Manager	
	1	Authorize Supplier's Invoice payment	2011/01/04 15:13:00.000	2011/01/04 15:13:00.000	Karalda Nimwada	Financial Manager	
	1	Pay invoice	2011/01/04 15:22:00.000	2011/01/04 15:31:00.000	Pedro Alvares	Financial Manager	
	2	Create Purchase Requisition	2011/01/01 00:16:00.000	2011/01/01 00:29:00.000	Immanuel Karagianni	Requester	
	2	Create Request for Quotation Requester	2011/01/01 08:16:00.000	2011/01/01 08:26:00.000	Alberto Duport	Requester	
	2	Analyze Request for Quotation	2011/01/01 09:34:00.000	2011/01/01 09:38:00.000	Karel de Groot	Purchasing Agent	

Abbildung 4: Auszug eines Event Logs (eigene Darstellung)

Neben den Mindestanforderungsattributen Case ID, Aktivität und Zeitstempel sind darüber hinaus oftmals zusätzliche optionale Attribute in Datentabellen gespeichert. Diese können den

Event Logs angereichert werden und erlauben somit umfangreichere Filterungen und Analysen. Je nach Anwendungsfall sind Ressourcen, Rollen, Kosten, Anfangs- und Endzeitstempel sinnvoll und bei manchen Process Mining-Algorithmen auch notwendig. Ressourcen sind beispielsweise Angaben über eine Person, die eine Aktivität ausgeführt hat. Eine Rolle kann die Position einer Person innerhalb eines Netzwerks beschreiben. Personelle Abhängigkeiten innerhalb eines Prozesses können mit diesen Angaben analysiert werden. Kosten und Anfangs- und Endzeitstempel lassen eine genauere Effizienzbetrachtung zu, beispielsweise Analysen zu Lagerkosten pro Lagerplatz, Engpässe, Durchlauf-, Liege- und Wartezeiten. Diese können Process Mining-Vorhaben qualitativ aufwerten und mehr Betrachtungsmöglichkeiten bieten. Im Zuge dessen muss immer auch die Performance bedacht werden, welche nachweislich mehr belastet wird, je mehr Attribute und Filterungen verwendet werden. Diese Problematik hat sich jedoch mit heutigen fortschrittlicheren und robusteren Algorithmen deutlich abgeschwächt.

2.4 Einsatzgebiete des Process Mining

Dem Market Guide for Process Mining [Ker19] zufolge ist der Einsatz von Process Mining-Software weiterhin von der Prozessentdeckung (Process Discovery) geprägt. Man sehe aber gemäß Abbildung 5 einen signifikanten Trend zu einer stärkeren Fokussierung auf die Konformitätsprüfung (Conformance Check) und Prozesserweiterung (Process Enhancement). Die Studie von Gartner [Ker19] liefert weitere Ergebnisse zum Anwendungsbereich der Process Mining-Software. Weiterhin wird ein Großteil der Software mit Blick auf Geschäftsprozessverbesserungen angewendet. Dieser Anteil ist jedoch rückläufig. Der Anteil an Implementierungen für Prozessautomatisierungen und Ansätze digitaler Transformation nehmen hingegen zu (Abbildung 6).

Die Konformitätsprüfung wird für Organisationen in Zeiten von Governance, Risk & Compliance (GRC) zur Absicherung immer wichtiger. Die Konformitätsprüfung ist von besonderer Bedeutung im Hinblick auf die Prüfung der Einhaltung von internen und externen Regelungen. Die Qualität von Prozessen wird damit gesichert und die Gesetzeskonformität unterstützt. Externe und interne Regelungen müssen zwingend eingehalten werden, um in Zeiten von deutlich strengerer Nachverfolgung möglichen Strafen zu entgehen und um den Prozessausführungen innerhalb des durch die Interessensgruppen definierten Rahmens zu entsprechen [Aal16, S. 244f.], [Aal16, S. 268].

Die Durchführung von Process Mining ist unabhängig vom Quellsystem möglich. Typische Quellsysteme sind unter anderem die in einer Vielzahl am Markt positionierten Enterprise-Ressource-Planning-Systeme (ERP-Systeme), Customer-Relationship-Management-Systeme (CRM-Systeme) und Lagerverwaltungssysteme, aber auch übergreifende Systeme. Ne-

ben den Mindestanforderungen an ein Event (siehe Abschnitt 2.3) ist ein transaktional-prozessualer Vorgang, eine qualitativ hohe und quantitativ ausreichende Datenbasis zur Verarbeitung des Algorithmus und Erstellung eines Prozessgraphen, erforderlich. Diese Tatsache macht Process Mining zu einem gewinnbringenden Tool für nahezu jede Organisationsform und Branche. In [Gar19, S. 46f.] werden die Anwendungs- und Forschungsbereiche ausführlich beschrieben. Demnach zählen das Gesundheitswesen, die Informations- und Kommunikationstechnik, Logistik, Finanz- und Versicherungswesen und die öffentliche Verwaltung zu den häufigeren Sektoren.

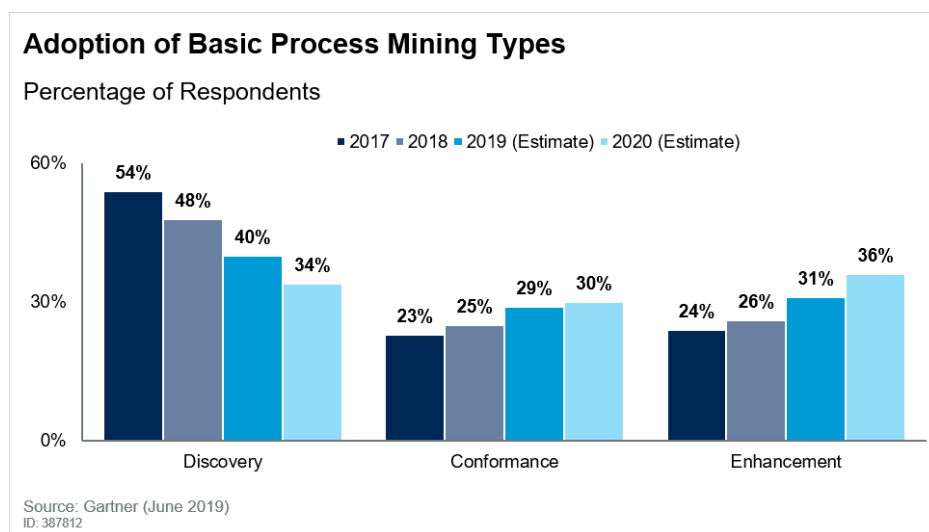


Abbildung 5: Einführung von Process Mining-Typen zwischen 2017 und 2020 [Ker19]

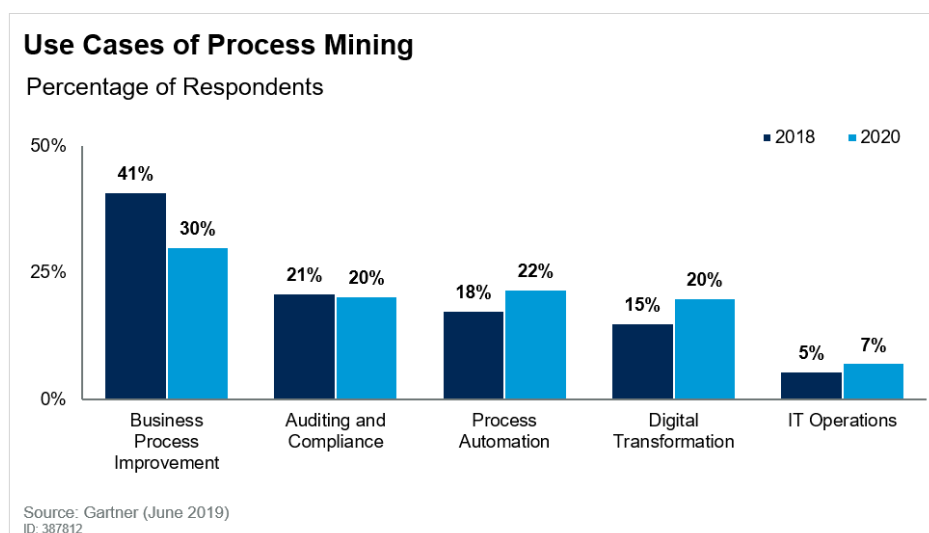


Abbildung 6: Veränderung des Einsatzzweckes von Process Mining zwischen 2018 und 2020 [Ker19]

2.5 Praxisbeispiel aus der Jahresabschlussprüfung

Als wichtiges Einsatzgebiet von Process Mining wird die Anwendung in Jahresabschlussprüfungen gesehen. Bei der Konformitätsprüfung setzen Revisionsabteilungen nicht mehr nur auf traditionelle Verfahren wie Stichprobentests, Walkthrough und Talkthrough, Beobachtungen, Interviews und Inspektionen, sondern auch vermehrt auf die „hypothesenfreie[...] Analyse der digitalen Spuren im Unternehmen“ mithilfe von Process Mining. Folgend werden Revisoren mitunter als Data Analysten ausgebildet, um die Ergebnisse des Process Mining adäquat interpretieren zu können [Len18, S. 202f.]. Die Entdeckung von Prozessen und deren Rekonstruktion zur Erlangung eines Überblicks aller zugrundeliegenden Geschäftsprozesse stellt eine deutliche Verbesserung zu früheren manuellen Prüfungsansätzen dar. Diese Abhängigkeit von manuell erhobenen Informationen wird durch den Einsatz von Process Mining deutlich reduziert, die Transparenz erhöht und die wichtigen Ressourcen Kosten und Zeit für die Prüfungshandlung reduziert. Letztere kann dadurch vermehrt für die detailliertere Analyse von ungewöhnlichen Geschäftsvorgängen verwendet werden. Qualität und Zuverlässigkeit der Prüfungsergebnisse steigen deutlich. Immer komplexer werdende betriebliche Vorgänge und die Vielzahl von Transaktionen mit der Integration in verschiedenste Informationssysteme wären mit herkömmlichen manuellen Prüfmethoden nur unter großem Aufwand bewältigbar. Die Durchführung von Interviews ist sehr zeitintensiv. Darauf folgt eine stichprobenartige Prüfung, was bei einer großen Grundgesamtheit an Daten und der Wichtigkeit der Thematik als unzureichend erachtet werden kann. Zu beachten gilt auch, dass Prüfer bei manueller Vorgehensweise nicht die Existenz von Parallelläufen und Verzweigungen vollumfänglich erfassen können. Diese Überlegungen machen das Process Mining für die Anwendung durch Wirtschaftsprüfer zur Rekonstruktion der Geschäftsvorfälle im Rahmen der Jahresabschlussprüfung sehr attraktiv [Wer11], [Wer12]. Die größeren Wirtschaftsprüfungsgesellschaften haben auch bereits Plattformen aufgebaut, welche Process Mining enthalten, um ihre Auditaktivitäten effizienter zu gestalten [Ker19]. Damit die Transparenz der Prüfungshandlungen sichergestellt wird, dürfen die Ereignisdaten in den Informationssystemen nicht manipulierbar sein [Aal16, S. 302].

2.6 Process Mining-Algorithmen

Algorithmen sind nicht nur entscheidend bei der Datenanalyse, auch beim Process Mining spielen sie eine wichtige Rolle. „Algorithmen sind informatische Werkzeuge, um mathematische Probleme automatisiert zu lösen. Sie berechnen zuverlässig eine Lösung für ein Problem, wenn sie die dafür nötigen Informationen bekommen, den sogenannten „Input“. Das mathematische Problem definiert, welche Eigenschaften der dazugehörige Output, also das Resultat der Berechnung, haben soll“ [Zwe18] und finden beim Process Mining beim Datenimport in ein Process Mining-Tool Anwendung. Umso größer und komplexer die Datenmenge, desto höher wird in der Regel auch die Komplexität der Algorithmen. Dies stellt eine Hauptherausforderung

für Software-Anbieter dar [Aal16, S. 12]. Im Laufe der Zeit wurden viele Process Mining-Algorithmen weiterentwickelt und Neue angefertigt. Zudem repräsentieren sie das am aktivsten bearbeitete Forschungsthema beim Process Mining [Gar19, S. 17f.]. Die Software von ProM bietet eine große Auswahl an, unter anderem heuristische, induktive und genetische Verfahren. Das Ergebnis beim Process Mining hängt entscheidend vom angewendeten Algorithmus ab. Für verschiedene Anforderungen gibt es unterschiedliche Algorithmen. Dabei wird auf das Referenzpapier [Gar19] verwiesen, in dem die Entwicklung der Process Mining-Algorithmen ausführlich thematisiert wird. In [Roz07] wird dies für das Vorhaben einer Prozesserkennung verdeutlicht. Abbildung 7 zeigt, dass das gleiche Ereignisprotokoll für unterschiedliche Algorithmen auch differenzierte Visualisierungsergebnisse liefert.

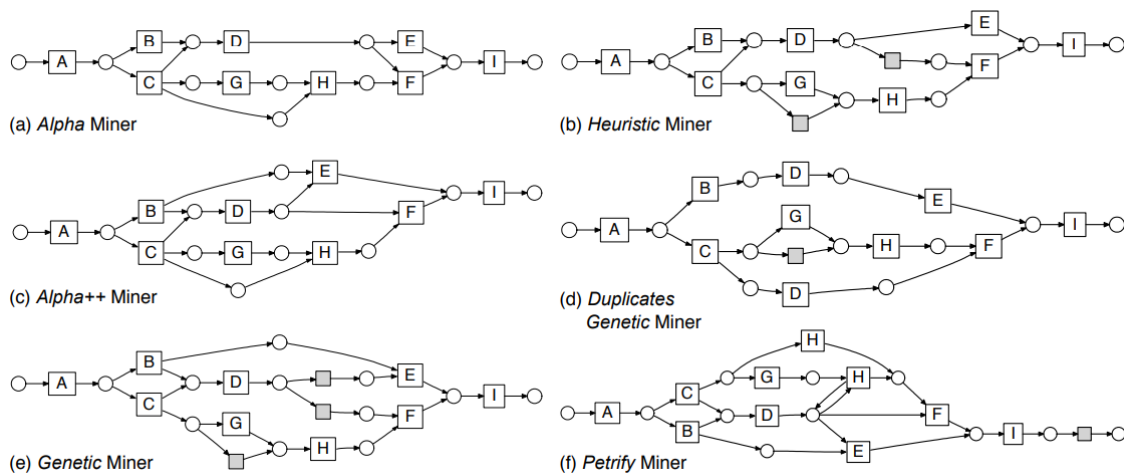


Abbildung 7: Dasselbe Event Log angewendet auf unterschiedliche Algorithmen [Roz07]

Der initiale Algorithmus war der α -Algorithmus, welcher aber für heutige Zwecke ungeeignet ist [Aal16, S. 182]. Im Rahmen dieser Arbeit wird der sehr fortschrittliche Inductive Miner-Algorithmus anhand eines Beispiels erklärt. Der dafür verwendete, exemplarische Beispieldatensatz ist aus praktischen Gründen im Rahmen des Verständnisses bewusst in moderatem Umfang gehalten. Um einen Process Mining-Algorithmus in der Praxis anwenden zu können, muss dieser einigen Qualitätskriterien gerecht werden. Dies soll im nächsten Unterkapitel verdeutlicht werden.

2.6.1 Qualitätskriterien von Process Mining-Algorithmen

Process Mining-Algorithmen kommen beim Import von Event Logs in eine Process Mining-Software zum Einsatz. Um Process Mining zielgerichtet und effizient einsetzen zu können, muss ein Algorithmus einigen Qualitätskriterien gerecht werden. In [Aal16, S. 166f.], [Roz07] und [IEE12] werden Qualitätskriterien für eine Prozesserkennung aus der Kontrollflussperspektive beschrieben, welche sich teilweise konträr zueinander verhalten und deshalb bei der Entwicklung eines Algorithmus eine Abwägung getroffen werden muss (Abbildung 8).

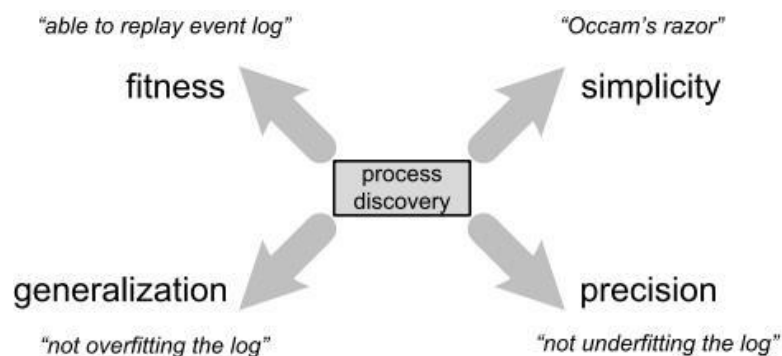


Abbildung 8: Die vier Qualitätskriterien der Prozessentdeckung [Aal16, S. 189]

- **Fitness:** Das Prozessmodell soll ein Verhalten gemäß Event Log darstellen. Die Fitness gibt das Ausmaß des erfassten Verhaltens im Prozessmodell an.
- **Einfachheit:** Das entdeckte Prozessmodell soll sich so einfach wie möglich darstellen und der Theorie von Occam's Razor folgen.
- **Präzision:** Das Prozessmodell soll kein zum Ereignisprotokoll konträres Verhalten zulassen. Alle deutlich abweichenden Vorgänge bleiben unberücksichtigt.
- **Generalisierung:** Das Prozessmodell soll eine verallgemeinerte Darstellung des im Event Log ersichtlichen Verhaltens generieren, welches nicht durch nachfolgende Beobachtungssätze entkräftet wird.

Zur Messung der genannten Kriterien gibt es mehr und weniger praktikable Ansätze. In [Lee16] werden einige nicht praktikable Vorgehen früherer Techniken erläutert sowie neuere und bessere Vorgehen wie die Dekompositionstechnik, die es erlaubt, mit komplexen Ereignisprotokollen und Modellen umgehen zu können.

In diesem Zusammenhang muss auf die Begrifflichkeiten Noise (Rauschen), und Incompleteness (Unvollständigkeit) näher eingegangen werden. Von Noise ist die Rede, wenn das Event Log „seltenes und unregelmäßiges Verhalten, das nicht repräsentativ für das typische Verhal-

ten des Prozesses ist“, enthält [Aal11, S. 147]. Darunter fallen falsch protokollierte Daten, Unzuverlässigkeit des Übertragungskanal oder Aufzeichnungsverzögerungen bei manuellen Tätigkeiten [Bur15, S. 61]. Process Mining-Algorithmen müssen auch mit Unvollständigkeit umgehen können, was bedeutet, dass auch auf nicht Eintretenes nicht die Annahme folgen darf, dass dies gar nicht erst eintreten kann [Aal11, S. 184]. Sobald ein Protokoll eine große Variantenvielfalt an Traces aufweist, ist es unmöglich, alle Verhaltensausrprägungen im Modell zu berücksichtigen [Aal16, S. 38f.]. Dabei muss der Algorithmus einen Ausgleich zwischen „Overfitting“ und „Underfitting“ anstreben. Overfitting (Überdeckung) bedeutet eine nicht genügend generalisierte Prozessdarstellung. Das entdeckte Prozessmodell wird zu spezifisch und präzise dargestellt. Underfitting (Unterdeckung) hingegen bedeutet eine starke Generalisierung der Prozessdarstellung. Das Modell ist zu allgemein und das Verhalten der Ereignisprotokolle wird zu unpräzise dargestellt [Aal16, S. 167]. Zudem ist der Einsatz der bekannten Workflow-Strukturen wie Entscheidungen (OR-/XOR-Splits und -Joins), Sequenzen, Parallelitäten (AND-Splits und -Joins) und Schleifen von den Algorithmen zu unterstützen [IEE12].

Es ist festzuhalten, dass für die verschiedensten Anwendungsfälle von Process Mining eine Vielzahl an Algorithmen entwickelt wurden. Dabei hat sich gezeigt, dass sich mehrere Ansätze für Produktivumgebungen in der Industrie als nicht oder nur sehr schwierig anwendbar erweisen [Wan13]. Dem Inductive Miner-Algorithmus wird nachfolgend ein Kapitel gewidmet, da er zu den fortschrittlicheren und ausgereiften Algorithmen zählt [Bet16].

2.6.2 Inductive Miner-Algorithmus

Der Inductive Miner-Algorithmus ist ein Algorithmus zur Prozesserkennung (Process Discovery), welcher auf den Ansatz von Prozessbäumen zurückzuführen ist. Der vergleichsweise neue Algorithmus stellt eine Verbesserung des α - und Heuristics Miner-Algorithmus dar und nimmt derzeit aufgrund großer Flexibilität, Skalierbarkeit, formaler Sicherheit und Genauigkeit eine führende Rolle zur Prozesserkennung im Process Mining ein. Als großer Vorteil erweist sich auch die Tatsache der einfachen Umwandlung der Ergebnisse in Petri-Netze, Business Process Management Notation (BPMN)-Modellen und Zustandsdiagramme. Der Algorithmus basiert auf einem Divide-and-Conquer-Ansatz, bei dem anstatt eines einzigen Durchlaufs durch ein Ereignisprotokoll das ganze Protokoll in Teilprotokolle rekursiv aufgeteilt und die jeweiligen Beziehungen mit Operatoren versehen werden, bis jedes Teilprotokoll nur noch aus einer einzigen Aktivität besteht. Anfangs wird ausgehend von einem Event Log ein Directly-Follows-Graph erstellt. Der Algorithmus arbeitet mit vier verschiedenen Arten von Schnitten und charakterisierenden Operatoren. Dieser Ansatz verfolgt eine definierte Vorgehensweise [Aal16, S. 200, S. 222ff.]:

1. Wenn es einen nicht-trivialen *exclusive-choice cut* gibt, wird ein maximaler *exclusive-choice cut* angewendet, der das Ereignisprotokoll in kleinere Ereignisprotokolle aufteilt.

Ein *exclusive-choice cut* ist ein Schnitt $(\times, A_1, A_2, \dots, A_n)$, für den gilt:

- $\forall_{i,j \in (1,\dots,n)} \forall_{a \in A_i} \forall_{b \in A_j} i \neq j \Rightarrow a \not\rightarrow_L b.$

2. Wenn es keinen nicht-trivialen *exclusive-choice cut* gibt, aber es sich um einen nicht-trivialen *sequence cut* handelt, dann wird ein maximaler *sequence cut* angewendet, der das Ereignisprotokoll in kleinere Ereignisprotokolle aufteilt.

Ein *sequence cut* ist ein Schnitt $(\rightarrow, A_1, A_2, \dots, A_n)$, für den gilt:

- $\forall_{i,j \in (1,\dots,n)} \forall_{a \in A_i} \forall_{b \in A_j} i < j \Rightarrow a \not\rightarrow_L^+ b \wedge b \not\rightarrow_L^+ a$

3. Wenn es keine nicht-trivialen *exclusive-choice cuts* und *sequence cuts* gibt, aber einen nicht-trivialen *parallel cut*, dann wird ein maximaler *parallel cut* angewendet, der das Ereignisprotokoll in kleinere Ereignisprotokolle aufteilt.

Ein *parallel cut* ist ein Schnitt $(\wedge, A_1, A_2, \dots, A_n)$, für den gilt:

- $\forall_{i \in (1,\dots,n)} A_i \cap A_L^{\text{start}} \neq \emptyset \wedge A_i \cap A_L^{\text{end}} \neq \emptyset$ und
- $\forall_{i,j \in (1,\dots,n)} \forall_{a \in A_i} \forall_{b \in A_j} i \neq j \Rightarrow a \rightarrow_L b.$

4. Wenn es keine nicht-trivialen *exclusive-choice cuts*, *sequence cuts* und *parallel cuts* gibt, aber einen *redo-loop cut*, dann wird ein maximaler *redo-loop cut* angewendet, der das Ereignisprotokoll in kleinere Ereignisprotokolle aufteilt.

Ein *redo-loop cut* ist ein Schnitt $(\cup, A_1, A_2, \dots, A_n)$, für den gilt:

- $n \geq 2,$
- $A_L^{\text{start}} \cup A_L^{\text{end}} \subseteq A_1,$
- $\{a \in A_1 \mid \exists_{i \in (2,\dots,n)} \exists_{b \in A_i} a \rightarrow_L b\} \subseteq A_L^{\text{end}},$
- $\{a \in A_1 \mid \exists_{i \in (2,\dots,n)} \exists_{b \in A_i} b \rightarrow_L a\} \subseteq A_L^{\text{start}},$
- $\forall_{i,j \in (2,\dots,n)} \forall_{a \in A_i} \forall_{b \in A_j} i \neq j \Rightarrow a \not\rightarrow_L b,$
- $\forall_{i \in (2,\dots,n)} \forall_{b \in A_i} \exists_{a \in A_L^{\text{end}}} a \rightarrow_L b \Rightarrow \forall_{a' \in A_L^{\text{end}}} a' \rightarrow_L b$ und
- $\forall_{i \in (2,\dots,n)} \forall_{b \in A_i} \exists_{a \in A_L^{\text{start}}} b \rightarrow_L a \Rightarrow \forall_{a' \in A_L^{\text{start}}} b \rightarrow_L a'.$

Dieser Vorgang der Aufteilung des Event Logs wird solange iterativ wiederholt, bis nur noch Teilprotokolle mit einer einzigen Aktivität vorhanden sind. Durch diese Algorithmik wird ein strukturiertes Prozessmodell erzeugt. Das standardmäßige Vorgehen des Inductive Miner-Algorithmus wird anhand eines Beispiels dargestellt. Das gewählte Beispiel zur Erläuterung wurde dankenswerterweise von Herrn Wil van der Aalst zur Verwendung in dieser Abschlussarbeit zugänglich gemacht [Aal19]. Die graphischen Darstellungen sind gemäß dem Beispiel eigens konzipiert.

Das Ereignisprotokoll zeigt die nachfolgend erfassten Ausprägungen des Prozessablaufs. Dabei steht jeder Buchstabe für eine Aktivität.

a	b	d	e	f
a	c	d	e	f
a	d	b	e	f
a	d	c	e	f
a	b	d	e	g
a	c	d	e	g
a	d	b	e	g
a	d	c	e	g

Die gesamte Spanne der Aktivitäten beträgt $\{a, b, c, d, e, f, g\}$. Jeder Trace besteht jedoch immer aus fünf Aktivitäten. Es folgt eine Splittung mittels *sequence cut* in die Teilmengen $\{a, b, c, d\}$ und $\{e, f, g\}$. Der Algorithmus erkennt die Aktivitäten der ersten Teilmenge immer an den Positionen eins bis drei und die Aktivitäten der zweiten Teilmenge immer an den letzten beiden Positionen.

a	b	d	e	f
a	c	d	e	f
a	d	b	e	f
a	d	c	e	f
a	b	d	e	g
a	c	d	e	g
a	d	b	e	g
a	d	c	e	g

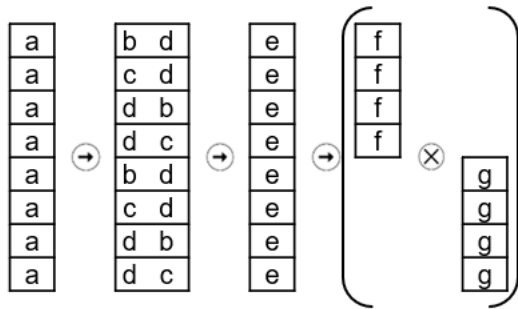
Aktivität a wird beständig an erster Position ausgeführt. Somit folgt eine Teilung in $\{a\}$ und $\{b, c, d\}$ mit einem *sequence cut*.

a	b	d	e	f
a	c	d	e	f
a	d	b	e	f
a	d	c	e	f
a	b	d	e	g
a	c	d	e	g
a	d	b	e	g
a	d	c	e	g

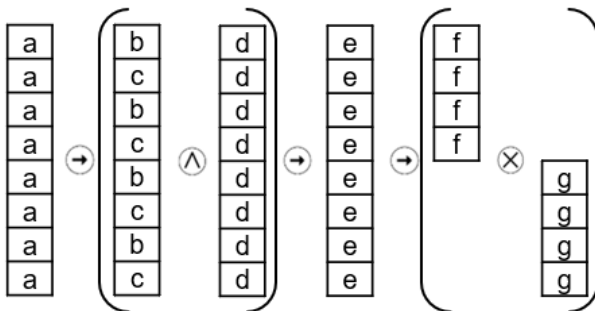
In Teilmenge $\{e, f, g\}$ verhält sich dies deckungsgleich mit Aktivität e, welche immer an vierter Position durchgeführt wird. Es folgt erneut ein *sequence cut*.

a	b	d	e	f
a	c	d	e	f
a	d	b	e	f
a	d	c	e	f
a	b	d	e	g
a	c	d	e	g
a	d	b	e	g
a	d	c	e	g

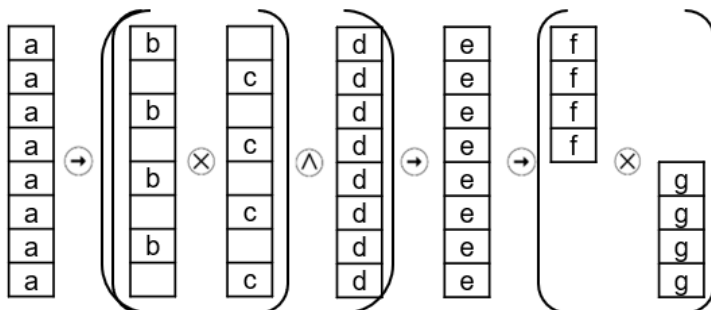
An Position fünf ist ein *exclusive-choice cut* erkennbar, da stets entweder Aktivität f oder Aktivität g ausgeführt wird.



Nun muss nur noch {b, c, d} betrachtet werden. Aktivität d kommt in jedem Trace vor, jedoch variiert deren Position zwischen zwei und drei. Realisiert wird dieser Umstand mit einem *parallel cut* in {b, c} und {d}.



Nun kann in einem letzten Schritt die Teilmenge {b, c} mit einem weiteren *exclusive-choice cut* aufgeteilt werden in {b} und {c}. Dies ist die letzte durchführbare Zerlegung.



Die resultierende Prozessbaumdarstellung ergibt das Bild gemäß Abbildung 9.

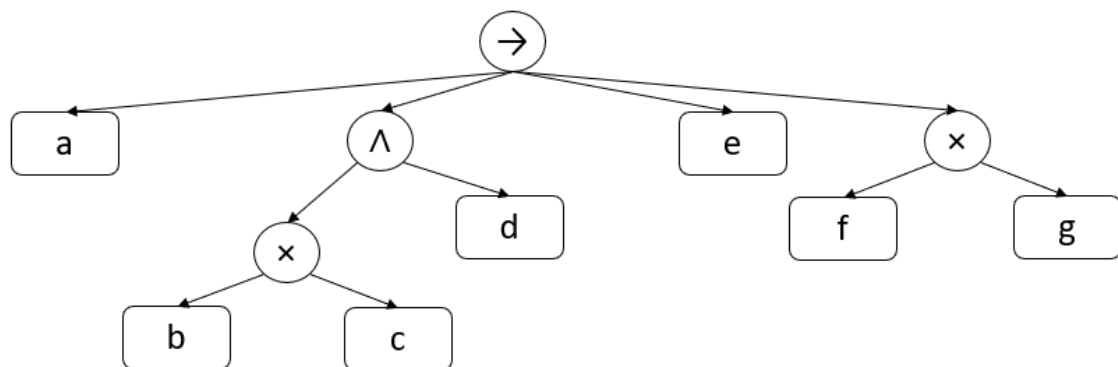


Abbildung 9: Prozessbaumdarstellung gemäß Beispiel (eigene Darstellung gemäß ProM-Plattform)

Wie bereits erwähnt, ist der Inductive Miner-Algorithmus für die Prozesserkennung konzipiert. Prozesslücken, zum Beispiel durch manuelle Tätigkeiten ohne digitale Erfassung, werden bei diesem Algorithmus mit Aktivitätsname tau (T) belegt [Aal16, S. 227ff.]. Würde beispielsweise bei Aktivität e die digitale Erfassung durch den zuständigen Mitarbeiter teilweise vernachlässigt werden, reagiert der Inductive Miner-Algorithmus mit einem *exclusive-choice cut* zwischen Aktivität e und Aktivität tau (T) (Abbildung 10).

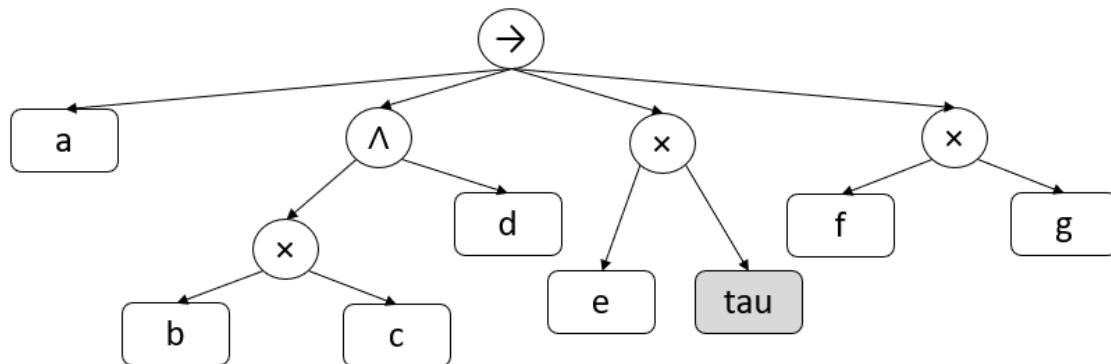


Abbildung 10: Prozessbaumdarstellung mit *exclusive-choice cut* zwischen Aktivität e und tau-Aktivität (eigene Darstellung gemäß ProM-Plattform)

Die Frage nach der Notwendigkeit des Einsatzes von Process Mining und deren Algorithmen ist eine Grundsatzfrage, denn das bekannte Prinzip „Malen nach Zahlen“ würde - umgemünzt auf die Prozessdarstellung - durchaus funktionieren. Dabei wäre die Vorgehensweise folgendermaßen: Alle Aktivitäten und Prozessschritte werden händisch aufgezeichnet. Treten Verbindungen zwischen einzelnen Prozessschritten häufiger auf, so würde die Verbindungslinie dicker werden und der am häufigsten auftretende Prozessdurchlauf könnte auch auf diese Weise entdeckt werden. Diese Vorgehensweise ist jedoch nur für kleine Datenmengen und wenige Events denkbar. Für große und komplexe Prozessdaten wäre die beschriebene Methodik zu zeitaufwändig und unstrukturiert. Ein weiteres triftiges Argument für den Einsatz von Process Mining-Algorithmen ist die Möglichkeit, die konkreten Beziehungen der einzelnen Unternehmensaktivitäten innerhalb eines Prozesses durch allgemein verständliche Operatoren ausfindig zu machen. Auch der Inductive Miner-Algorithmus wird nicht der perfekte und absolute Ansatz sein. Es treten immer neue Forschungsansätze für Process Mining-Algorithmen auf. Allein für den Inductive Miner gibt es bereits viele Abwandlungen, mit denen auch die Problematik mit der Entdeckung von Parallelitäten weitestgehend gelöst werden kann, ohne ein Prozessmodell unzureichend und unpräzise darstellen zu lassen [Aal16, S. 347]. Auch sogenannte Hybrid-Modelle wurden als Plug-ins in die ProM-Plattform eingeführt, die sowohl für reale als auch für synthetische Logs deutliche Verbesserungen in Sachen Präzision aufweisen [Sch18b]. Weitere Process Mining-Algorithmen werden unter anderem in [Aal16] beschrieben und stehen zur kostenfreien Nutzung auf der ProM-Plattform als Plug-ins bereit.

3 Methodisches Vorgehen beim Process Mining

Von den Rohdaten zur Visualisierung des Prozessgraphen sind einige Schritte notwendig. Wenngleich auch relativ schnell ein Prozessgraph durch Laien erzeugt werden kann, ist es dennoch ratsam, zu Beginn des Process Mining im Rahmen eines Projektes voranzuschreiten (Abbildung 11). Nur so wird sichergestellt, dass alle wichtigen Anforderungen erfüllt werden. Der Eingriffszeitpunkt des Endanwenders ist dabei abhängig von den zugrunde liegenden Kenntnissen und wird zu Beginn eines Process Mining-Vorhabens festgelegt. Die anschließenden Abschnitte geben eine methodische, standardmäßige Vorgehensweise zur Durchführung eines Process Mining-Projektes im L*-Lebenszyklusmodell wieder [Eck15], [IEE12].

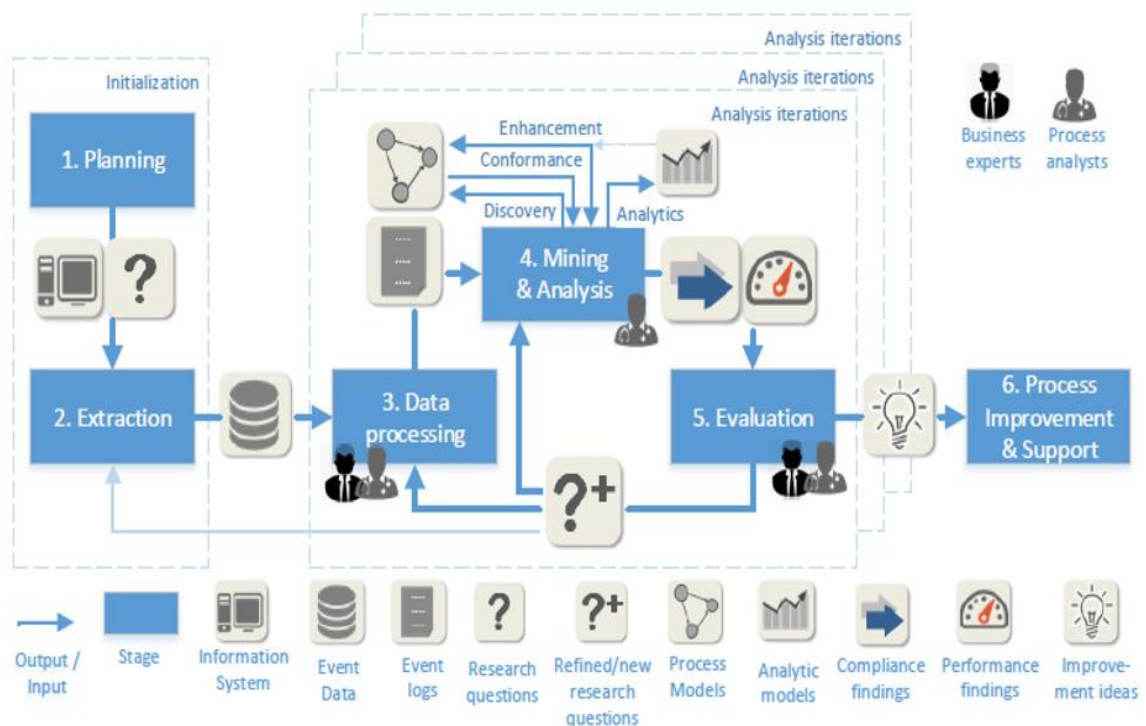


Abbildung 11: Methodisches Vorgehen beim Process Mining [Eck15]

3.1 Planung und Modellierung

In der Planungsphase wird zunächst der zu analysierende Geschäftsprozess ermittelt und eine konkrete Forschungsfrage zum ausgewählten Geschäftsprozess definiert, welche sich auf die Faktoren Kosten, Qualität, Ressourcen und Zeit stützen kann. Von Bedeutung ist auch die wohlüberlegte Zusammenstellung des Projektteams. Geschäftsprozessverantwortliche und -experten sollen mit Prozessanalysten in enger Zusammenarbeit stehen [Nom18]. Die Planung ist wichtig in Bezug auf die korrekte Darstellung der Unternehmensprozesse. Dazu bedarf es einer Modellierung, welche im Zusammenhang mit Process Mining einer vorherigen Festlegung aller sich auf die konkrete Fragestellung beziehenden Daten entspricht. Die Process Mining-Ergebnisse stehen in direkter Abhängigkeit zur Modellierung. Die Modellierung ist der

Schritt beim Process Mining, der die größte Erfahrung und Kenntnis von Strukturen und Datentabellen des Quellsystems voraussetzt. Es sollten stets nur die Daten berücksichtigt werden und in die Analyse einfließen, die gemäß Zielstellung von Bedeutung sind. Es findet folglich eine Abgrenzung statt. Andererseits kann die Software gegebenenfalls keine performante Datenverarbeitung durchführen oder es entstehen zu komplexe Prozessgebilde, die nicht oder nur sehr schwer zu interpretieren sind. Somit ist die Modellierung ein entscheidender Schritt zur Darstellung der real in einer Organisation ablaufenden Prozesse. Allein dieser Schritt setzt schon fundiertes Wissen in Bezug auf das Quellsystem voraus. Diesem Schritt folgt die Datenextraktion aus dem Quellsystem.

3.2 Datenextraktion aus Quellsystem

Prozessbezogene Daten weisen verschiedene Quellsysteme auf. Diese Systeme unterliegen oftmals einer vordefinierten Darstellung der beinhalteten Daten. Aus diesem Grund ist der Einsatz eines adäquaten Extraktionstools von großer Bedeutung bei der Anwendung von Process Mining. Viele Softwarehersteller bieten Tools mit eigens entwickelten Extraktionsskripten für diverse Quellsysteme an, welche für das Process Mining großen Nutzen bringen. Die Ereignisdaten werden dadurch aus den unterschiedlichen Systemen zu einer einzigen Sammlung zusammengeführt. Process Mining muss sich mit Problemen der Datenqualität auseinandersetzen [Aal16, S. 125]. Bei der Extraktion und Zusammenführung von Daten spielen sowohl Syntax als auch Semantik eine wichtige Rolle. Als Syntax werden „formale Regeln über die zulässigen Sprachelemente einer Programmiersprache und über zulässige Möglichkeiten ihrer Verwendung in einem Programm“ [Sie18] bezeichnet. Die Semantik liefert „Aussagen über die Bedeutung der Sprachelemente und der zulässigen Kombinationen von Elementen“ [Sie18]. Insbesondere sollte auch der Zeitraum der Extraktionsdaten sorgfältig gewählt werden. Process Mining ist generell dann durchführbar, sobald eine Datenextraktion und eine Event Log-Generierung möglich sind. Letzteres wird im nun folgenden Schritt erläutert.

3.3 Datenverarbeitung und Event Log-Generierung

Die Datenverarbeitung und Generierung von sogenannten Event Logs stellen die Basis für einen Import von Daten in ein Process Mining-Tool dar. Ziel ist es dabei, für das definierte Process Mining-Vorhaben alle von den verschiedenen Quellsystemen extrahierten Daten in eine einheitliche Form zu bringen und Cases sowie Eventklassen zu definieren. Dies ist essentiell, damit die Process Mining-Software die Spalten direkt den Attributen zuordnen kann. In der Datenverarbeitung können Ereignisse aggregiert werden und Event Logs mit zusätzlichen für die Analyse relevanten Attributen angereichert werden. Verschiedene Filter-Ansätze werden häufig angewandt, um die Komplexität zu reduzieren. In der Praxis hat es sich bewährt, die Belegnummer und die Positionsnummer zur Case ID zu kombinieren. Dies stellt im Ge-

gensatz zu künstlich generierten und nicht sprechenden Nummern eine verbesserte Nachvollziehbarkeit sicher. Jedes Ereignis sollte einen eigenen Datensatz im Ereignisprotokoll darstellen. Somit wird pro Position ein eigenständiger Prozessdurchlauf unter einer eindeutigen Case ID durchgeführt. Event Logs werden üblicherweise als XES-, MXML-, Excel- oder CSV-Dateien zum Import bereitgestellt. Aber auch der direkte Bezug aus Datenbanken oder Anwendungen ist möglich [Aal16, S. 330]. In Abbildung 12 sind die Möglichkeiten in den Phasen Datenverarbeitung und Event Log-Generierung sowie Mining und Analyse visuell dargestellt.

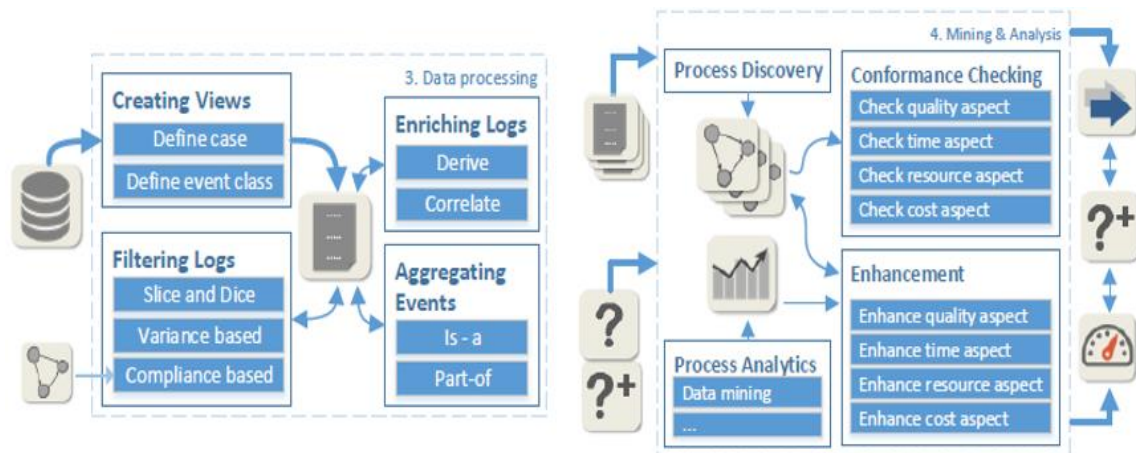


Abbildung 12: Phasen der Datenverarbeitung, Mining und Analyse [Eck15]

3.4 Mining und Analyse

Sobald die Event Log-Generierung erfolgreich abgeschlossen ist, kann dieses in ein Process Mining-Tool importiert werden, welche die Ereignisprotokolle unter Berücksichtigung der Mindestanforderungen (Case ID, Aktivität, Zeitstempel) plus zusätzlich definierter Attribute in Form einer Prozessvisualisierung darstellt. Beim Import der Protokolle überführt das Tool diese durch Algorithmen und festzulegende Schwellenwerte (Thresholds) in einen Graph. Dies ist die Mining-Phase, in der gemäß Forschungsfrage die Prozessentdeckung, die Konformitätsprüfung oder die Erweiterung von Prozessmodellen als Output generiert wird. Häufig kommen weitere Analysetechniken wie etwa aus dem Data Mining-Bereich und Visualisierungen wie Dashboards hinzu. Kommerzielle Process Mining-Tools beinhalten bereits vorher definierte Darstellungsformen, wie beispielsweise Dashboards mit Key Performance Indikatoren (KPIs) für Prozesse, Automatisierungsraten, Laufzeiten sowie standardisierte Berichte. Im Analyseschritt können nun einzelne Prozessdurchläufe und vorliegende Mengengerüste analysiert werden. Eine vollumfängliche Analyse der eigenen Prozesse wird derzeit mit transaktionsbasierter Datenanalyse sichergestellt. Dieser Punkt wird in 4.5 bei den Herausforderungen von Process Mining vorgestellt. In Abschnitt 6 werden einige Anbieter von Process Mining-Tools vorgestellt, welche sich auch in weitergehenden Analysemöglichkeiten unterscheiden.

3.5 Evaluierung

Die Bewertung von Ergebnissen eines Process Mining-Vorhabens wird in der Regel von Angestellten durchgeführt, welche mit der Thematik Geschäftsprozessmanagement betraut sind. Nur so können adäquate Schlussfolgerungen aus den Ergebnissen gezogen werden. Dabei ist es wichtig, eine oder mehrere Ursachen für häufig auftretende Abnormitäten ausfindig zu machen und bereits Forschungsfragen für weitere Iterationen des kontinuierlichen Process Mining-Prozesses zu erkennen und anzustoßen.

3.6 Umsetzung von Prozessverbesserungen

Die bei der Analyse und Bewertung gewonnenen Erkenntnisse sollen in einem abschließenden Schritt zur konkreten Umsetzung von Prozessverbesserungen führen. Dessen Umsetzung wird meist im jeweiligen Fachgebiet im Rahmen eines eigenständigen Projektes realisiert. Process Mining sollte einen kontinuierlichen Prozess darstellen, bei dem Geschäftsprozessverantwortliche und Prozessanalysten eng miteinander interagieren. Eine proaktive Nutzung und Entwicklung der toolbasierten Methode öffnet die Tür zur operativen Unterstützung.

4 Herausforderungen bei der Anwendung von Process Mining

Dieses Kapitel gibt einen Überblick darüber, welche Herausforderungen die Anwendung von Process Mining mit sich bringen kann. Herausforderungen ergeben sich auf unterschiedlichen Ebenen und aus verschiedenen Sichten. Zum einen sehen sich Anbieter und Entwickler von Process Mining-Software bezüglich der eingesetzten Techniken einigen speziellen Anforderungen gegenübergestellt. Zum anderen hat der Endanwender Herausforderungen in Bezug auf die Definition bestimmter Parameter und die Interpretation der Ergebnisse zu bewältigen. Zu erwähnen ist dabei, dass die Ausprägung der Herausforderungen oftmals vom eingesetzten Tool abhängt und die nun folgenden Punkte als nicht abschließend anzusehen sind.

4.1 Schaffung der Voraussetzungen für Process Mining

Process Mining basiert auf in Informationssystemen gespeicherten Ereignisdaten. Organisationen haben häufig für verschiedene Anforderungen mehrere Informationssysteme im gleichzeitigen Einsatz und die einzelnen Prozessausführungsschritte erzeugen darin temporäre und permanente Spuren [Beh16, S. 128]. Diese Tatsache macht das Vorhaben einer Gestaltung größtmöglicher Transparenz der eigenen Prozesse zu einem schwierigen Unterfangen. Oftmals stehen Organisationen vor der Herausforderung, die Vielzahl an verstreuten Daten zusammenzuführen und valide Handlungsempfehlungen daraus zu generieren. Unterschiede zwischen Unternehmen ergeben sich beim Stellenwert der eigenen Daten in Informationssystemen. In [Bur15, S. 59ff.] wird das Konzept des Prozessbewusstseins wie folgt charakterisiert:

- Prozessbewusstsein des Unternehmens: Geschäftsprozessmanagement wird täglich praktiziert
- Prozessbewusstsein der im Unternehmen eingesetzten Informationssysteme: Sind die vorhandenen Informationssysteme für den prozessualen Einsatz geeignet?

Während sich die Prozesskultur in mittleren und großen Unternehmen etabliert hat und Informationssysteme die Prozessausführungen gänzlich unterstützen, sind in kleineren Unternehmen verwaltende Systeme für die Abwicklung aller Aktivitäten nicht durchgängig im Einsatz. Nicht in Systemen erfasste Arbeitsschritte werden auch als Schattenaktivitäten bezeichnet. Dokumentiert werden diese ganz klassisch auf Papier. Dies hat wirtschaftliche Gründe, da Informationssysteme sowohl in der Implementierung als auch in der Erfassung der Daten teuer sind und Process Mining somit für kleine Betriebe häufig keine Rolle spielt. Die durchgängige digitale Erfassung von Arbeitsschritten ist dann schlichtweg nicht rentabel.

Ereignisdaten aus Informationssystemen liegen in unterschiedlichen Ausprägungen vor. Bei deren Aufbereitung können diese zu überwältigend großen Ereignisprotokollen mit sehr vielen

Attributen führen. Anderen hingegen fehlen sogar die Mindestanforderungen. Protokolle, welche die Mindestanforderungen nicht abbilden, sind nicht für Process Mining anwendbar [Bur15, S. 4 und S. 59f.]. Um diesen Umstand obsolet zu machen, muss organisationsweit ein Verständnis für die Wichtigkeit des richtigen und verantwortungsvollen Umgangs mit Daten und deren Erfassung erreicht werden. Jedoch liegt der Großteil an Daten immer noch in unstrukturierter Form vor, sodass eine der größten Herausforderungen für Unternehmen zukünftig darin bestehen wird, die in ihren Informationssystemen massenhaft gespeicherten Daten mit der dynamischen Prozesswelt zu vereinen und dadurch wertvolle Informationen und Wissen zu gewinnen [Aal16, S. 4]. Die Qualität des Process Mining ist von der zuverlässigen Eingabe der Protokolldaten abhängig [IEE12]. Es sollte eine Kultur vermittelt werden, nach der Ereignisdaten als kostbare Ressource anzusehen sind. Qualitativ hochwertige Daten weisen eine präzise Semantik auf, sind zuverlässig richtig und vollständig. Zudem berücksichtigen sie Sicherheitsaspekte bei der Aufzeichnung. Im Process Mining Manifest [IEE12] werden Daten in fünf Reifegrade eingeteilt. Je höher der Reifegrad, desto zielgerichteter können Ereignisdaten für das Process Mining eingesetzt werden. Zur Verbesserung der Datenqualität in Unternehmen wären Kontrollfelder denkbar, welche alle notwendigen Informationen aufzeigen und vom beteiligten Mitarbeiter nochmals bestätigt werden müssen. Dies ist ein möglicher Ansatz zur Reduzierung fehlerhafter Daten. Bei unvollständigen Daten kann es passieren, dass Ereignisse nicht ausdrücklich auf Prozessinstanzen verweisen. Die sogenannte Ereigniskorrelation, durch die solche Ereignisse zu Prozessinstanzen zuordnenbar werden, wird in [Bur15, S. 71ff.] anhand eines Beispiels zu einem Dokumentenmanagementsystem aufgezeigt.

4.2 Zielfestlegung

Das methodische Vorgehen beim Process Mining wurde in Abschnitt 3 vorgestellt und begleitet das Vorhaben in strukturierter Weise. Dieser Projektgedanke ist für die Einführung einer für Process Mining relevanten Thematik empfehlenswert, um etwaigen Fallstricken zu entgehen. Langfristig sollte Process Mining als ein kontinuierlicher Prozess verstanden werden, bei dem Fragestellungen in gewissen zeitlichen Abständen wiederholt analysiert werden. Zu unterscheiden sind hierbei die datengetriebene, die fragengesteuerte oder die zielgetriebene Fragestellung [Aal16, S. 394f.].

Während der Planungsphase eines Process Mining-Vorhabens sollte das Ziel präzise definiert werden und ein gemeinsamer Konsens vorliegen. Dies umfasst die Entscheidung darüber, für welche betrieblichen (End-to-End)-Prozesse (zum Beispiel Beschaffung oder Produktion) eine genauere Betrachtung durchgeführt werden soll. Dabei sollte der Nutzen der Investition - der Return on Investment (ROI) - die veranschlagten Kosten (unter anderem Implementierungs-, Lizenzierungs- und Personalkosten) mindestens decken. Es ist eine Einschätzung des ROI

erforderlich, was eine nicht-triviale Angelegenheit ist. In der Prozessentdeckung und -erweiterung spielen sicherlich Einsparungen durch Effizienzverbesserungen eine große Rolle. Folgender Gedankengang scheint für den Produktionsbereich möglich: Um wieviel kann der Produktionsoutput gesteigert werden, wenn durch schlankere Prozesse die Durchlaufzeiten erheblich reduziert werden können?

Eindeutige Zielfestlegungen sollten immer mit Effizienz- und Qualitätssteigerungen oder einer Risikominimierung für die Organisation einhergehen. Hohe organisationale Qualität generiert auch positiven Einfluss auf die Zufriedenheit aller Mitarbeiter im Wissen, dass Prozesse ordnungsgemäß und durchdacht ablaufen. Engpässe in jeglicher Form sind nicht nur ineffizient. Sie werfen auch ein schlechtes Licht auf ein Unternehmen, sobald Kunden direkt davon betroffen sind.

4.3 Problematiken der einzelnen Prozessschritte

Die einzelnen Schritte eines Process Mining-Vorhabens wurden in Kapitel 3 aufgeführt. In jeder Phase treten gewisse Herausforderungen auf. Die Kapitel 4.3.1 bis 4.3.6 geben darüber Aufschluss.

4.3.1 Planung und Modellierung

Die für einen Geschäftsprozess relevanten Daten aus einer Vielzahl von heterogenen Quellsystemen ausfindig zu machen, setzt zum einen großes Prozesswissen voraus und zum anderen technische Domainkenntnisse in den beteiligten Systemen [Eck15], [IEE12]. Als Beispiel dazu soll ein Purchase-to-Pay-Prozess (P2P-Prozess) visualisiert werden. Die dazugehörigen Daten entstammen einem SAP-System als Quelle. Im Rahmen der Planung und Modellierung wird nun ermittelt, in welchen Daten- und Änderungstabellen sich die für einen P2P-Prozess relevanten Daten befinden. Da eine komplette SAP-Implementierung viele tausende Datentabellen umfasst, ist dieser Schritt alles andere als einfach [Aal16, S. 126f.]. Tabelle 1 für das SAP-Modul Materialmanagement (SAP-MM) und Tabelle 2 für das SAP-Modul Finanzen (SAP-FI) bilden die Reihenfolge eines kompletten P2P-Prozesses ab. Im Anhang in Abbildung 22 sind die Tabellen mit Beziehungen zueinander graphisch dargestellt.

Tabelle 1: Vereinfachte Darstellung der relevanten Tabellen aus SAP-MM für P2P-Prozess [dab:GmbH]

Prozessschritt	Tabelle	Beschreibung	Feldname	Beschreibung
Bestellanforderung	EBAN	Bestellanforderung	BANFN BNFPO	Bestellanforderungsnummer Bestellanforderungspositionsnummer
Bestellung	EKKO	Einkaufsbelegkopf	EBELN	Einkaufsbelegnummer
	EKPO	Einkaufsbelegposition	EBELN	Einkaufsbelegnummer

			EBELP	Einkaufsbelegpositionsnummer
Wareneingang	MKPF	Belegkopf Materialbeleg	MBLNR MJAHR	Materialbelegnummer Materialbelegjahr
	MSEG	Belegsegment Material	MBLNR MJAHR ZEILE	Materialbelegnummer Materialbelegjahr Materialbelegposition
Rechnungseingang	RBKP	Belegkopf Eingangsrechnung	BELNR GJAHR	Belegnummer der Rechnung Geschäftsjahr
	RSEG	Belegposition Eingangsrechnung	BELNR GJAHR BUZEI	Belegnummer der Rechnung Geschäftsjahr Belegposition im Rechnungsbeleg

Tabelle 2: Vereinfachte Darstellung der relevanten Tabellen aus SAP-FI für P2P-Prozess [dab:GmbH]

Prozessschritt	Tabelle	Beschreibung	Feldname	Beschreibung
Verbindlichkeit	BKPF	Belegkopf für Buchhaltung	BURKS BELNR GJAHR	Buchungskreis Belegnummer Geschäftsjahr
	BSIK	Buchhaltung: Sekundärindex für Kreditoren	BURKS BELNR GJAHR BUZEI	Buchungskreis Belegnummer Geschäftsjahr Buchungszeile
Zahlung	REGUH	Regulierungsdaten aus Zahlprogramm	LAUFD LAUFI ZBUKR VBLNR	Laufdatum Identifikation Zahlender Buchungskreis Zahlbelegnummer
	REGUP	Bearbeitete Positionen aus Zahlprogramm	LAUFD LAUFI ZBUKR VBLNR BUZEI	Laufdatum Identifikation Zahlender Buchungskreis Zahlbelegnummer Position
Ausgleichsbuchung	BKPF	Belegkopf für Buchhaltung	BURKS BELNR GJAHR	Buchungskreis Belegnummer Geschäftsjahr
	BSAK	Buchhaltung: Sekundärindex für Kreditoren (ausgegl. Posten)	BURKS BELNR GJAHR BUZEI	Buchungskreis Belegnummer Geschäftsjahr Buchungszeile

4.3.2 Datenextraktion aus Quellsystem

Beim Extrahieren der Daten aus dem Quellsystem ist der Einsatz eines Extraktionstools zielführend, welches zuverlässig und vollständig große Datenmengen extrahieren kann. Dabei sind Extraktionstools nicht per se für Process Mining-Vorhaben konzipiert, bieten sich aber natürlich dafür an. In der Regel haben Process Mining-Anbieter solche Datenextraktionstools nicht selbst im Produktportfolio. Es wird oft auf Partnerunternehmen verwiesen oder, falls möglich, eine manuelle Extraktion aus dem Quellsystem vorgeschlagen. In SAP kann über die Transaktionen SE16 oder SE16N die manuelle Datenextraktion realisiert werden. Dafür sind einige Schritte händisch zu tätigen. Hersteller von Datenextraktionstools punkten dahingegen zumeist mit Support und großer Bedienerfreundlichkeit der Software, um unproblematisch an die Daten zu gelangen und diese automatisch in verschiedene Datenanalyseformate umzuwandeln. Selbst größte Datenmengen, bei denen ein SAP-System an seine Grenzen stößt, können zuverlässig verarbeitet werden. SAP selbst hat überdies Schnittstellen für externe Datenextraktionstools im eigenen ERP-System eingerichtet.

4.3.3 Datenverarbeitung und Event Log-Generierung

Das Zusammenführen der Extraktionsdaten zu einem einheitlichen Event Log ist die Basis für einen reibungslosen Import in die Software. Bei der Datenverarbeitung ist auf eine einheitliche Darstellung zu achten, damit der Algorithmus zwischen den Attributen unterscheiden kann. Nicht einheitliche Schlüsselfelder der Quellen für ein und denselben Sachverhalt machen die korrekte Zusammenführung schwierig [IEE12]. Die Case ID sollte einfach nachzuvollziehen sein. Dies kann durch eine Verkettung von Beleg- (beispielsweise Bestellbelegnummer) und Positionsnummer realisiert werden. In der Praxis hat sich dies bewährt, da Cases ohnehin meistens auf Positionsebene betrachtet werden.

4.3.4 Mining und Analyse

Die Herausforderungen des Mining sind in erster Linie mit dem Algorithmus verbunden und werden in Kapitel 4.4 ausführlich behandelt. Der Punkt der Ergebnisanalyse betrifft Anbieter und Endanwender zugleich. Um die Verbreitung des Tools weiter positiv voranzutreiben, sollten die Anbieter auf eine möglichst intuitive Benutzeroberfläche achten. Ideal sind sich nahezu selbsterklärende Darstellungen und Bearbeitungsmöglichkeiten. Dabei sollte der Endanwender nicht zu intensiv mit Parametern arbeiten müssen. Für den Analysezweck gestaltete Techniken sollten durch Erkennung bestimmter Muster in den Daten automatisch sinnvolle Parameter definieren. Klar strukturierte, einfach zu bedienende und nicht zu überladene Oberflächen laden zur täglichen Arbeit mit einer Process Mining-Software ein. Kommerzielle Anbieter sind schon sehr weit bezüglich einer ansprechenden Gestaltung ihrer Tools. Dashboards mit KPIs sind bereits weit verbreitet. Letztendlich sollten möglichst viele Erkenntnisse bereits durch das Tool und im Tool selbst gewonnen werden können.

4.3.5 Evaluierung

Die gewonnenen Erkenntnisse aus der Ergebnisanalyse müssen einer Bewertung unterzogen werden. Der Prozessschritt der Evaluierung bietet vielfältige Fallstücke und bedarf umfangreicher Kenntnisse über das Geschäftsprozessmanagement und das Process Mining-Tool. Werden falsche Rückschlüsse gezogen oder der Hebel beim falschen Bearbeitungsschritt angesetzt, so kann dies zu teuren Verzögerungen des Verbesserungsvorhabens führen. Als Anregung soll im Rahmen dieser Arbeit die Bewertung und der Umgang mit Schleifen im Prozessgraph dienen. Der Analyst wird Fälle vorfinden, bei denen eine Schleife auf ein korrektes Verhalten aber auch auf einen Schwachpunkt im Prozess hinweisen kann. Schleifen sind sich selbst wiederholende Aktivitäten. Da ein Case gewöhnlicherweise pro Position definiert wird, kann bei Aktivität „Wareneingang buchen“ die Situation von vereinbarten Teillieferungen entstehen. Beispielsweise werden 150 Einheiten eines Artikels auf drei Teillieferungen zu je 50 Einheiten geliefert. Die Systembuchung des Wareneingangs wird zu den drei unterschiedlichen Anlieferungszeitpunkten vorgenommen. Die Rechnungsstellung erfolgt als Gesamtrechnung nach vollständiger Lieferung. Dieser beispielhafte Fall würde im Prozessgraphen, gemäß der Annahme von häufigeren Vorfällen dieser Art, zu prozessvaliden Schleifen bei der Aktivität „Wareneingang buchen“ führen. Dem entgegengesetzt eine Situation, wonach bei der Aktivität „Bestellposition hinzufügen“ unberechtigterweise bei einer noch offenen Bestellung weitere Positionen erfasst werden. Ein mögliches Beispiel lässt sich bei der Umgehung von Genehmigungsgrenzen bei Bestellungen konstruieren, bei der anstatt einer genehmigungspflichtigen Bestellung einige Bestellpositionen auf „alte“, noch nicht vollständig fakturierte Bestellungen beim selben Lieferanten gebucht werden. Diese beiden konträren Absichten zur gleichen Thematik der Schleifen verdeutlichen die Schwierigkeit bei der Analyse und Bewertung von Process Mining-Graphen [Beh16, S. 128].

4.3.6 Umsetzung von Prozessverbesserungen

Dieser Prozessschritt sollte direkt nach erfolgter Evaluierung umgesetzt werden, um schnell von den gewonnenen Erkenntnissen profitieren zu können. Beispielsweise wird von einem Unternehmen als Ziel eines konkreten Process Mining-Vorhabens vorgegeben, die Durchlaufzeiten der zehn häufigsten Varianten der Produktionsprozesse um jeweils 5 Prozent zu verkürzen. Dabei werden mithilfe der toolbasierten Methode auffällig lange Liegezeiten zwischen den Prozessschritten ausfindig gemacht. Bei der Analyse und Bewertung werden Maßnahmen zur Verbesserung geprüft und festgelegt. Letztendlich wird eine konkrete Prozessverbesserung angestoßen, um den im Produktionsprozess erkannten Flaschenhals durch die Anschaffung einer zweiten Maschine den betreffenden Arbeitsschritt auf zwei Maschinen parallel ausführen zu lassen.

Eingefahrene Dinge sind erfahrungsgemäß schwer änderbar. Dies ist bei Prozessen oft nicht anders. Es fehlt die Akzeptanz für Veränderungen. „Das haben wir schon immer so gemacht!“, hört man sehr häufig in diesem Zusammenhang. Dabei ist die Gesamtunternehmenssicht immer wichtiger als die Abteilungssicht, wenn es um Prozessverbesserungen geht. So kann es vorkommen, dass Prozessverbesserungen für die betroffene Abteilung gar nicht als solche wahrgenommen werden, sich aber aus Unternehmenssicht deutliche Vorteile ergeben. Generell ist bei der Umsetzung von Prozessverbesserungen eine transparente Kommunikation von großer Wichtigkeit.

4.4 Ausgereiftheit der Algorithmen

Das Prozessbild, welches als Startpunkt für weitergehende Analysen im Rahmen eines Process Mining-Vorhabens dient, wird aus einem vordefinierten Algorithmus erzeugt, welcher beim Import der Ereignisprotokolle Anwendung findet. Wichtig ist hierbei, dass die Event Logs fehlerfrei verarbeitet werden können, um ein aussagekräftiges Gesamtprozessbild zu erhalten. Dabei muss der Algorithmus einige, teils konträre Anforderungen erfüllen. Algorithmen müssen so konzipiert sein, dass sie die Ausprägungen von Noise und Unvollständigkeit in angemessenem Rahmen beachten und Prozesslücken beziehungsweise Schattenaktivitäten berücksichtigen. Zudem bedarf es einer Festlegung, ob Schwellenwerte vorgegeben oder diese benutzerdefiniert zu setzen sind.

Different Informationssysteme bedingen zwangsläufig auch ungleiche Granularitäten. Zu beobachten ist dies oftmals beim Zeitstempel. Grobe Datumsinformationen geben nur das Datum und keinerlei Zeitangaben an, während feine Datumsinformationen Einträge im Millisekundenbereich liefern können [IEE12]. Process Mining-Techniken sollten für weiterführende Analysen in der Lage sein, viele unterschiedliche Attribute in Protokollen verarbeiten zu können. So liefert die Angabe von Start- und Endzeitstempel weitaus bessere Analysemöglichkeiten als nur die Angabe einer der beiden Zeitstempel.

Während eine zunehmende Entwicklung weiterer Algorithmen festzustellen ist, fehlt weiterhin die Einigkeit zu deren Qualitätsbewertung. Die diversen Ausgestaltungen der Algorithmen lassen einen Vergleich über die Qualität kaum zu. Benchmarks sollten hierbei Abhilfe schaffen. Jedoch wird auch erkannt, dass Standards bei der Prozessmodellierung schlicht schwieriger umzusetzen sind als bei Tabellen. Bekannte und etablierte Qualitätskriterien wie Einfachheit, Fitness, Generalisierung und Präzision sollten erst den Anfang weiterer, zu entwickelnder Kriterien darstellen [IEE12]. Die Annahme, die Technik müsse stets eine Balance zwischen den vier Qualitätskriterien finden, stimmt dahingegen nur zum Teil, denn je nach Anwendungszweck muss ein Algorithmus beispielsweise unterschiedlich mit Noise umgehen. Wir erinnern uns: Noise bedeutet abnormes Verhalten. Während bei der Prozessentdeckung zur Modellie-

rung eines neuen Prozessmodells keinerlei Noise erwünscht ist und diese im Modell nicht berücksichtigt werden sollten, wird bei der Konformitätsprüfung der Fokus vor allem auch auf diese ungewöhnlichen Vorgänge, auch Ausreißer genannt, gelegt. Mining-Algorithmen für die Konformitätsprüfung müssen abnormes Verhalten berücksichtigen, solche für die Prozessentdeckung nicht. In Abschnitt 2.6.1 wurden die Kriterien vorgestellt, welche die Process Mining-Algorithmen in der Prozessentdeckung zu beachten haben. Zur Abwägung der Kriterien wurden Parameter (Thresholds) eingeführt, deren Anwendung jedoch weitere Herausforderungen mit sich bringt: Die Handhabung dieser Parameter. Die diversen Algorithmen benötigen oftmals Schwellenwert-Angaben, vor allem die Ausgereifteren, wie etwa der Heuristics Miner-Algorithmus [Bur15, S. 98ff.]. Schwellenwerte steuern den Umgang des Algorithmus mit Noise, Unvollständigkeit, Abhängigkeiten, Schleifen et cetera. Umso leistungsfähiger und universell einsetzbarer die Algorithmen sind, desto mehr Parameter müssen gesetzt werden. Bei deren Konfiguration haben typische Benutzer oder Laien häufig Schwierigkeiten, diese richtig zu setzen [Bur15, S. 62]. In Abbildung 13 ist dies beispielhaft für die ProM-Plattform bei Verwendung des Heuristics Miner dargestellt. Bei Anwendung dieser Schwellenwerte wird der Algorithmus nur Abhängigkeitswerte ab 90 im Modell berücksichtigen. Schleifen der Länge eins und zwei werden ebenfalls nur dann in das Modell übernommen, wenn der mit entsprechender Formel ermittelte Wert über 90 liegt [Bur15, S. 99]. In [Bur15, S. 101ff.] wird anhand des Heuristics Miner ++ ein fortschrittlicher Ansatz beschrieben, bei dem der Parameterraum aufgrund der Protokollspuren strukturiert und eine lokale Suche in diesem Raum durchgeführt wird, um die Komplexität einer Erforschung aller möglichen Prozessmodelle meistern zu können.

Abbildung 13: Einzustellende Parameter zum Heuristics Miner in ProM

Abhilfe bei der Bewältigung dieser Herausforderung sollte dabei der Austausch mit Daten- und Prozessanalysten schaffen, die oftmals bereits umfangreiche Erfahrung im Umgang mit Process Mining-Tools haben und Auswirkungen bestimmter Parametereinstellungen problemlos nachvollziehen können. Denkbar ist auch eine benutzergeführte Vorgehensweise für die Konfiguration der Parameter, wie in [Bur15, S. 113ff.] beschrieben. In [Lee16] werden Ansätze vorgeschlagen, mit denen Skalierbarkeit und starke Qualitätsgarantien für sehr große Datensätze in Einklang gebracht werden können. Diese Kombination wird von bestehenden Prozesserkennungsalgorithmen nicht zufriedenstellend gewährleistet. Der Inductive Miner-Algorithmus bietet hohe Qualitätsgarantien, kann jedoch nur mit mittelgroßen Protokollen umgehen. Die α - und Heuristics Miner-Algorithmen bieten diese hohen Qualitätsgarantien nicht, können aber mit großen Protokollen umgehen. Hierfür könnte die Entwicklung von hybriden Algorithmen eine Lösung darstellen, welche die Vorteile der einzelnen Algorithmen miteinander vereinen. Zukünftig ist aufgrund der gewaltigen Datenzunahme zu erwarten, dass Unternehmen noch viel mehr Daten aufzeichnen und die Protokolle dementsprechend größer und komplexer werden. Aufgrund dessen wird die Frage nach der Skalierbarkeit sehr wichtig für zukünftige Process Mining-Vorhaben sein. Bezugsrahmen wie das Inductive Miner directly-follows-Framework für die Prozesserkennung und das Projected Conformance Checking-Framework zur Konformitätsprüfung lassen eine Bewertung neuer Entdeckungs- und Konformitätsalgorithmen zu. Mit gleicher Modellqualität bieten neue Algorithmen dadurch eine wesentlich bessere Skalierbarkeit. Beim Einsatz in Echtzeitanwendungen müssen Unmengen an Daten kontinuierlich verarbeitet werden und die Ausfallsicherheit garantiert werden. Aufgrund der bereits erwähnten zunehmenden Datenmenge und komplexer werdender Organisationsstrukturen mit stetiger Berücksichtigung von Wechselwirkungen innerhalb von Organisationen werden Entwickler von Algorithmen mit Kontext Process Mining stark gefordert sein [IEE12].

4.5 Einbindung von transaktionaler Datenanalyse

Unabhängig vom eingesetzten Process Mining-Tool lässt sich bei der Analyse des Prozessgraphen und darauffolgender Interpretation beobachten, dass die Process Mining-Ergebnisse noch nicht vollumfänglich transparent sind. Hierfür ist die transaktionsbasierte Datenanalyse erforderlich. Der Einsatz einer solchen Analyse lässt sich als nachgelagerter Schritt des Process Mining rechtfertigen. Das Portfolio von Prüfungsschritten ist lang: fehlende Währungsrechnungen, Kunden oder Lieferanten mit Bankkonten in Hochrisikoländern oder Steuerparadiesen, Nichteinhaltung von Embargos, dubiose Transaktionen – um nur einige zu nennen. Die genannten Prüfungsschritte haben miteinander gemein, dass sie nicht aus einem Prozessgraphen heraus zu interpretieren sind. Somit reichen die Prozessgraphen und deren Untersuchung nicht für umfassendere Informationen über die eigenen Unternehmensprozesse aus. Einige Tool-Anbieter haben Lösungsansätze für diese Problematik, eine vollumfängliche

Abdeckung ist aber nicht erkennbar. In der Praxis begegnet man dieser Problematik durch Konsultieren von Datenanalysefirmen, welche mit ausgeklügelten Skripten etwaige Prüfungsschritte aus den Quellsystemen identifizieren können.

Als Beispiel kann der Prüfschritt „Rechnungsdatum vor Bestelldatum“ herangezogen werden (Abbildung 14). Dabei wird das Rechnungsbelegdatum dem Systemdatum der Bestellerfassung gegenübergestellt. Der Prüfschritt selektiert alle Bestellungen mit Erfassungsdatum gleich oder nach dem Belegdatum der ersten Rechnung. Im Prozessgraph wird der betreffende Vorgang prozessual richtig dargestellt und dieser Umstand würde einen Prozessanalysten nicht zum Handeln verleiten lassen. Grund dafür ist die Nutzung der automatisch generierten Systemerfassungsdaten beim Process Mining. Die Belegdatumsangaben bleiben hingegen gewöhnlich außen vor. Hierbei lässt sich erkennen, dass eine Prozessumgehung auch bei der Anwendung von Process Mining durchaus möglich ist.

Invoice date before Purchase Order date

Caseld	Invoice Document Date	Purchase Order Creation	Net value in RC
800414660000010	08.03.2010	21.03.2010 15:36	6.398,13 €
800414660000020	08.03.2010	21.03.2010 15:36	4.418,02 €
800414660000030	08.03.2010	21.03.2010 15:36	247,00 €
800414890000010	22.03.2010	24.03.2010 07:57	889,99 €
800414930000010	24.03.2010	24.03.2010 16:09	11.713,00 €
800414930000020	24.03.2010	24.03.2010 16:09	29.999,00 €
800415030000010	28.03.2010	30.03.2010 10:19	47.861,26 €
800415030000020	28.03.2010	30.03.2010 10:19	30.792,54 €
800415030000030	28.03.2010	30.03.2010 10:19	19.333,75 €
800415030000040	28.03.2010	30.03.2010 10:19	14.150,00 €
800415030000050	28.03.2010	30.03.2010 10:19	798,90 €
Gesamt			166.601,59 €

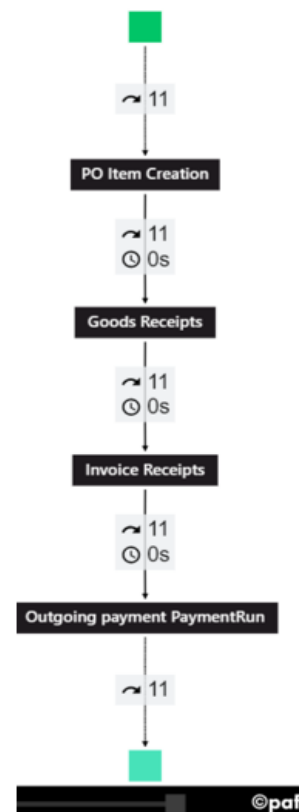


Abbildung 14: Ergebnisse von dab-Prüfschritt "Invoice date before Purchase Order date"

Die Kombination aus Process Mining und transaktionsbasierter Datenanalyse bietet somit eine noch transparentere Einsicht in die eigenen Prozesse und wird empfohlen.

4.6 Einhaltung von Datenschutzbestimmungen

Datenschutz spielt in der heutigen Gesellschaft eine viel größere Rolle als früher. Dazu trägt die Datenschutz-Grundverordnung (DSGVO) bei, welche am 25. Mai 2018 verbindlich in allen Mitgliedsstaaten der Europäischen Union in Kraft getreten ist. Sicherheitsthemen sind in der datengetriebenen Welt essentiell und übertreffen in Sachen Wichtigkeit sogar die Digitalisierungsthemen (siehe Anhang Abbildung 23), da aufgrund von Gesetzesänderungen bei einer Nichtbeachtung hohe Sanktionen drohen [Cap19].

Die Grundlage für Process Mining sind die während der Ausführung von Geschäftsprozessen erfassten Daten in Informationssystemen. Darunter befinden sich auch viele personenbezogene Daten. So werden beispielsweise die Mitarbeiter erfasst, welche Dokumente angelegt und bearbeitet haben. Aber auch Kontaktpersonen im Kunden- und Lieferantenumfeld finden sich in den Unternehmensdaten. Bedenken gibt es, wenn beim Process Mining hochsensible und schützenswerte Ereignisdaten von Quellsystemen extrahiert und in die Software eines Drittanbieters importiert werden. Gegebenenfalls wird die weitergehende Analyse und Bewertung der Process Mining-Ergebnisse sogar ausgelagert und von einem Drittanbieter durchgeführt. Dies erhöht die Möglichkeit von Verstößen gegen die DSGVO-Compliance [Cap19]. Unternehmen müssen Bestimmungen zum Datenschutz strikt befolgen, bei Verstößen sind empfindliche Strafen zu erwarten. Grundsätzlich wird unterschieden, ob Daten primär oder sekundär genutzt werden. Bei primärer Nutzung ist die ausdrückliche Zustimmung der Beteiligten der Normalfall, um den ordnungsgemäßen Ablauf von Geschäftsprozessen zu gewährleisten. Bei sekundärer Datennutzung, wie es beim Process Mining der Fall ist, gestaltet es sich schwieriger mit der Zustimmung. Die sekundäre Nutzung von Daten bezeichnet die Weiterverwendung der Daten außerhalb eines geschäftlichen Vorganges. Das bedingt eine Verschlüsselung der Daten. Es gilt: je höher die Datenschutzanforderung, desto höher auch der Nutzungsverlust und desto schwieriger wird eine Analyse [Aal16, S. 288 ff.]. In [Fah19] werden zwei Ansätze vorgestellt. Zum einen die Bereinigung von Ereignisprotokollen, bei der die Event Logs so vorverarbeitet werden, dass sie selbst bereits bestimmte Datenschutzstufen gewährleisten. Als Modell dienen dabei die k-Anonymität mit Einschränkungen für den Datensatz sowie Differential Privacy mit Einschränkungen für Abfragen über einen Datensatz, um eine Zuordnung unmöglich zu machen. Der zweite Ansatz ist das privatisierte Process Mining, bei dem der verwendete Algorithmus nur die Abfragen erkennt und gebraucht, bei denen der Datenschutz gewährleistet ist. Der Process Mining-Algorithmus muss den neuen Vorschriften angepasst werden. Die Vor- und Nachteile beider Ansätze werden im Referenzpapier erläutert. Die Entwicklung von Datenschutztechniken für die Konformitätsprüfung wird als deutlich weniger sinnvoll erachtet als für die anderen beiden Typen Prozessentdeckung und Prozesserverweiterung. Weitere Herausforderungen bestehen in Bezug auf Handlungsempfehlungen und

Vorhersagen mittels Process Mining, bei der die Einhaltung des Datenschutzes bei Anwendung in Echtzeit schwierig umzusetzen ist. Darüber hinaus ist eine wichtige Anforderung, dass die Process Mining-Ergebnisse weiterhin erzielt werden können und eine entsprechende Interpretation möglich ist [Fah19], [Raf18]. Wird Process Mining über Organisationsgrenzen hinweg im Rahmen einer ganzheitlichen Betrachtung der Supply Chain zur Effizienzsteigerung verwendet, muss im Vorfeld ein einheitlicher Konsens über die Bedingungen der Durchführung geschaffen werden. Unter dem Gesichtspunkt des Datenschutzes gilt es insbesondere an den organisationsübergreifenden Schnittstellen, erhöhte Aufmerksamkeit walten zu lassen. In der Praxis kann diesem Problem begegnet werden, indem die Daten bereits in verschlüsselter Form geladen werden.

Der verantwortungsvolle Umgang mit Daten adressiert aber nicht nur externe Prozessbeteiligte. Auch die eigenen Mitarbeiter haben Anspruch auf Datenschutz. Process Mining sollte zur Entdeckung und Verbesserung von Prozessen angewendet werden. Keinesfalls darf es Anwendung zur Überwachung und Leistungsbewertung von Mitarbeitern finden, was in Deutschland ohnehin nicht erlaubt ist. In diesem Zusammenhang kommt nochmals die Wichtigkeit der zielgerichteten Datenextraktion auf. Nur die für das konkrete Analyseziel notwendigen Daten sollten extrahiert werden.

5 Potentiale von Process Mining

Wie alle neuartigen Methoden ist auch Process Mining mit innovativen Forschungsansätzen konfrontiert und wird rasch weiterentwickelt. Process Mining ist ein toolbasiertes Werkzeug, mit dem die digitale Transformation von Geschäftsprozessen unterstützt werden kann [Ker19]. Potentiale zur Verbesserung von Effizienz, Agilität und Einhaltung der Compliance werden identifiziert und im Nachgang umgesetzt. Dieses Kapitel wird einige interessante zukünftige Entwicklungen von Process Mining beleuchten, welche aus der einschlägigen Literatur entnommen sind.

5.1 Vorstufe zu innovativen Technologien

Die Process Mining-Anbieter verstehen die toolbasierte Methode zunehmend als Vorstufe von aufstrebenden Technologien der digitalen Transformation. Für Data Mining wurden bereits viele dieser Techniken umgesetzt. Jedoch lassen sich diese nicht ohne Weiteres auf Process Mining übertragen, da sich Prozesse komplexer als Tabellen verarbeiten lassen. Bei möglichen Anbindungstechnologien werden häufig Robotic Process Automation (RPA) und Digital Twin of an Organization (DTO) genannt. Sowohl RPA als auch DTO sind gemäß der Capgemini-Studie „IT-Trends 2019“ weit vorne in den Tech-Trends des Jahres 2019 [Cap19]. Diese Technologien und im speziellen die Einbindung und gleichzeitige Nutzung von Process Mining werden nun thematisiert.

Process Mining in Verbindung mit RPA wird in [Gey18] beschrieben. Mit RPA wird der Ansatz nach Automatisierung und Standardisierung von Prozessen durch Software-Roboter verfolgt. Dafür eignen sich insbesondere sich häufig wiederholende, fehleranfällige, skalierbare und nicht zu komplex ablaufende Aktivitäten in standardisierten Geschäftsprozessen. Der Einsatz von RPA ist für diese Prozessschritte konzipiert, um sie effizienter und konsistenter zu gestalten. RPA benötigt zumeist unterstützende Technologien. Process Mining wird als Wegbereiter und Schlüsselfaktor zu einer erfolgreichen Umsetzung von RPA-Initiativen angesehen. Der kombinierte Einsatz wird bereits praktiziert, bietet aber zukünftig noch deutlich größeres Potential in Akzeptanz und Wachstum. Gemäß einer Capgemini-Studie [Cap19] wird RPA in Bezug auf Geschäftsprozesse eines der wichtigen Themen für Unternehmen in den kommenden Jahren. Beim RPA werden Prozessabläufe aufgezeichnet und menschliche Aktivitäten durch Roboter nachgebildet. Konkret wird Process Mining hierbei zur Bewertung der Reife und des Automatisierungspotentials von Geschäftsprozessen herangezogen. Mit höherer Automatisierungsreife steigt auch die Bereitschaft zu RPA. Vor einer RPA-Implementierung werden mithilfe von Process Mining verschiedene Roboterautomatisierungsansätze simuliert und verglichen, was so zur faktenbasierten Auswahl und Priorisierung des leistungsstärksten Ansatzes beiträgt. Diese objektive Vorgehensweise weist deutliche Kosten- und Zeiteinsparungen im

Vergleich zur subjektiven manuellen Auslotung von Automatisierungspotentialen und dem darauffolgenden Leistungsvergleich auf. Implementierte Roboter können mit Process Mining sowohl überwacht werden als auch anderen Robotern als Benchmarks gegenübergestellt und bewertet werden.

Dabei wird ein Ablauf in drei Stufen empfohlen [Gey18]. Als erster Schritt wird das RPA-Potential abgeschätzt. Geachtet wird dabei vor allem darauf, dass die infrage kommenden Vorgänge nicht zu komplex und variantenreich sind. Einer wirtschaftlichen Betrachtung folgt eine Einschätzung zur Standardisierbarkeit des Prozesses. Die Komplexität von Prozessen sollte vor der RPA nach Möglichkeit weiter homogenisiert werden, um einen höheren Profit aus der RPA-Initiative zu ziehen. Daraufhin werden Ziele definiert und eine RPA-Applikation entwickelt, bei dem Roboter auf die Ausübung der Aufgaben trainiert werden. Bei dem Schritt zur Auswertung der erzeugten Vorgangsinstanzen kommt Process Mining zum Einsatz. Roboterleistungen werden verglichen und der effizienteste Roboter wird für diese Instanz eingesetzt. Durch die ständige Überwachung der Roboter mithilfe von Process Mining können Kosteneinsparungen beziffert werden. Selbsterklärend lässt sich feststellen, dass Prozessvarianten mit niedrigerer Automatisierungsrate höheres Einsparpotential bieten als welche, die bereits hoch automatisiert sind. In [Gey18] wurde zudem eine RPA-Initiative anhand eines P2P-Prozesses aus SAP und unter Bezugnahme der Process Mining-Software von Celonis bildhaft dargestellt. IT-Landschaften werden immer komplexer und vernetzter. Um dies im Geschäftsalltag optimal bewältigen zu können, nutzen etwa drei Viertel der Organisationen Möglichkeiten zur höheren Automatisierung [Cap19]. Die Einbindung von Process Mining in RPA-Initiativen oder RPA-Projekten bietet sich an. Der Prozessgraph mit Zeitinformationen lässt erkennen, welche Aktivitäten häufig durchgeführt werden. Daraus muss eine Beurteilung erfolgen, ob jene Aktivität automatisiert werden kann. Ist dies machbar, können dabei deutliche Reduzierungen der Durchlaufzeiten erreicht werden. Die Verknüpfung von Process Mining und RPA befindet sich unter anderem bereits bei der Process Analytics Factory GmbH in Anwendung. Insgesamt lässt sich ein zunehmender Aufbau von Kooperationen zwischen Process Mining- und RPA-Anbietern verzeichnen [Ker19]. Damit wird die Entscheidung gestützt, welche Prozesse überhaupt automatisiert werden können. Bei bereits automatisierten Vorgängen kann mit Process Mining ermittelt werden, ob eine RPA effizient arbeitet oder nicht. Auch eine Konformitätsprüfung ist hierbei denkbar. Die Automatisierung von Prozessen wird laut Umfrage (siehe Anhang Abbildung 23) als wichtigstes Thema des Prozessmanagements gesehen [Cap19].

Von der Technologie DTO ist ebenso wie bei RPA immer wieder die Rede im Zusammenhang mit den Entwicklungen im Rahmen der Industrie 4.0. Dabei werden reale Unternehmensstrukturen und -vorgänge in der digitalen Welt nachgebildet. Mit dieser Nachbildung können alle Aus- und Wechselwirkungen infolge von Entscheidungen simuliert und getestet werden und so den Entscheidern einen besser kontrollierbaren Rahmen ermöglichen. Eine Intensivierung von prädiktiven Analysen und Szenario-Tests wird erwartet. Die Steuerung der Digitalinitiative mit der Datenverbindung könnte dabei dem Process Mining zugesprochen werden. Zwar fehlen bei der Verknüpfung von Process Mining mit der DTO-Technologie noch Best Practice-Fälle und sie befindet sich noch in der Anfangsphase, nichtsdestotrotz denken Process Mining-Anbieter über Ansatzpunkte und Implementierungen für diese Simulationen nach und erkennen den vielversprechenden Ansatz zur zukünftigen Bewältigung der Herausforderungen, welche die digitale Transformation mit sich bringen [Bay19], [Ker19].

5.2 Anwendung in Echtzeit zur operativen Unterstützung

Process Mining wird meist mit historischen Daten durchgeführt und es werden bereits abgeschlossene Aktivitäten analysiert, um die Prozesse für zukünftig auftretende Cases robuster und effizienter zu gestalten. Die Aktualisierung von Datenquellen in Echtzeit und ausreichend hohe Rechenleistungen zur Analyse der einzelnen Aktivitäten und Cases lassen die Erwartungshaltung an die toolbasierte Methode weiter wachsen. Nach [Cap19] nutzten im Jahr 2018 bereits 37,7 Prozent der befragten Unternehmen Realtime Intelligence (Informationen in Echtzeit), respektive befanden sich in dessen Implementierungsphase. Das Echtzeit-Monitoring von Durchlaufzeiten ist nur einer von vielen Ansätzen bei der Unterstützung von Entscheidungen. Diese Online-Entscheidungsunterstützung lässt Mitarbeitern direkt Abweichungen anzeigen, die dann zielgerichtete Maßnahmen ergreifen können. Signalmeldungen dafür gibt es bereits [Aal16, S. 304], ihr Einsatz zur operationellen Unterstützung könnte aber zukünftig intensiviert werden. Potential wird auch bei Vorhersagen und gar Handlungsempfehlungen für laufende Cases aus der Process Mining-Software heraus gesehen. Cases sollen dort über die gesamte Prozesslaufzeit begleitet werden [Eck15]. Für das Erlernen von Prognosen werden historische mit aktuellen Informationen kombiniert [Aal16, S. 301f.]. Durchgehend hohe Rechenleistungen zur Verknüpfung von Echtzeitdaten und eine stets hohe Qualität der zugrundeliegenden Daten sind dafür Voraussetzung [Eck15], [IEE12]. Es ist davon auszugehen, dass die Anwendung von Process Mining in Echtzeit zur operativen Unterstützung zunächst von großen Organisationen umgesetzt wird [Bur15, S. 4].

5.3 Machine-Learning-Algorithmen und Künstliche Intelligenz (KI)

In Kapitel 4.4 wurden Herausforderungen zu den Process Mining-Algorithmen ausgearbeitet. Die dortigen umfangreichen Erläuterungen lassen vermuten, dass in Zukunft viel Wert auf die Weiterentwicklung der Algorithmen zu legen sein wird. In [Bay19] wird das Vorhaben vom Anbieter Lana Labs erläutert, Machine-Learning-Algorithmen für die Ursachenerkennung von Ineffizienzen in die Software zu implementieren und durch Künstliche Intelligenz (KI) die Phasen der Planung, Modellierung und Datenaufbereitung zu Event Logs problemloser zu verwirklichen. Damit können die heute notwendigen und umfangreichen Prozesskenntnisse und technischen Domainkenntnisse der Quellsysteme durch selbstlernende Ansätze und Handlungsempfehlungen unterstützt werden. Dies würde Process Mining für viele weitere Zielgruppen interessant erscheinen lassen. Die Anbieter erwarten sich durch Machine-Learning und KI für den Einsatz beim Process Mining eine noch fortschrittlichere und deutlich effizientere Analyse [Ker19]. Vor allem auch die Fehlerhäufigkeiten, welche bei der Modellierung auftreten können, würden wohl drastisch abnehmen. Process Mining in Verbindung mit KI in der Umgebung der Industrie 4.0 weist ein spannendes Zukunftsszenario auf, bei dem neuwertige und effektvolle Erkenntnisse für die Unternehmensorganisation generiert werden können.

5.4 Organisationsübergreifendes Process Mining

In der heutigen Unternehmensumwelt ist eine zunehmende Verflechtung existierender Unternehmensnetzwerke zu beobachten. Ein bekanntes Beispiel für diese Kollaborationen sind die Automobilhersteller und ihre Zulieferer. Hier werden Informationssysteme von Zulieferern direkt an die der Automobilhersteller angebunden, um die für die Automobilindustrie so wichtige Just-In-Time-Lieferung gewährleisten zu können. In anderen Fällen machen es Autohersteller sogar zur Bedingung, dass das eigene ERP-System von den Zulieferern mitgepflegt wird. Dazu hat sich der Begriff Supply Chain oder mittlerweile auch das Supply Network in der Organisationswelt etabliert, welcher die ganzheitliche Betrachtung der Prozesskette vom Lieferanten zum Endverbraucher miteinschließt. Im Zuge dessen wird organisationsübergreifendes Process Mining an Bedeutung gewinnen. Das ermöglicht Optimierungspotentiale in erweitertem Umfang. Dabei ist nicht lediglich die Kosteneinsparung ein Thema, sondern das gesamte Vorhaben lebt vom Erfahrungs- und Wissensaustausch über die Thematik Process Mining. Um die Verwässerung und den unkontrollierten Ablauf von fachspezifischen Interna jedoch einzudämmen, sind die in Kapitel 4.6 genannten Punkte zur Einhaltung von Datenschutzbestimmungen zwingend einzuhalten.

6 Anbieter von Process Mining-Tools

Dieses Kapitel soll einen Überblick über den Process Mining-Markt verschaffen (Abschnitt 6.1). Darüber hinaus werden führende Anbieter mit deren kostenfreien Versionen vorgestellt und diese abschließend einer Bewertung unterzogen (Abschnitt 6.2).

6.1 Process Mining Markt

Der Markt für Process Mining-Software entwickelt sich schnell. Organisationen fragen verstärkt Process Mining Tools nach. Es ist eine exponentielle Zunahme des Bewusstseins für diese toolbasierte Methode zu vernehmen. Folgerichtig gerät es vermehrt in den Fokus vieler Unternehmen [Ker19]. Dies ist auch der Grund, warum viele Softwareanbieter ihre bereits bestehenden Tools um Process Mining-Funktionalitäten erweitern und stetig neue kommerzielle Tools in den Markt eintreten [Aal16, S. 21]. Auch in der Forschung wird dieses Gebiet stark bearbeitet, was an den zunehmend erscheinenden Publikationen abzuleiten ist [Aal15], [Gar19, S. 15]. Im Gartner Market Guide for Process Mining [Ker19] wird der Markt für das Jahr 2018 auf 160 Millionen US-Dollar beziffert. Im Jahr 2017 wurde laut Gartner noch eine Erhöhung bis zum Jahr 2019 um den Faktor 3 bis 4 erwartet, was jedoch deutlich verfehlt wurde. Demnach wird sich das starke Wachstum vor allem deshalb verzögern, weil die Anbieter die große Nachfrage derzeit noch nicht bedienen können. Grund für die ambitionierte Wachstumsprognose bietet insbesondere der US-Markt, der die Nachfrage nach Process Mining-Software nun erst intensiviert. Ausgenommen von diesen Zahlen ist der vielversprechende Folgemarkt mit Beratung und Dienstleistung rund um die Implementierung und Anwendung der Tools. Im Wissen dessen, dass der Folgemarkt die Erlöse für die Software bei vielen Anbietern teils deutlich übersteigt, ist der durch Process Mining generierte Gesamtmarkt erheblich höher. Eine Übersicht der Anbieter ist in Abbildung 15 dargestellt.



Abbildung 15: Process Mining Anbieter [Rob18]

Der Gesamtmarkt wird auf etwa 25 Anbieter geschätzt. Demnach ist, gemessen an der reinen Anzahl, in Deutschland die größte Verbreitung von Process Mining-Softwareanbietern zu verzeichnen, gefolgt von den Niederlanden und den USA. Deutschland und die Niederlande steuern zudem die meiste Innovationsarbeit im Gebiet des Process Mining bei. Die höchsten Wachstumsraten und größten Marktchancen werden in Nordamerika verortet [Rob18], [RAM18].

6.2 Führende Produkte

Wie im vorherigen Kapitel verdeutlicht wird, ist der Markt für Process Mining-Software in den letzten Jahren kontinuierlich gestiegen. Die Prognose für die nächsten Jahre ist weiterhin positiv. Diese lukrative Marktaussicht führt zu einem zunehmenden Wettbewerbsaufkommen im Process Mining-Markt. Im Folgenden werden ausgewählte, führende Anbieter und deren Tools vorgestellt und verglichen. Von allen kommerziellen Anbietern (6.2.1 – 6.2.4) wurden akademische Testlizenzen für den Zeitraum der Ausarbeitung dieser Arbeit zur Verfügung gestellt. Diese, und nicht die umfangreicheren kommerziellen Vollversionen, werden für den folgenden Vergleich als Grundlage verwendet. Zusätzlich wird das Open-Source-Tool ProM der Process Mining Group (6.2.5) für den Vergleich herangezogen.

6.2.1 Celonis

Die Celonis SE mit Hauptsitz in München hat derzeit die Marktführerschaft im Process Mining-Markt inne und beschäftigt weltweit bereits über 500 Mitarbeiter². Die im Jahr 2011 gegründete Firma wurde 2015 rasch zur schnellst wachsenden Tech-Firma in Deutschland [DDW15] und ist eine von nur sechs Unternehmen in Deutschland mit einer Milliarden-Bewertung (knapp über 300 weltweit) und verdiente sich die Auszeichnung „Einhorn“, welches junge, innovative Unternehmen mit einer Bewertung von über einer Milliarde US-Dollar erhalten, die noch vor dem Börsengang oder einem Exit stehen [WIW19], [Ben19].

Celonis Snap wird über eine eigene Intelligent Business Cloud realisiert. Der Datenimport kann über Drag & Drop erfolgen. Ein Feld-Separator für CSV-Dateien muss umgehend festgelegt werden. Im weiteren Vorgehen werden für die einzelnen Attribute sowohl Datentyp als auch Format (beispielsweise Datums- und Uhrzeitangaben für den Zeitstempel) vorgeschlagen, welche jedoch geändert werden können. Nach Festlegung der Mindestanforderungen an ein Event Log wird online ein Prozessmodell im ersten Reiter „Overview“ erzeugt (Abbildung 16). Dabei kann das Ereignisprotokoll als Animation wiedergegeben werden. Dies lässt eine sekundengenaue Analyse und Nachverfolgung einzelner Prozessschritte zu. Cases oder Aktivitäten können auch einer Einzelanalyse unterzogen werden.

² www.celonis.com/de/company

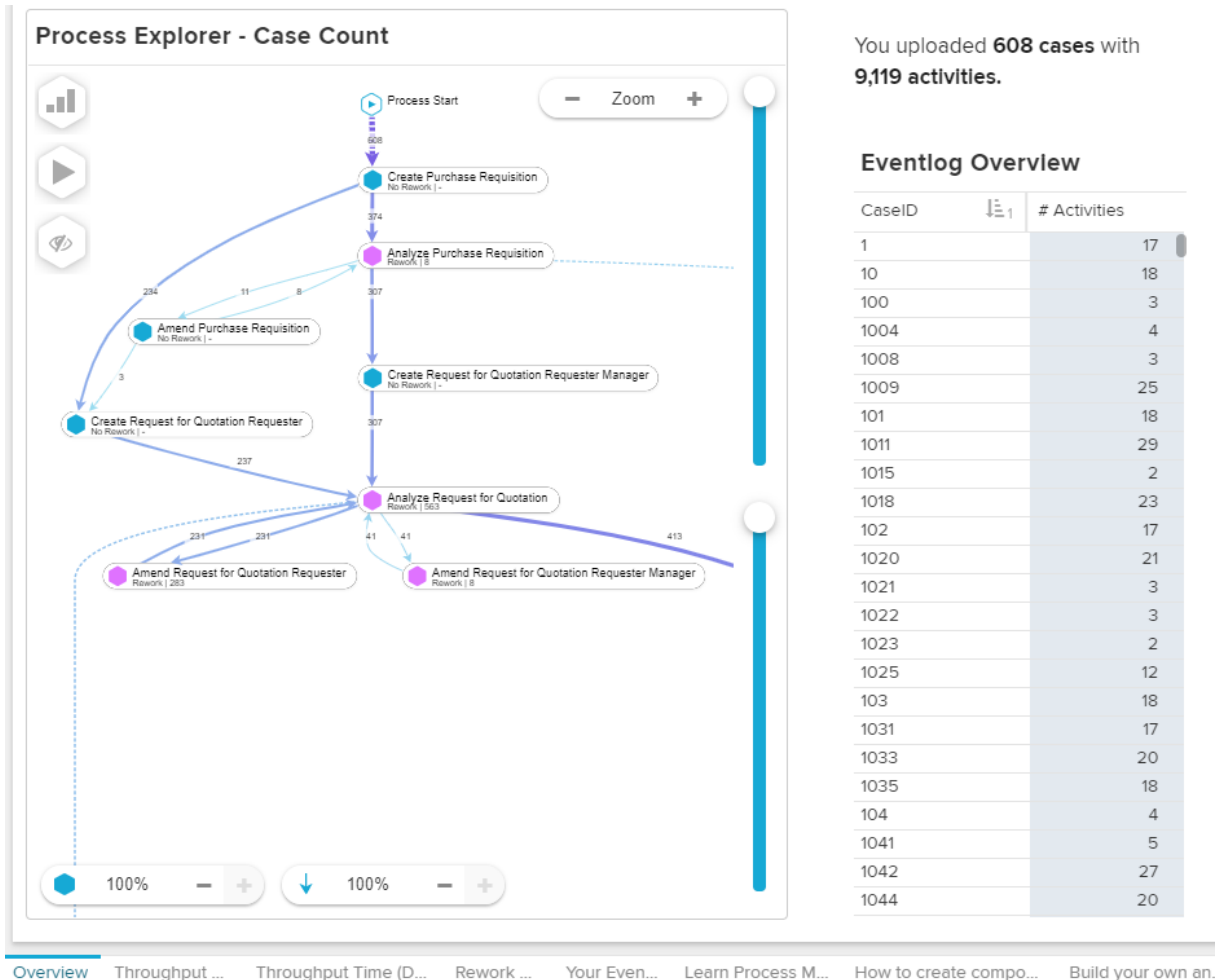


Abbildung 16: Reiter "Overview" von Celonis Snap

Im Reiter „Throughput“ werden die Durchlaufzeiten und die durchschnittlichen Durchlaufzeiten aller Cases angegeben. Der Reiter „Rework Rate“ (siehe Anhang Abbildung 26) gibt alle Cases an, bei denen mindestens eine Aktivität im Prozessverlauf häufiger als einmal ausgeführt wurde. Auf unseren Beispieldatensatz bezogen beträgt die Rate auffällige 41 Prozent.

In jedem Reiter sind zwei Regler für Aktivitäten und Pfade zwischen null und hundert Prozent einstellbar. Darüber hinaus sind Pfeilinformatoren als Häufigkeits- und Zeitangaben wählbar. Die weiteren Reiter geben Informationen rund um die Anwendung von Celonis Snap.

6.2.2 Fluxicon

Fluxicon BV ist seit vielen Jahren im Process Mining-Markt mit dem Tool Disco vertreten. Das Unternehmen mit Gründungsjahr 2009 ist in Eindhoven in den Niederlanden beheimatet. Hervorgegangen ist das Unternehmen aus dem Open-Source-Tool ProM, um den Fokus von der Entwicklersoftware hin zur kommerziellen Endanwendersoftware zu legen. Fluxicon trägt weiterhin zur Entwicklung der ProM-Plattform bei. Als Mitglied der Institute of Electrical and

Electronics Engineers (IEEE) Task Force on Process Mining ist Fluxicon stark in Forschungsaktivitäten involviert³.

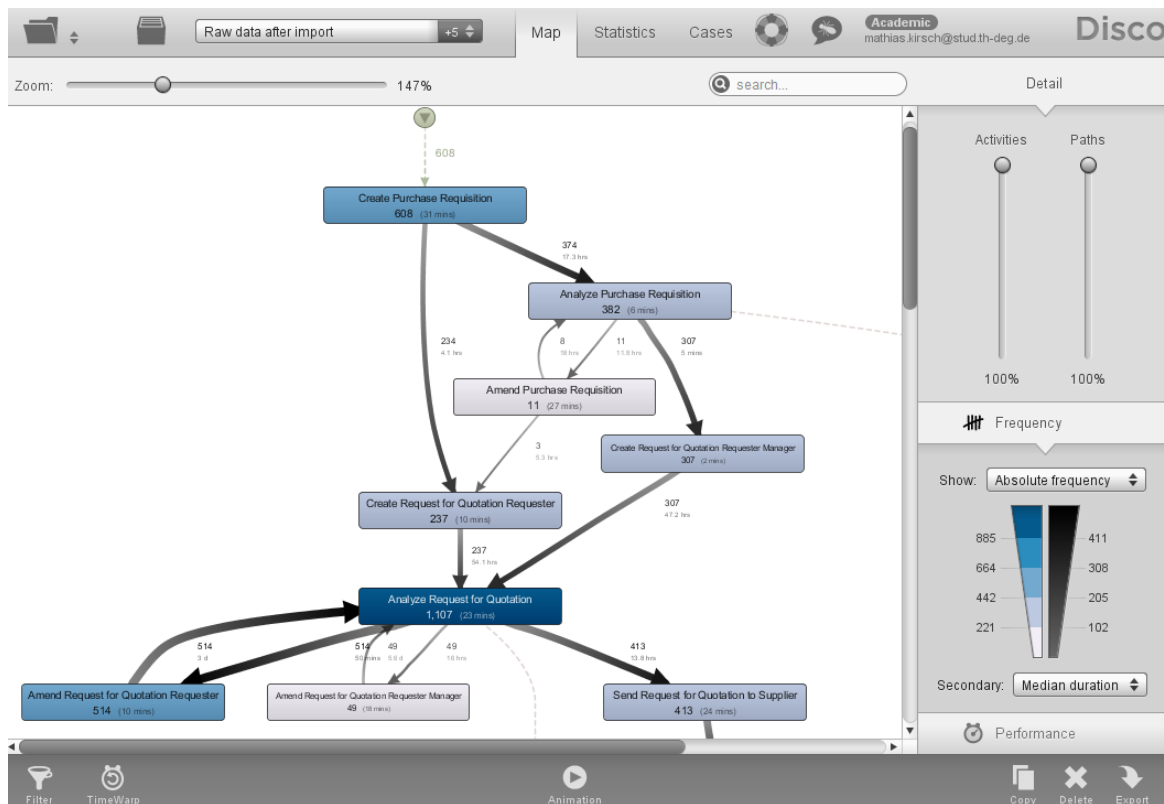


Abbildung 17: Oberfläche des Disco-Tools von Fluxicon

Disco wird über eine Applikation aufgerufen. Die Oberfläche von Disco ist dreiteilig aufgebaut. Der Reiter „Map“ (Abbildung 17) liefert die graphische Prozessdarstellung des importierten Event Logs. Die Einfärbung der Aktivitäten gibt visuellen Aufschluss über die Häufigkeit der Durchführungen innerhalb des Event Logs. Je dunkler die Einfärbung, desto höher die Aktivitätshäufigkeit. Dies gilt für die beiden Darstellungsformen Häufigkeit und Leistung. Ähnlich verhält es sich mit den Pfaden zwischen den Aktivitäten. Häufig auftretende Verbindungen sind dunkler und dicker dargestellt. Dabei ist die genaue Anzahl neben den Verbindungspfeilen angegeben. Ein Mausklick auf eine Aktivität oder einen Pfeil öffnet ein Fenster, welches Informationen zu absoluter Häufigkeit, Fallhäufigkeit, maximale Wiederholungen sowie die Angabe, wie häufig eine Aktivität eine Start- oder Endaktivität repräsentiert, liefert. Dazu kommen Leistungsdaten wie Gesamtdauer, mediane Dauer, durchschnittliche Dauer sowie die minimale und maximale Dauer von Aktivitäten und Pfaden. Über dieses Fenster kann in den umfangreichen Filterbereich gewechselt werden. Der Detailgrad kann über Regler für Aktivitäten und

³ www.fluxicon.com

Pfade zwischen null und hundert Prozent angepasst werden. Eine Animation des Ereignisprotokolls im Graphen kann abgespielt und als Video exportiert werden. Dies lässt eine sekundengenaue Analyse und Nachverfolgung einzelner Prozessschritte zu.

Der zweite Reiter „Statistics“ (siehe Anhang Abbildung 24) hält sehr umfangreiche Leistungszahlen sowohl zum Gesamtprozess als auch für die einzelnen Cases und Varianten für den Nutzer bereit. Informationen können bis auf die einzelne Aktivität, Ressource und Rolle innerhalb eines Cases heruntergebrochen werden. Der dritte Reiter „Cases“ (siehe Anhang Abbildung 25) listet alle Varianten und die dazugehörigen Cases einzeln als Graph und Tabelle auf. Ein Zeitstrahl mit allen Zeitinformationen zur Einordnung eines Cases in die Gesamtzeitspanne des betrachteten Event Logs ist gegeben. Neben den drei Reitern ist ein Rettungsring-Symbol angebracht, welches dem Anwender Informationen zur Bedienung und den Möglichkeiten des Tools bietet.

6.2.3 Lana Labs

Die Lana Labs GmbH mit Sitz in Berlin ist seit dem Jahr 2016 auf dem Process Mining-Markt vertreten und beschäftigt über 30 Mitarbeiter. LANA Process Mining ist mit einem einzigartig im Markt vertretenen komplexen Machine-Learning-Algorithmus versehen⁴.

LANA Process Mining ist eine cloudbasierte Lösung. Beim Datenimport werden zunächst die Protokollattribute konfiguriert und es folgt ein Abgleich zwischen den Attributen aus dem Protokoll mit den vorgegebenen Attributen. Der Anwender kann bei Bedarf Änderungen vornehmen. Grundsätzlich wird zunächst definiert, ob man mit einem generalisierten Prozessmodell arbeiten möchte oder einem Prozessmodell, welches sämtliche entdeckte Ausprägungen beinhaltet. Sechs Reiter stehen in der kostenfreien Testlizenz zur Verfügung.

Der Reiter „Model“ (Abbildung 18) zeigt das entdeckte Prozessmodell an. Auffällig ist die horizontale Prozessdarstellung. Visuelle Instrumentarien wie dickere Pfade für häufiger auftretende Verbindungen und rote Pfade für Engpässe helfen bei der Analyse und Ergebnisinterpretation. Häufigkeits- und Zeitangaben werden angegeben, dabei gibt ein Mausklick auf die Aktivitäten noch konkretere Informationen zu minimalen und maximalen, durchschnittlichen und medianen Durchführungszeiten. Ferner finden sich auf dieser Ebene auch Filtermöglichkeiten. In diesem Reiter kann ein Vergleich zwischen dem entdeckten und einem Ziel-Prozessmodell durchgeführt werden. Das Zielmodell kann dabei entweder hochgeladen werden oder mit dem integrierten BPMN-Tool von bpmn.io selbst erstellt werden.

⁴ www.lanalabs.com/unternehmen



Abbildung 18: Reiter "Model" von LANA Process Mining

Im Reiter „Variants“ wird die Häufigkeit aller entdeckten Varianten angegeben und es kann direkt auf Case- und Event-Ebene heruntergebrochen werden. Leistungsdaten zu den einzelnen Events und deren Folgeevents werden im Reiter „Statistics“ (siehe Anhang Abbildung 27) zur Verfügung gestellt. Die Reiter „Action“ und „Root Cause“ bieten umfangreiche Analyse-möglichkeiten, welche Antworten auf konkrete Fragestellungen über die eigenen Prozesse ge-ben. Der Reiter „Dashboard“ erlaubt es zudem, Visualisierungen nach eigenen Wünschen zu kreieren und so den eigenen Bedürfnissen nach Dashboards mit entsprechenden KPIs gerecht zu werden. Generell bietet LANA Process Mining dem User Hilfestellungen in Form von Chat-boxen an, bei denen direkt mit Support-Mitarbeitern interagiert werden kann.

6.2.4 Process Analytics Factory

Die Process Analytics Factory GmbH, kurz PAF, hat ihren Firmensitz in Darmstadt. Das 2014 gegründete Unternehmen hat derzeit über 25 Mitarbeiter beschäftigt. Die Besonderheit ist die Implementierung der Process Mining-Software mit eigenen Visualisierungen in die Business Intelligence (BI)-Lösung von Microsoft, genannt Microsoft Power BI. PAF macht sich zu Nutze, dass sich Microsoft Power BI bereits in vielen Unternehmen im Einsatz befindet. Power BI ist ein intuitives Business Analytics-Tool, mit dem Daten visualisiert und in Berichten und Dash-boards verarbeitet werden können. Mit der Plattform belegt Microsoft seit Jahren den Spitzen-platz im Gartner Magic Quadrant im Bereich „Analytics and Business Intelligence Platforms“ [How19]. Nach Aussage von Tobias Rother, Geschäftsführer von PAF, verfolge man die De-mokratisierung von Process Mining. Dies impliziert die Verwendung des Tools direkt durch den Endanwender, um fachspezifische Fragestellungen innerhalb eines Prozesses zielgerichteter beantworten zu können⁵.

⁵ www.pafnow.com/de/paf

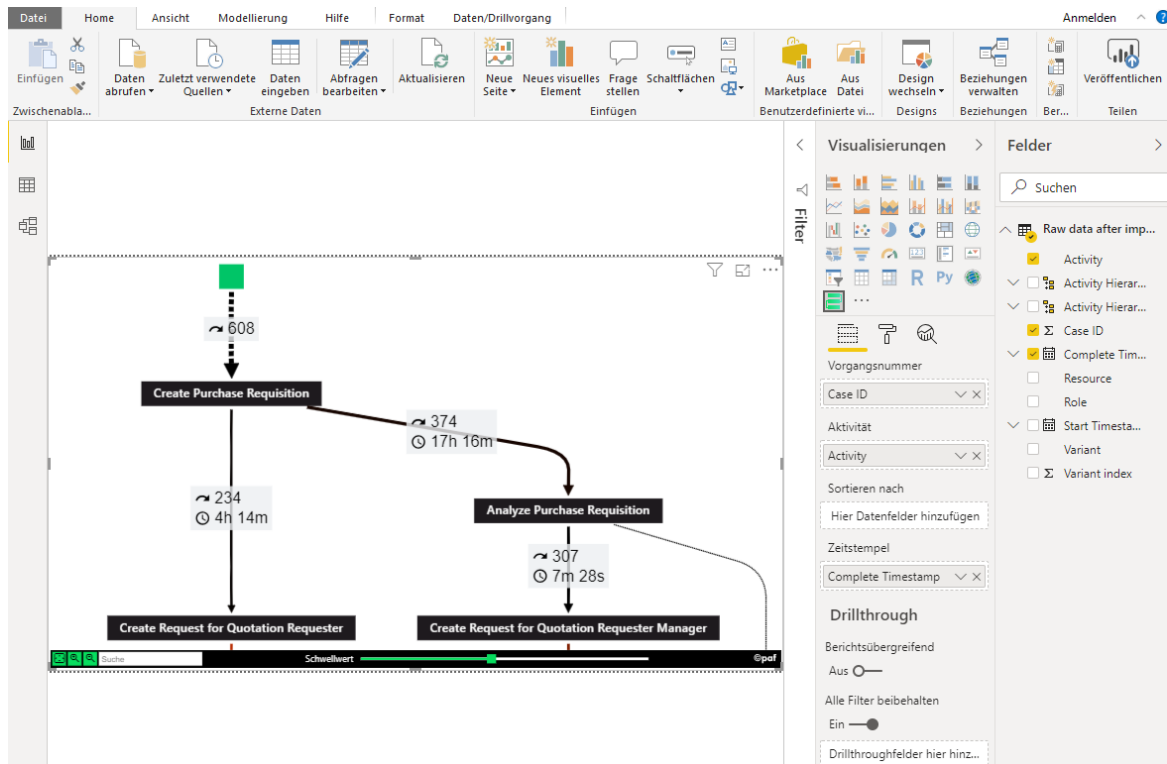


Abbildung 19: PAFnow in Microsoft Power BI

Zunächst wird die Plattform Power BI von Microsoft gestartet. Um Process Mining mit PAFnow anwenden zu können, muss das Custom Visual PAFnow Process Mining vom Microsoft App-Store heruntergeladen werden. Daraufhin können die Quelldaten direkt in das Custom Visual geladen werden. Die Auswahl der Daten erfolgt über die Schaltfläche „Daten abrufen“. Dabei kann eine große Auswahl an Quellen gewählt werden – verschiedene Dateitypen, Datenbanken, Onlinedienste und viele weitere. Es folgt die Verbindung und der Import der Quelldaten. Der Reiter „Discovery“ ist standardmäßig in der kostenlosen Variante von PAFnow Process Mining verfügbar. Die Darstellung kann der Abbildung 19 entnommen werden. Dieser verfügt über einen Regler für die Angabe von Schwellenwerten sowie Häufigkeits- und Zeitinformationen zu den einzelnen Pfaden. Dabei wird automatisch mit visuellen Hilfen unterstützt. Etwa werden häufig auftretende Verbindungen dicker und dunkler dargestellt und Engpässe in roter Signalfarbe. Bei der Navigation des Cursors über einen Pfad oder eine Aktivität werden weiterführende Informationen angezeigt. Bei der kostenfreien Variante ist auffällig, dass nur der Start- oder Endzeitstempel von Aktivitäten für die Darstellung angewählt werden kann und nicht beide. Dies führt zu Einschränkungen bei einer Leistungsanalyse.

6.2.5 Process Mining Group

Die Process Mining Group der Technischen Universität Eindhoven bietet mit ProM eine kostenfreie Plattform für Process Mining an. Es ist die Standardanwendung für wissenschaftliche

Ansätze und richtet sich primär an Entwickler, die für akademische Zwecke eigene Algorithmen anfertigen oder Bestehende erweitern wollen⁶. Die Software ragt vor allem durch eine große Anzahl von über 1.500 Plug-ins heraus, welche im Rahmen der Forschung laufend entwickelt werden [Aal16, S. 334]. Der Großteil der Forschungspublikationen bedient sich der ProM-Plattform aufgrund des einfachen Zugangs [Gar19, S. 37]. Die Benutzeroberfläche ist von Fluxicon BV konzipiert⁷.

Das Programm zeichnet sich durch eine Dreiteilung im oberen Bereich aus, welche die Schrittfolge darstellt. Jedes resultierende Ergebnis kann für weitere Aktionen und Ansichten verwendet werden. So ist eine immer detailliertere Analyse möglich. Im Reiter „Workspace“ findet die Datenauswahl statt. Dies kann über Favoriten, einen bereits importierten Datensatz, eine Auswahl oder einen neuen Import in den Workspace geschehen. Im Reiter „Actions“ (Abbildung 20) wird zunächst die importierte CSV-Datei direkt in eine XES-Datei konvertiert. XES ist die Kurzform für Extensible Event Stream und wurde von der IEEE Task Force of Process Mining als Standardformat für Process Mining festgelegt [Aal16, S. 138]. Zu den vielen Plug-ins gilt es bei der jeweiligen Auswahl zu erwähnen, dass grün markierte Aktionen mit der XES Event Log-Inputdatei kompatibel und somit ausführbar sind, während gelb markierte Aktionen mit der gegebenen Inputdatei nicht ausgeführt werden können.

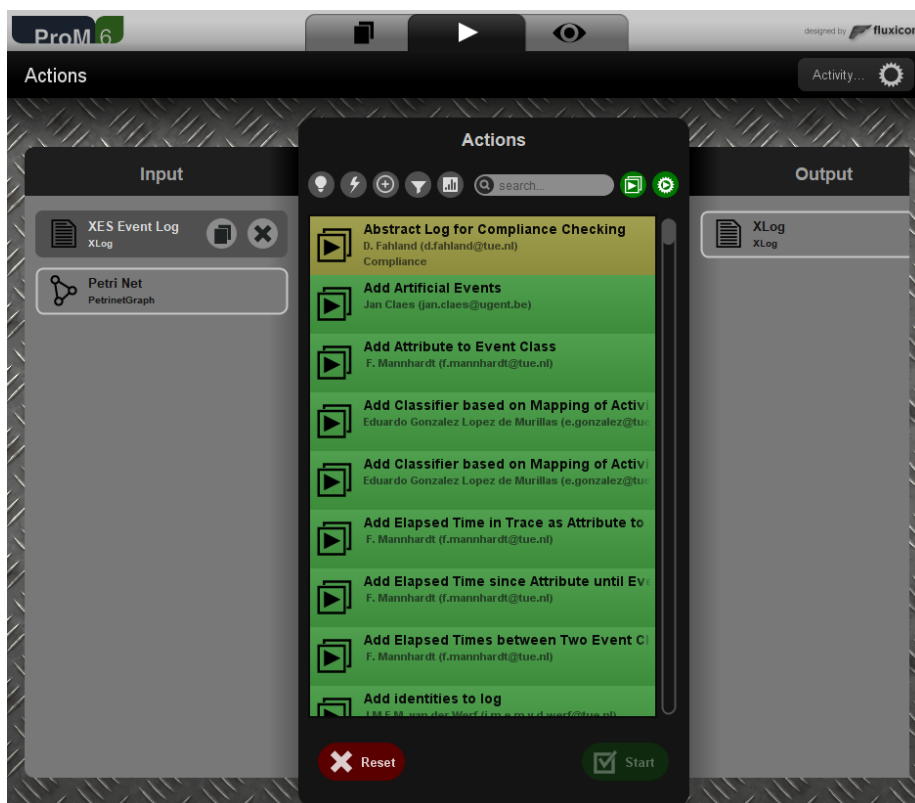


Abbildung 20: Reiter "Actions" von ProM

⁶ www.processmining.org/start

⁷ www.fluxicon.com

Der Reiter „View“ (siehe Anhang Abbildung 28) bietet zunächst erneut eine Dreiteilung. Die Dashboard-Ansicht liefert eine Übersicht zu Events und Eventklassen pro Case sowie Schlüsseldaten zum verarbeiteten Protokoll. Die Ansicht „Inspector“ lässt eine Analyse auf Case-Ebene zu mit chronologischen Prozessangaben und allen Attributdaten. Die Summary-Ansicht stellt die Protokollzusammenfassung dar. Absolute und prozentuale Häufigkeiten von Eventklassen und Traces sowie alle Events und deren Auftretenshäufigkeit als Start- oder End-Event werden ersichtlich. Für die konkrete Darstellung des Prozessgraphen sind weitere Aktionen über den Button „Use resource“ erforderlich, wie etwa „Mine with Inductive visual Miner“ oder „Mine process tree with Inductive Miner“. Die Visualisierung „Inductive visual Miner“ liefert die bewährten Regler-Einstellungen für Aktivitäten und Pfade (siehe Anhang Abbildung 29).

6.3 Vergleich

Die Process Mining-Tools der vorgestellten Anbieter werden nun einem objektiven Vergleich unterzogen. Um der Gegenüberstellung eine einheitliche Basis zu bieten, ist die Verwendung eines einheitlichen Datensatzes zielführend. Dies wird in dieser Arbeit anhand eines Datensatzes mit 608 Cases mit 24 verschiedenen Einzelaktivitäten, 9.119 Events und 98 Prozessvarianten sichergestellt. Der Datensatz ist dem Beschaffungsbereich eines Unternehmens zuzuordnen. Im Kontext einer ERP-Umgebung hat sich dafür die Bezeichnung Purchase-to-Pay-Prozess (P2P-Prozess) etabliert. Der relativ einfach gehaltene Beispieldatensatz ist keineswegs mit realen Datensätzen in Produktivumgebungen zu vergleichen, er sollte jedoch zum Zwecke eines Vergleichs ausreichen.

Die vorgestellten Tools weisen unterschiedliche Implementierungsarten auf. Celonis Snap und LANA Process Mining sind cloudbasiert. Bei Disco und ProM handelt es sich um lokal zu installierende Applikationen. PAFnow wiederum zeichnet sich durch Implementierung in die Plattform Microsoft Power BI aus. Alle Tools bieten durchwegs eine wohlüberlegte Oberfläche zur Festlegung der Attribute zum Import des Event Logs. Die Tools Disco, LANA Process Mining und PAFnow schlagen selbständig eine Zuordnung der Attribute vor, dabei sind Änderungen problemlos möglich. Bei Celonis Snap ist bei einem Import einer CSV-Datei zunächst der entsprechende Feld-Separator anzugeben und auch die Zeichenformate pro Attribut/Spalte festzulegen. In der ProM-Plattform ist ebenfalls ein Feld-Separator anzugeben und die Attribute sind auszuwählen. Nach Auswahl der Start- und Ende-Zeitstempel ist das Tool in der Lage, das Datumsformat automatisch zu erkennen.

Die kostenfreien Versionen der kommerziellen Anbieter von Process Mining-Software bieten deutlich weniger Funktionen als es bei ProM der Fall ist, jedoch ist deren Anwendung viel intuitiver und verständlicher als die überwältigende Möglichkeiten bietende ProM-Plattform, welche für Laien anfangs schwierig zu erschließen ist. Bei Anwendung von ProM wird mehr

Zeit zur Einarbeitung benötigt. Ähnlich wird es auch in [Aal16, S. 325] interpretiert. Kommt Process Mining in Organisationen infrage, sollte auf die Art der Benutzergruppen Rücksicht genommen werden. Zum einen gibt es Benutzer, die zu einmaligen Fragestellungen mithilfe der Software eigenständige Lösungen entwickeln wollen. Zum anderen Benutzer, die bereits möglichst viele Anwendungsmöglichkeiten und vordefinierte, standardmäßige Use-Cases bereitgestellt haben möchten. In [Aal16, S. 328ff.] werden drei differente Benutzertypen vorgestellt sowie Ausprägungsmöglichkeiten der Tools. Vor der Entscheidung für einen Anbieter und dessen Tool sollte genau analysiert werden, für welche Zwecke und Benutzergruppen die Process Mining-Software im Unternehmen genutzt werden soll. Vordefinierte Use-Cases sind bei den kommerziellen Anbietern durchwegs verfügbar. Es ist festzustellen, dass alle vorgestellten Tools ihre generierten Prozessgraphen mit visuellen Hilfestellungen versehen. Dies erleichtert die Arbeit der Prozessanalysten deutlich, wenn im Prozess auftretende Flaschenhals-Aktivitäten direkt von anderen Aktivitäten farblich hervorgehoben werden. Zeit- und Performance-Informationen werden stets für alle Aktivitäten und Verbindungen angegeben. Ebenso sind die Schwellenwert-Regler und Zoomfunktionen durchwegs implementiert. In Sachen Performance wurden aufgrund des sehr kleinen Datensatzes keine nennenswerten Unterschiede zwischen den Tools festgestellt.

Unterschiede ergeben sich vor allem nach der Mining-Phase. Die Analyse-Phase wird häufig durch vordefinierte Dashboards mit benötigten Kennzahlen unterstützt. Durch die Implementierung in Microsoft Power BI bietet PAFnow die größten Möglichkeiten der Visualisierung mit großen Freiheiten der benutzerspezifischen Darstellung von Dashboards. Viele eigene Ideen sind einbringbar. Jedoch sind dafür zwingend auch grundlegende Kenntnisse in Microsoft Power BI erforderlich. Ein Blick auf die Webseite von PAF zeigt, welche umfangreiche Visualisierungen und Dashboards möglich sind. Diese Freiheiten sind bei den anderen Tools nur eingeschränkt gegeben. Zur Analysenerstellung wird bei Celonis Snap in den Edit-Mode gewechselt. Dabei stehen viele Komponenten zur Prozessanalyse, Tabellen und Diagramme, KPI-Komponenten und Design-Komponenten zur Verfügung. LANA Process Mining überzeugt zum einen mit vorgefertigten Statistiken, zum anderen aber auch mit selbst zu erstellenden Root Cause-Analysen und Dashboards. Diese werden in vier Kategorien eingeteilt, darunter Case-Analysen und Aggregationen. Das Disco-Tool von Fluxicon bietet lediglich vorgefertigte Darstellungen, welche von Statistiken geprägt sind. Als weniger praktikabel erweist sich in der Analyse-Phase das Open-Source-Tool ProM, da Darstellungen wie Dashboards mit KPIs nicht ohne Weiteres erstellt werden können.

Bei Anwendung der Tools kann es an unterschiedlichen Stellen vorkommen, dass der Endanwender Hilfe benötigt beziehungsweise die eigenen Lösungsansätze ausgereizt sind. Im Rahmen der Usability sollten hierfür in Softwares entsprechende Hilfestellungen gegeben sein.

Celonis Snap bietet neben einem Handbuch auch umfangreiche Tutorials für die eigene Analysenerstellung an. In Disco werden Informationen durch virtuelle Post-it-Notizen bereitgestellt. LANA Process Mining punktet mit direkter Interaktion über eine Chatfunktion. Bei PAFnow sind in Microsoft Power BI umfangreiche Dokumentationen, Tutorials, Schulungsvideos und vieles mehr verfügbar – jedoch lediglich allgemein zum Microsoft Power BI und nicht für das Custom Visual von PAFnow Process Mining. In ProM wird keine Hilfestellung gegeben. Es können jedoch freie Onlinekurse über die Webseite belegt werden. Auch das Forum bietet Möglichkeiten zum Austausch über das Tool.

Dem Gartner Market Guide for Process Mining [Ker19] zufolge, ist bei den kommerziellen Versionen der Process Mining-Anbieter eine Diskrepanz betreffend des Funktionsumfangs zu erkennen. Wie umfangreich die jeweiligen kommerziellen Lizenzen sind, wird im Rahmen dieser Arbeit nicht behandelt. Die kostenfreien Versionen bieten wie in den Unterkapiteln von Abschnitt 6.2 beschrieben, einen bezüglich der Grundfunktionen ziemlich ähnlichen Umfang an Möglichkeiten an.

7 Zusammenfassung und Fazit

Die vorgestellte, toolbasierte Methode des Process Mining erfreut sich steigender Akzeptanz. Die Thematik des Prozessbewusstseins und der Stellenwert von Unternehmensdaten findet in Führungsetagen von Organisationen durch die digitale Transformation von Geschäftsvorgängen großen Anklang. Hierbei gilt es, die Vorteile einer hohen Datenqualität zu kommunizieren und sowohl das Bewusstsein als auch die Akzeptanz auf breiter Basis zu verankern.

Ziel dieser Bachelorarbeit war es, die Herausforderungen bei der Anwendung und Entwicklung der toolbasierten Methodik Process Mining auszuarbeiten. Dabei gilt es zunächst die notwendigen Voraussetzungen in der Organisation zu schaffen. Betrachtet wurden insbesondere das Prozessbewusstsein und die Datenqualität. Die Wichtigkeit beider Aspekte sollte den Organisationen, welche sich mit der Nutzung von Process Mining auseinandersetzen, klar sein. Kommt die Nutzung infrage, sollte zielgerichtet vorgegangen werden und das Vorhaben initial einem Projektcharakter entsprechen und fortlaufend als iterativer Prozess betrachtet werden. Dafür wird in dieser Arbeit ein repräsentatives methodisches Vorgehen als Grundlage verwendet. Die einzelnen Schritte weisen wiederum gewisse Fallstricke auf, welche beachtet werden sollten. In Sachen Process Mining-Algorithmen wird in den nächsten Jahren viel Bewegung erwartet, zumal stets Neue entwickelt oder Bestehende verbessert werden. Die Leistungsfähigkeit und Skalierbarkeit stehen hierbei im besonderen Fokus. Die Anwendung von Machine-Learning und KI kann sich als stark unterstützender Faktor darstellen und bietet großes Potential. Es wurde eruiert, welche betrieblichen Vorgänge zwar dem Prozessmanagement zuzuordnen sind, jedoch durch das Vorliegen eines Process Mining-Graphen nicht vollumfänglich beantwortet werden. Herausgestellt hat sich dabei, dass sich eine Durchführung von Prüfschritten der transaktionsbasierten Datenanalyse im Nachgang zum Process Mining zur weiterführenden Analyse der eigenen Prozesse empfiehlt. In der heutigen Zeit, in der sich Organisationen aufgrund datenintensiver Prozesse vermehrt mit personenbezogenen und schützenswerten Daten auseinandersetzen, spielt der Datenschutz eine große Rolle. Seit Einführung der DSGVO hat sich die Rechtslage verschärft. Dies bedingt eine noch größere Beachtung dieser Thematik. Im Rahmen des Process Mining wird von sekundärer Nutzung von Daten gesprochen, was bei datenschutzkritischen Prozessanalysen einer vorherigen Verschlüsselung bedarf. Dabei ist es jedoch wichtig, dass weiterhin die bestmöglichen Process Mining-Ergebnisse erzielt werden können.

Ein weiteres Ziel der Arbeit bestand in der Herausarbeitung möglicher Potentiale dieser toolbasierten Methode. Vor allem die Vorteilhaftigkeit als Vorstufe zu innovativen Technologien wie RPA wird hervorgehoben. Organisationen sind erpicht, gewisse Prozesse zu standardisieren und zu automatisieren. Process Mining wird als Wegbereiter für RPA-Initiativen verstanden, mit dem das Automatisierungspotential von Vorgängen bewertet werden kann. Stärker in

den Fokus der Process Mining-Anbieter rückt mittlerweile auch die Technologie DTO. Die Anbieter loten aus, inwiefern Process Mining für die Simulation von Geschäftsprozessen eingesetzt werden kann. Durch stetig besser werdende Rechnerleistungen steigt auch die Nachfrage nach der Nutzung von Echtzeitinformationen, was die Verwendung von Process Mining zur operativen Unterstützung befürwortet. So kann dieses Tool zum täglichen Begleiter für Prozessmanager werden, die durch Warnmeldungen direkt in noch laufende Prozesse eingreifen können. Aufgrund steigender Verflechtung existierender Unternehmensnetzwerke kann Process Mining auch entlang der Wertschöpfungskette organisationsübergreifend eingesetzt werden. Das Optimierungspotential ist dadurch höher und Zeit- und Kostenvorteile können generiert werden. Das kann bestenfalls zu signifikanten Wettbewerbsvorteilen im kompetitiven Globalisierungsmarkt führen.

Für Process Mining wird ein sehr vielversprechender Zukunftsmarkt mit hohen Wachstumsraten prognostiziert. Das kann zu schnell wachsenden Unternehmen führen, wie es beim Marktführer Celonis der Fall ist. Im Rahmen dieser Ausarbeitung wurden exemplarisch fünf führende Anbieter kurz vorgestellt und der Funktionsumfang von deren kostenfreien Versionen beschrieben. Dabei nimmt das Open-Source-Tool ProM eine gesonderte Rolle ein, da alle Funktionen vollumfänglich verfügbar sind und es sich vor allem für akademische Zwecke anbietet und als Plattform für Entwickler dient. Die kommerziellen Tool-Anbieter gleichen sich an bei ihren Testversionen. Sie weisen durchgehend eine größere Bedienerfreundlichkeit auf und sind intuitiver handzuhaben als ProM. Dieser Umstand macht es vor allem für die Benutzergruppe interessant, welche die vordefinierten Use-Cases der Hersteller nutzen möchten. Unterschiede im Funktionsumfang werden vor allem in der Analyse-Phase innerhalb der Tools deutlich. Dort konnten bei den Tools PAFnow, Celonis Snap und LANA Process Mining die größten Funktionsmöglichkeiten verzeichnet werden.

Ein unternehmensweit gelebtes, gemeinsames Geschäftsprozessverständnis und deren Transparenz können positive Auswirkungen auf Beschäftigte, Lieferanten und Kunden haben. Durchdachte und effiziente Vorgehensweisen und Prozesse sprechen für die hohe Qualität einer Organisation, in der die Beschäftigten gerne arbeiten und Lieferanten sowie Kunden einen Mehrwert durch kürzer gestaltete Bearbeitungs- und Lieferzeiten erlangen. Die in dieser Arbeit aufgeführten Punkte begründen zweifellos die Daseinsberechtigung von Process Mining zur Entdeckung, Prüfung und Verbesserung der eigenen Geschäftsprozesse. Mit Process Mining wird ein effizientes und transparentes Geschäftsprozessmanagement ermöglicht.

Nichtsdestotrotz müssen Anbieter und Anwender die bestehenden Herausforderungen intensiv angehen sowie die Potentiale des Tools aktiv vorantreiben, um den größtmöglichen Nutzen daraus zu generieren zu können.

Anhang

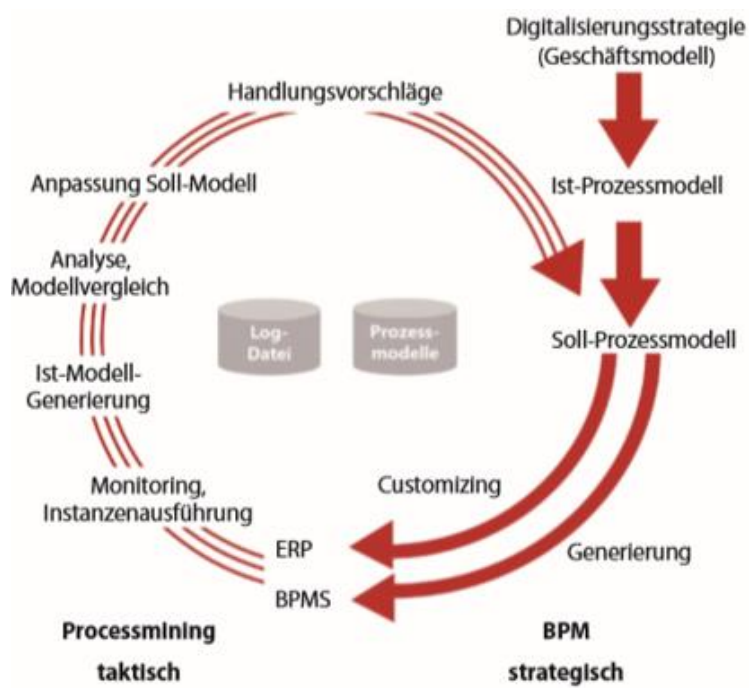


Abbildung 21: Process Mining in BPM [Sch18a]

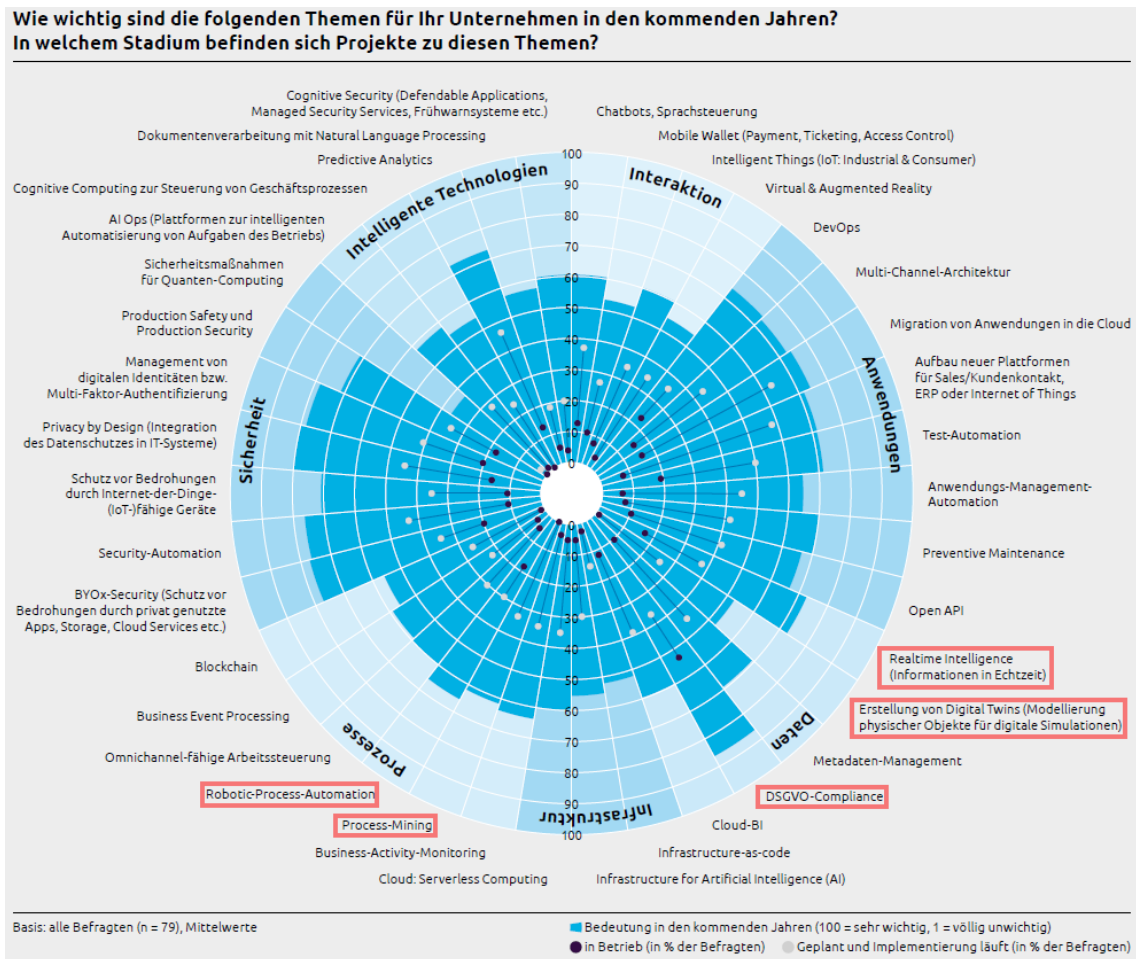


Abbildung 23: Zukunftsthemen im IT-Umfeld [Cap19]

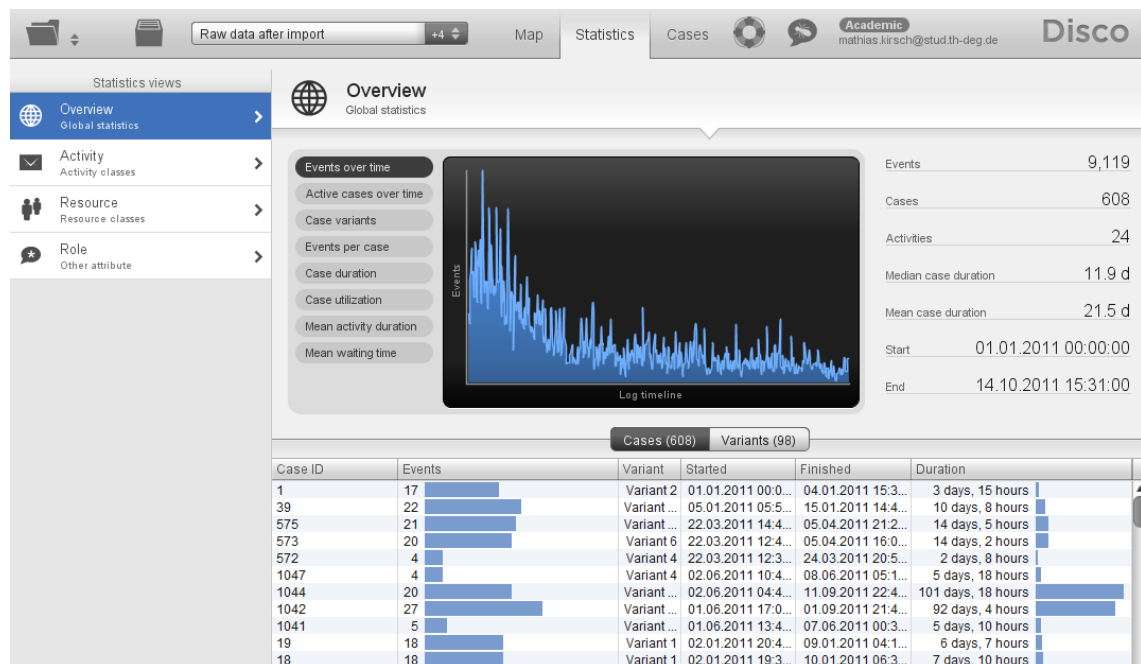


Abbildung 24: Reiter "Statistics" des Disco-Tools

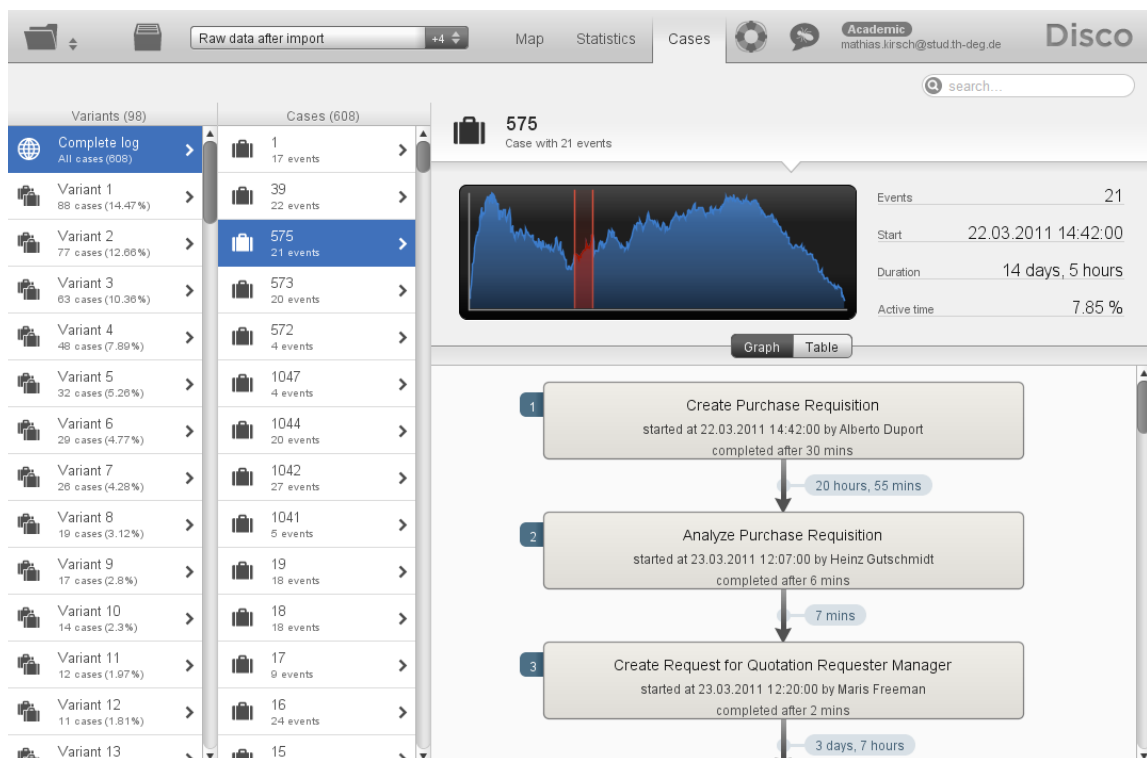


Abbildung 25: Reiter "Cases" des Disco-Tools

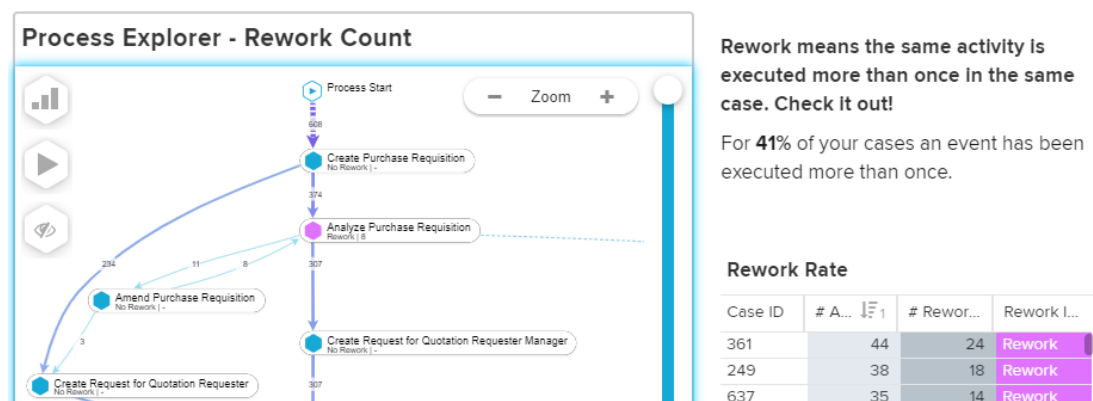


Abbildung 26: Reiter „Rework Rate“ von Celonis Process Mining

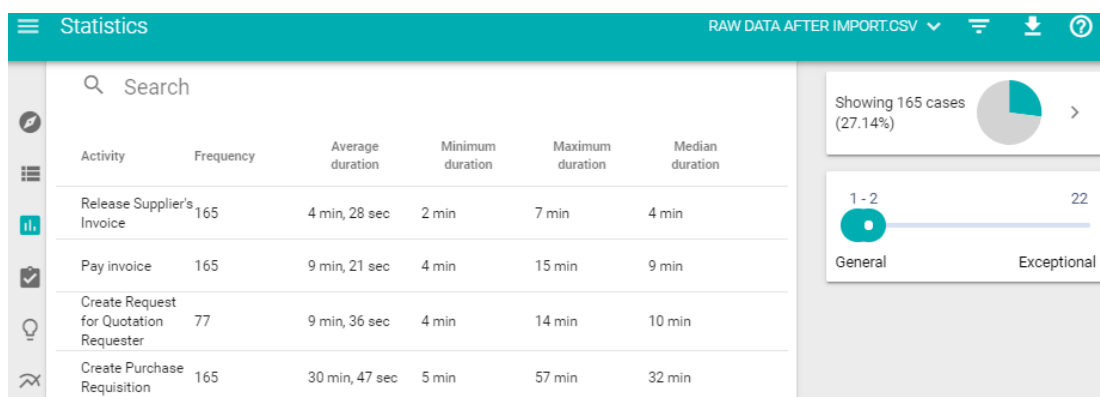


Abbildung 27: Reiter "Statistics" von LANA Process Mining

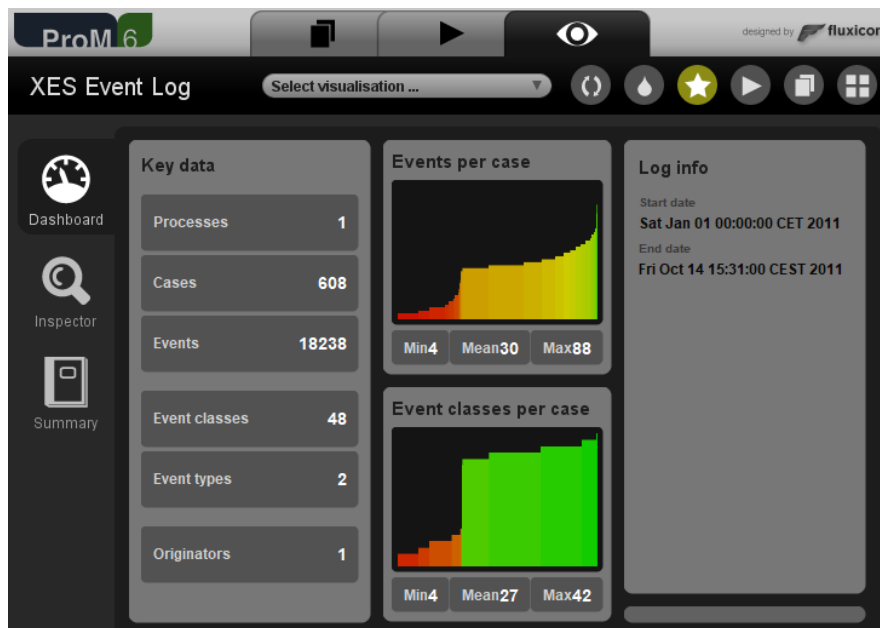


Abbildung 28: Reiter "View" von ProM

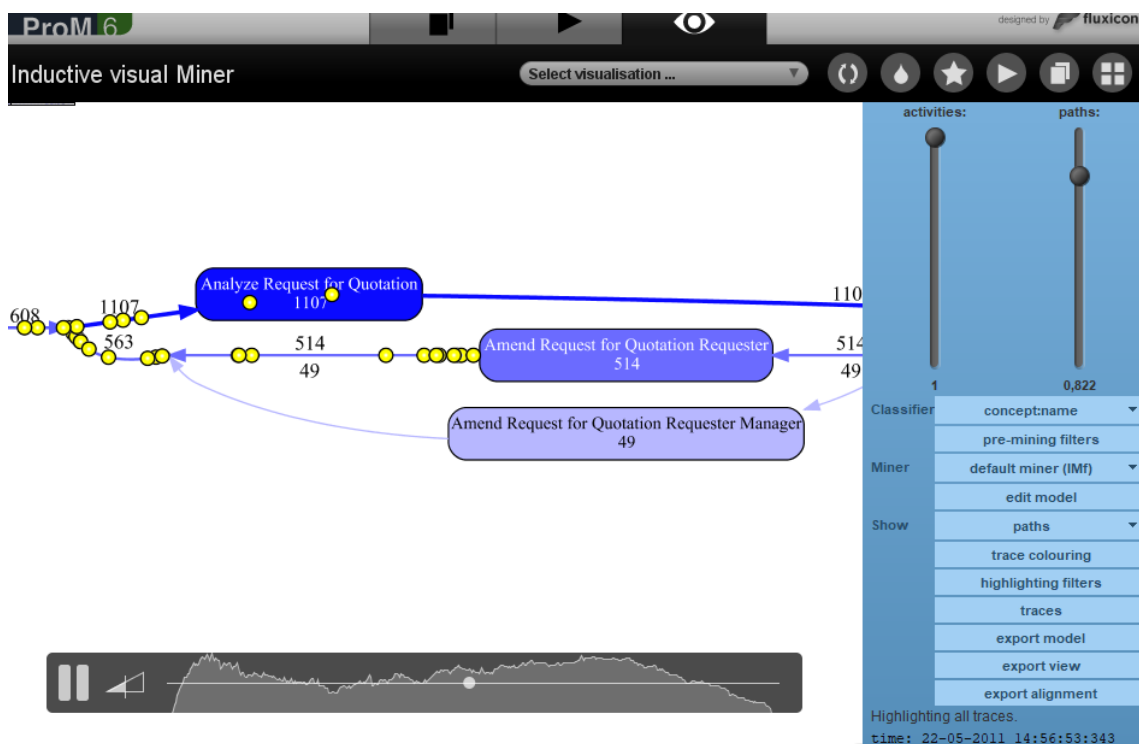


Abbildung 29: Inductive visual Miner in ProM

Literaturverzeichnis

- [Aal11] Aalst van der W.M.P.: Process Mining: Discovery, Conformance and Enhancement of Business Processes, Springer-Verlag, Berlin, Heidelberg, 2011
- [Aal15] Aalst van der W.M.P., Damiani E.: Processes Meet Big Data: Connecting Data Science with Process Science: IEEE Transactions on Services Computing, Volume: 8, Issue: 6 S. 810 - 819, 2015
<https://ieeexplore.ieee.org/document/7302592>, Aufruf: 14.08.2019
- [Aal16] Aalst van der W.M.P.: Process Mining: Data Science in Action, Springer-Verlag, Heidelberg, 2016
- [Aal19] Aalst van der W.M.P.: Alternative Process Discovery Techniques, 2019
<https://de.coursera.org/lecture/process-mining/4-2-alternative-process-discovery-techniques-4qURZ>, Aufruf: 21.07.2019
- [Bay19] Bayer M.: Process Mining verspricht Durchblick im Prozessdschungel: Prozesse visualisieren, analysieren, optimieren
<https://www.computerwoche.de/a/process-mining-verspricht-durchblick-im-prozessdschungel,3546949>, Aufruf: 21.07.2019
- [Beh16] Beheshti S.M.R., Benatallah B., Sakr S., Grigori D., Motahari-Nezhad H.R., Barukh M.C., Gater A., Ryu S.H.: Process analytics: Concepts and Techniques for Querying and Analyzing Process Data, Springer-Verlag Berlin Heidelberg, New York, 2016
- [Ben19] Bendel O.: Einhorn: Gabler Wirtschaftslexikon
<https://wirtschaftslexikon.gabler.de/definition/einhorn-119178/version-368879>, Aufruf: 05.08.2019
- [Bet16] Bettacchi A., Re B.: Understanding Production Chain Business Process Using Process Mining: A Case Study in the Manufacturing Scenario, International Conference on Advanced Information Systems Engineering, 2016
https://www.researchgate.net/publication/303823045_Understanding_Production_Chain_Business_Process_Using_Process_Mining_A_Case_Study_in_the_Manufacturing_Scenario, Aufruf: 26.08.2019
- [Bod06] Bodendorf F.: Daten- und Wissensmanagement, Springer-Verlag Berlin Heidelberg, Berlin, Heidelberg, 2006
- [Bur15] Burattin A.: Process Mining Techniques in Business Environments: Theoretical Aspects, Algorithms, Techniques and Open Challenges in Process Mining, Springer-Verlag, Cham, 2015
- [Cap19] Capgemini: Intelligente Technologien - Vorreiter erzielen bereits Ergebnisse, Studie IT-Trends 2019
<https://www.capgemini.com/de-de/resources/studie-it-trends-2019/> Aufruf: 21.07.2019
- [DDW15] DIE DEUTSCHE WIRTSCHAFT: Die 50 am schnellsten wachsenden Technologieunternehmen Deutschlands
<https://die-deutsche-wirtschaft.de/die-50-am-schnellsten-wachsenden-technologieunternehmen-deutschlands/>, Aufruf: 20.07.2019

- [Dud19] Duden.de: Daten
<https://www.duden.de/rechtschreibung/Daten>, Aufruf: 11.08.2018
- [Eck15] Eck van M.L., Lu X., Leemans S.J.J., Aalst van der W.M.P.: PM²: A Process Mining Project Methodology: Advanced Information Systems Engineering Workshops in CAiSE 2015 International Workshops, S. 297-313, Lecture Notes in Computer Science book series (LNCS, volume 9097), Springer-Verlag, Cham, 2015
- [Fah19] Fahrenkrog-Petersen S.A.: Providing Privacy Guarantees in Process Mining, International Conference on Advanced Information Systems Engineering - Doctoral Consortium, 2019
https://www.researchgate.net/publication/333080170_Providing_Privacy_Guarantees_in_Process_Mining, Aufruf: 14.08.2019
- [Fra18] Fraunhofer-Gesellschaft: Der Data Scientist – dank Big Data heiß begehrt!
<https://www.fraunhofer.de/de/presse/presseinformationen/2018/juli/der-data-scientist-dank-big-data-heiss-begehrt.html>, Aufruf: 12.08.2019
- [Gar19] Garcia C.d.S., Meincheim A., Faria Junior E.R., Dallagassa M.R., Sato D.M.V., Carvalho D.R., Santos, E.A.P., Scalabrin E.E.: Process Mining Techniques and Applications – A Systematic Mapping Study, Expert Systems With Applications, Volume 133, S. 260–295, 2019
<https://www.sciencedirect.com/science/article/pii/S0957417419303161?via%3Dihub>, Aufruf: 18.08.2019
- [Gey18] Geyer-Klingenberg J., Nakladal J., Baldauf F.: Process Mining and Robotic Process Automation: A Perfect Match, 16th International Conference on Business Process Management (BPM), 2018
https://www.researchgate.net/publication/326466901_Process_Mining_and_Robotic_Process_Automation_A_Perfect_Match, Aufruf: 14.08.2019
- [How19] Howson C., Richardson J., Sallam R., Kronz A.: Gartner Magic Quadrant for Analytics and Business Intelligence Platforms, Gartner Inc.
<https://www.gartner.com/doc/reprints?id=1-3TXXSLV&ct=170221&st=sb>, Aufruf: 24.08.2019
- [IDC18] IDC - International Data Corporation: Prognose zum Volumen der jährlich generierten digitalen Datenmenge weltweit in den Jahren 2018 und 2025 (in Zettabyte), Studie in Statista, 2018
<https://de.statista.com/statistik/daten/studie/267974/umfrage/prognose-zum-weltweit-generierten-datenvolumen/>, Aufruf: 02.07.2019
- [IEE12] IEEE CIS Task Force on Process Mining: Aalst van der W.M.P., Adriansyah A., Alves de Medeiros A.K., Arcieri F., Baier T., Blickle T., Bose J.C., Brand van den P., Brandtjen R., Buijs J., Burattin A., Carmona J., Castellanos M., Claes J., Cook J., Costantini N., Curbera F., Damiani E., Leoni de M., Delias P., Dongen van B., Dumas M., Dustdar S., Fahland D., Ferreira D.R., Gaaloul W., Geffen van F., Goel S., Günther C., Guzzo A., Harmon P., Hofstede ter A., Hoogland J., Ingvaldsen J.E., Kato K., Kuhn R., Kumar A., La Rosa M., Maggi F., Malerba D., Mans R., Manuel A., McCreesh M., Mello P., Mendling J., Montali M., Nezhad H.M., Muehlen zur M., Munoz-Gama J., Pontieri L., Ribeiro J., Rozinat A., Pérez H.S., Pérez R.S., Sepúlveda M., Sinur J., Soffer P., Song M., Sperduti A., Stilo G., Stoel C., Swenson K., Talamo M., Tan W., Turner C., Vanthienen J., Varvaressos G., Verbeek E., Verdonk M., Vigo R.,

- Wang J., Weber B., Weidlich M., Weijters T., Wen L., Westergaard M., Wynn M.: Process Mining Manifest: Business Process Management Workshops, BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, revised selected papers, Springer-Verlag, Berlin, 2012
https://www.win.tue.nl/ieeetfpm/doku.php?id=shared:process_mining_manifesto, Aufruf: 14.08.2019
- [Joe18] Jörgensen H.: Wie Unternehmen mehr und bessere Daten gewinnen, <https://www.capital.de/wirtschaft-politik/wie-unternehmen-mehr-und-bessere-daten-gewinnen>, 2018, Aufruf: 14.08.2019
- [Ker19] Kerremans M.: Gartner Market Guide for Process Mining, Gartner Inc.
- [Lau16] Laudon K.C., Laudon J.P., Schoder D.: Wirtschaftsinformatik: Eine Einführung, 3. vollständig überarbeitete Auflage, Pearson Verlag, Hallbergmoos, 2016
- [Lee16] Leemans S.J.J., Fahland D., Aalst van der W.M.P.: Scalable process discovery and conformance checking: Software & Systems Modeling Volume 17 Issue 2 S. 599-631, 2016
<https://link.springer.com/article/10.1007%2Fs10270-016-0545-x>, Aufruf: 14.08.2019
- [Len18] Lenz R., Druckenmüller J., Nominacher B.: Process Mining in der Revision: Ein Erfahrungsbericht, Zeitschrift Interne Revision ZIR.04.2018.202, 2018
- [Nom18] Nominacher B.: Motivationsfaktor Process Mining - Wie Effizienz das Miteinander verbessert
<https://www.computerwoche.de/a/motivationsfaktor-process-mining,3544808>, Aufruf: 10.07.2019
- [RAM18] Research and Markets: Process Analytics Market by Process Mining Type (Process Discovery, Process Conformance & Process Enhancement), Deployment Type, Organization Size, Application (Business Process, It Process, & Customer Interaction) & Region - Global Forecast to 2023
<https://www.researchandmarkets.com/reports/4576970/process-analytics-market-by-process-mining-type>, Aufruf: 20.07.2019
- [Raf18] Rafiei M., Waldthausen von L., Aalst van der W.M.P.: Ensuring Confidentiality in Process Mining
https://www.researchgate.net/publication/330042256_Ensuring_Confidentiality_in_Process_Mining, Aufruf: 15.08.2019
- [Rob18] Robledo P.: Process Mining plays an essential role in Digital Transformation
<https://medium.com/@pedrorobledobpm/process-mining-plays-an-essential-role-in-digital-transformation-384839236bbe>, Aufruf: 05.07.2019
- [Roz07] Rozinat A., Alves de Medeiros A.K., Günther C.W., Weijters A.J.M.M., Aalst van der W.M.P.: Towards an Evaluation Framework for Process Mining Algorithms
<https://pdfs.semanticscholar.org/d690/739283b54929cb9b536387e5775a8cd41d54.pdf>, Aufruf: 15.08.2019
- [Sch18a] Scheer A.W.: Den Daten der Geschäftsprozesse auf der Spur: Manager Wissen in Harvard Business Manager: Ausgabe Oktober 2018

auch online: http://manager-wissen.com/wp-content/uploads/2018/09/manager-wissen_im_harvard-business-manager_Oktober2018.pdf, Aufruf: 15.08.2019

- [Sch18b] Schunselaar D.M.M., Slaats T., Maggi F.M., Reijers H.A., Aalst van der W.M.P.: Mining Hybrid Business Process Models: A Quest for Better Precision https://www.researchgate.net/publication/325784858_Mining_Hybrid_Business_Process_Models_A_Quest_for_Better_Precision, Aufruf: 15.08.2019
- [Sie18] Siepermann M., Lackes R.: Syntax einer Programmiersprache: Gabler Wirtschaftslexikon <https://wirtschaftslexikon.gabler.de/definition/syntax-einer-programmiersprache-47018/version-270289> <https://wirtschaftslexikon.gabler.de/definition/semantik-einer-programmiersprache-42816/version-266158> Stand: 15.08.2019
- [Wan13] Wang J., Wong R.K., Ding J., Guo Q., Wen L.: Efficient selection of process mining algorithms, IEEE Transactions on Services Computing, Volume 6, Issue 4, S. 484-496, 2013 https://www.researchgate.net/publication/260710998_Efficient_Selection_of_Process_Mining_Algorithms, Aufruf: 18.08.2019
- [Wer11] Werner M., Gehrke N.: Potentiale und Grenzen automatisierter Prozessprüfungen durch Prozessrekonstruktionen https://www.researchgate.net/publication/266078219_Potentiale_und_Grenzen_automatisierter_Prozesspruefungen_durch_Prozessrekonstruktionen, Aufruf 20.07.2019
- [Wer12] Werner M.: Einsatzmöglichkeiten von Process Mining für die Analyse von Geschäftsprozessen im Rahmen der Jahresabschlussprüfung https://www.researchgate.net/publication/266078166_Einsatzmoeglichkeiten_von_Process_Mining_fur_die_Analyse_von_Geschäftsprozessen_im_Rahmen_der_Jahresabschlussprüfung, Aufruf: 20.07.2019
- [WIW19] Wirtschaftswoche: Diese deutschen Unternehmen sind mindestens eine Milliarde Dollar wert: Deutsche Unicorns <https://www.wiwo.de/unternehmen/handel/deutsche-unicorns-diese-deutschen-unternehmen-sind-mindestens-eine-milliarde-dollar-wert/23904116.html>, Aufruf: 20.07.2019
- [Zwe18] Zweig K.A., Fischer, S., Lischka K.: Wo Maschinen irren können – Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung: Bertelsmann Stiftung, Gütersloh, 2018
auch online: <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/wo-maschinen-irren-koennen/>, Aufruf: 15.08.2019

