Digital Audit with Process Mining

Weaknesses of current process mining techniques as an instrument for the data-driven audit of financial statements

Dissertation

zur Erlangung des Doktorgrades
Dr. rer. pol.

der Fakultät für Wirtschaftswissenschaften
der Universität Duisburg-Essen

vorgelegt
von

Annalouise Maas
aus
Hagen

Essen, August 2023

Erstgutachter:	Prof. Dr. Ludwig Mochty
Zweitgutachter:	Prof. Dr. Stefan Eicker
Dritter Prüfer:	Prof. Dr. Erwin Amann
Datum der Disputation:	29. Februar 2024

DOI: 10.17185/duepublico/81667
URN: urn:nbn:de:hbz:465-20240305-065407-5

Table of contents

List of abbreviations

ACFE	Association of Certified Fraud Examiners
ACM	Association for Computing Machinery
AG	Aktiengesellschaft (Public Limited Company)
AICPA	American Institute of Certified Public Accountants
AIS	Accounting Information System
AS	Auditing Standard
BaFin	Bundesanstalt für Finanzdienstleistungsaufsicht (Federal Financial Supervisory Authority)
BMF	Bundesfinanzministerium (Federal Ministry of Finance)
BMJ	Bundesministerium der Justiz (Federal Ministry of Justice)
BPM	Business Process Management
CAATs	Computer Assisted Audit Techniques
CAD	Cash Against Documents
CICA	Canadian Institute of Chartered Accountants
COSO	Committee of Sponsoring Organizations of the Treadway Commission
CPA	Certified Public Accountant
CpD	Conto pro Diverse
CRM	Customer Relationship Management
CY	Current Year
DCGK	Deutscher Corporate Governance Kodex (German Corporate Governance Code)
ERP	Enterprise Resource Planning
EUR	Euro
EY	Ernst & Young GmbH, Wirtschaftsprüfungsgesellschaft
FI	Finance Invoice (SAP FI module)

FISG	Finanzmarktintegritätsstärkungsgesetz (Financial Market Integrity Strengthening Act)
GAAP	Generally Accepted Accounting Principles
GR	Goods Receipt
GR/IR	Goods Received/Invoice Received
IAASB	International Auditing & Assurance Standards Board
IC	Intercompany
IEEE	Institute of Electrical and Electronic Engineers
IFAC	International Federation of Accountants
INV	Invoice
ISA	International Standard on Auditing
ISRN	International Standard Technical Report Number
IT	Information Technology
JE	Journal Entry
JPY	Japanese Yen
KPI	Key Performance Indicator
KPMP	Klynveld Peat Marwick Goerdeler AG, Wirtschaftsprüfungsgesellschaft
MM	Material Management (SAP MM module)
Para.	Paragraph
PCAOB	Public Company Accounting Oversight Board
PO	Purchase Order
PwC	PricewaterhouseCoopers AG, Wirtschaftsprüfungsgesellschaft
PY	Prior Year
SAP	Systeme, Anwendungen und Produkte (SAP AG)
SAS	Statement on Auditing Standard
SCOT	Significant Class of Transactions

SLA	Service Level Agreement
SOX	Sarbanes-Oxley-Act
SQL	Structured Query Language
SRM	Supplier Relationship Management (SAP SRM module)
SSC	Shared Service Center
VAT	Value-Added Tax

List of tables

List of figures

1 Introduction

1.1 Problem definition

The objective of an audit of financial statements is to obtain reasonable assurance that the financial statements as a whole are in line with regulatory requirements, in particular free from material misstatement.[1] Over decades, this audit conclusion has been achieved predominantly based on samples by applying the audit risk model,[2] including samples with regard to the execution of internal controls over financial reporting and samples with regard to substantive procedures. The scope of applying analytical procedures was limited to specific questions and separated areas of the financial statements, primarily in the phase of audit planning rather than the audit execution.

Over the last couple of years, audit procedures summarized under the concept “data analytics” receive increasing attention by audit theory and audit practice. To some extent, “data analytics” include methods that have already been previously suggested with the concept of “computer assisted audit techniques” (CAATs)[3], but have not been accepted due to the lack of data availability, high-performing computer systems that support evaluation as well as psychological resistance of the audit practice.[4] The novel aspect of data analytics is that in contrast to the traditional sample-based audit, the concept presumes an approximate full audit, or at least achieving the audit opinion based on the total population of transactions.[5] This presumption is made because today, the accounting relevant data is predominantly available in electronic format and by this accessible for a holistic analysis instead of a sample-based inspection only. Recently, the revised version of International Standard on Auditing (ISA) 315 introduced the concept “automated tools and techniques”[6] to particularly address the psychological resistances and enable a broad adoption by the audit practice. However, apart from the guidance for applying analytical procedures in ISA 520, a methodical integration of the concepts “data analytics”, “automated tools and techniques” and “CAATs” is missing.

---

1 Cf. IFAC (2021), ISA 200, para. 3 and 11(a).
2 Cf. ibid., para. 5 and 17.
3 Cf. CICA (2003), pp. 77ff.
4 Cf. BIERSTAKER, JAMES/JANVRIN, DIANE/LOWE, D. JORDAN (2014), pp. 5ff.; DEBRECENY, ROGER et al. (2005), pp. 8ff.; BRAUN, ROBERT L./DAVIS, HAROLD E. (2003), p. 725.
5 Cf. IFAC (2021), ISA 315 (Revised 2019), para. A21; IFAC (2021), ISA 330, para. A16.
6 Cf. IFAC (2021), ISA 315 (Revised 2019), para. A21.

A study of the FINANCIAL REPORTING COUNCIL shows that despite the wide availability of data, the analysis of mass data in an audit of financial statements is widely limited to journal entry testing using general ledger data or partly automated analytical procedures using subledger data, addressing management's assertions for individual account balances.[7] To a large extent, the audit of internal controls is performed manually by inquiries and observations (design effectiveness of controls according to ISA 315) and tests of samples (operating effectiveness of controls according to ISA 330). However, the increasing use of information technology (IT) and increasing controlling and documentation requirements steadily increased the complexity of business processes in the past.[8] An increasing process complexity leads to an increased susceptibility to errors of a manual audit of internal processes and controls. Over the past years, a large portion of the findings of audit firms' internal and external professional inspections is related to an insufficient or defective audit of the entity's system of internal control.[9] Next to these findings are the auditees' expectations of a digital audit, that particularly include information about the control safety of business processes. Often, business processes are historically grown and have to be continuously adjusted to changing economic, regulatory, technological and organizational conditions. As a consequence, there are often significant differences between the process design and the actual execution of a process.[10]

In the recent past, methods and techniques subsumed under the term "process mining" receive increasing attention by audit theory and audit practice.[11] With process mining it should be possible to transfer the idea of data analytics to the audit of internal processes and integrated controls.[12] The technology uses data stored in the auditees' IT systems to identify and analyze the entities' accounting relevant processes and controls. The methodical integration and conformance with national and international auditing standards has been unclear as of now.[13] Acceptance problems are also foreseeable with regard to the practical application of process mining in the audit. Since many years, the audit of internal controls is performed almost unchanged from a methodical

[7] Cf. FINANCIAL REPORTING COUNCIL (2017), pp. 12f.
[8] Cf. WIESE, MICHAEL (2013), p. 1; MOCHTY, LUDWIG/WIESE, MICHAEL (2012), pp. 487f.
[9] Cf. PCAOB (2018), pp. 7ff.; PCAOB (2016), p. 10.
[10] Cf. BODENMANN, ANDREAS/MAAS, ANNALOUISE/SCHWAN, MORITZ (2018), p. 41.
[11] Cf. DELOITTE (2020); EY (2019); PWC (2019); KPMG (2018).
[12] Cf. VAN DER AALST, WIL M. P. et al. (2010), p. 90.
[13] Cf. WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021), p. 2.

perspective. At the same time, the concept of process mining reveals several weaknesses of the audit risk model[14] introduced 1983.[15] Regardless of that, process mining is continuously propagandized as an audit instrument especially by the Big Four audit firms.[16]

Early 2017, one of the Big Four audit firms[17] has started an extensive piloting of process mining to establish the technology in the audit practice and investigate related questions empirically based on concrete examples. In 2020, the firm's process mining application for the purchase to pay process has been globally certified, that is, approved for the worldwide application in the firm's audits of financial statements, provided that certain conditions are met. As part of the research, development and piloting processes, several expected but also surprising methodical, technical and regulatory questions emerged.

This thesis includes a systematic identification of the methodical requirements of a data-based process analysis in an audit of financial statements. However, evaluating the ability of process mining to meet these requirements as well as identifying and evaluating upcoming questions and practical challenges cannot be theoretical tasks only. Therefore, the thesis includes an empirical part that describes and evaluates the piloting of process mining in the audit of numerous entities of varying sizes and industries across Europe, Middle East, Africa and Japan. Subsequently, the findings obtained and challenges and points of criticisms identified with regard to the current process mining are critically analyzed and evaluated. Finally, propositions for dealing with the identified challenges are developed.

The objective of the thesis is to highlight key points of criticism and acceptance challenges based on empirical findings that conflict a broader dissemination of process mining as an instrument of the financial statement audit. By evaluating the weaknesses identified and developing solution approaches, the thesis aims to contribute to establishing process mining as one of the key technologies of the audit of the future.

14 Cf. AICPA (1983), SAS No. 47.
15 The traditional audit procedures performed for assessing control risk do not support a robust assessment of the reliability of the system of internal control – because this is not their purpose. Tests of controls are designed to reduce the auditor's detection risk (cf. IFAC (2021), ISA 315 (Revised 2019), para. 34; Chapter 2.3.1) and by this the extent of remaining substantive procedures required to keep the audit risk at an acceptable low level.
16 Cf. DELOITTE (2020); EY (2019); PWC (2019); KPMG (2018).
17 This audit firm is referred to as „the audit firm“ or „the case audit firm“ throughout this thesis.

The central research questions of this thesis are:

(1) Which requirements have to be addressed by a data-based process analysis in an audit of financial statements?

(2) Which findings and challenges can be identified based on the empirical evaluation of the implementation of process mining in the audit practice?

(3) Provided the methodical requirements on the one hand and the practical challenges on the other hand, which modifications of process mining are necessary to address the key points of criticism?

1.2 Course of the investigation

The course of the investigation is following the three research questions raised:

At (1): Requirements that have to be addressed by a data-based process analysis in an audit of financial statements

Chapter 2 deals with the integration of process mining into the context of an audit of financial statements. In order to explore the possible application areas of process mining in the different phases of the audit, Chapter 2.1 introduces the fundamentals of a process analysis with process mining. This includes the basic functionality of the technology, the data structures required as well as common areas of application. Chapter 2.2 provides an overview of the state of scientific research related to (a) process mining in general and (b) the suitability of the technology for use in an audit of financial statements in particular. Chapter 2.3 commences with the theoretical integration of the data-based process analysis into the audit of financial statements and the identification of methodical requirements that need to be addressed. As prerequisites for embedding process mining into the audit, Chapter 2.3.1 covers the fundamentals of the risk-based audit approach and Chapter 2.3.2 summarizes the different audit phases, including (1) audit planning, (2) risk identification and risk assessment, (3) addressing identified risks and (4) concluding the audit and reporting. In its current state of development, process mining is used for the analysis of individual processes. By this, the main areas of application in an audit of financial statements are particularly the phase of risk identification and risk assessment and the phase of addressing the identified risks. Chapter 2.3.3 and Chapter 2.3.4 deal with embedding process mining into these phases, respectively, by analyzing how the audit procedures prescribed by related ISA may be supported with process mining. Chapter 2.3.5 closes the discussion with an interim conclusion about the suitability of process mining as an audit instrument from a methodical point of view.

At (2): Empirical evaluation of the implementation of process mining in the audit practice

The theoretical integration of process mining into the ISA in Chapter 2 is used as a basis in Chapter 3 to empirically evaluate the suitability of process mining to support an audit of financial statements.[18]

Chapter 3.1 starts with evaluating the audit firm's initial feasibility assessment that has been conducted in 2017 in Germany, Switzerland and Austria. In Chapter 3.1.1, the scope and design of the feasibility assessment are introduced, including the process mining application used and the key modifications made to adjust the vendor's software to the requirements of an audit of financial statements. Chapter 3.1.2 summarizes the feedback obtained from the audit teams who applied the technology alongside their traditional financial statement audit. It further provides insights obtained by the author of this thesis as part of her quality assurance procedures performed during the piloting. Chapter 3.2 is structured similar to the previous chapter and covers the extended piloting of the audit firm's process mining solution conducted in 2018 with audit teams from Europe, Middle East and Africa. Chapter 3.2.1 includes the modifications made to the application based on the feedback and insights obtained from the initial feasibility assessment. Chapter 3.2.2 deals with the feedback obtained from the audit teams and observations made during quality assurance procedures. Based on the results of the feasibility assessment and the piloting of process mining in an audit of financial statements, the audit firm decided to conduct a global certification of its process mining analyzer built for the purchase to pay process. Chapter 3.3 includes the results from the certification process and the scope and design of the subsequent implementation of the certified process mining analyzer in Europe, Middle East, Africa and Japan (Chapter 3.3.1). A summary and evaluation of the audit team feedback and the results of the quality assurance procedures performed over the course of the implementation is provided in Chapter 3.3.2. As part of the different piloting projects, various challenges in using process mining to support audit procedures are identified that are not yet discussed in available audit related process mining research. Chapter 3.4 summarizes the implementation process and completes the empirical evaluation with an interim conclusion on methodical and technical key questions and challenges that

[18] As part of her employment at the international audit firm, the author of this thesis supervised the development of the process mining application from an audit methodology perspective, wrote related methodical guidance, lead the piloting projects and was involved in the global certification.

need to be answered or resolved, respectively, to enable a broader adoption of process mining in the audit profession.

At (3): Identification of required modifications and development of solution approaches to address the key points of criticism for applying process mining in the audit

Based on the theoretical analysis in Chapter 2 and the empirical evaluation in Chapter 3, the main points of criticism for applying process mining in the audit are subject to a systematic analysis in Chapter 4. The chapter focuses on the theoretical and practical challenges that are of significant importance from an audit methodology and application perspective. It further includes the development of solution approaches and proposes modifications of the current process mining to address each challenge identified.

The term “mining” in “process mining” implies that the technology has an immanent explorative character. However, an extensive process understanding and knowledge of the relevant IT systems are needed already in preparation of the actual process analysis. This discrepancy is discussed in Chapter 4.1. Chapter 4.1.1 introduces the decisions and challenges in preparing the data, designing the analyses and customizing the application that have been identified as part of the empirical evaluation. Considering the enterprise resource planning (ERP) system, interfaces to other IT applications and configurations and customizations of each individual auditee poses challenges to the development of a process mining solution that is scalable to many different audit engagements (Chapter 4.1.2). Chapter 4.1.3 discusses the findings of the empirical evaluation relating to decisions on the activities to include (or not to include) in the event log. The choice of the process instance that is followed through the process execution is a fundamental prerequisite for assembling the input data for process mining. Related implications on the audit evidence that may be obtained with process mining are covered in Chapter 4.1.4. Chapter 4.1.5 introduces the challenges in capturing the date and timestamp information for certain activities that need to be considered when analyzing the throughput times between activities, durations of activities and the sequence of activities in the process flow. The definition of the data extraction period, resulting inclusion criteria for cases in the analyzer and the implications on the audit procedures that may or may not be performed are discussed in Chapter 4.1.6. Limitations regarding the completeness and accuracy of the input data further reduce the reliability and relevance of process mining for the audit and are discussed in Chapter 4.1.7. After summarizing the identified challenges, in Chapter 4.1.8, a solution approach is developed that contributes to continuously lifting the assumptions and prerequisites related to current process mining techniques that limit the current scope of application in auditing.

Subject of the audit are the financial statements of the auditee.[19] Thus, one of the fundamental conditions for the use of process mining in a financial statement audit is the enrichment of the process data with financial data. Chapter 4.2 addresses the integration of process data and financial data and discusses related challenges and opportunities. To avoid over-auditing of classes of transactions with a limited risk of material misstatement and under-auditing of significant classes of transactions, the process data needs to be reconciled to the related transaction movements on the balance sheet and income statement accounts in the general ledger (Chapter 4.2.1). The integration of journal entry recording events enables the auditor to benefit from the data-driven process analysis along the entire critical path of a process, i.e., beyond the initiating and processing activities only (Chapter 4.2.2). As metrics related to a process instance are stored on the level of the case and thus keep constant throughout the period, current process mining applications do not support to infer from a sequence of events performed to the value of the related process instance at a specific point in time. Chapter 4.2.3 uses the invoice amount of a transaction as an example to propose a solution that is based on the integration and reconciliation of process and subledger data. The proposed concept is applied in Chapter 4.2.4 to reconstruct the vendor balances and open items at a given point in time, differentiating between transaction volumes that are covered by the process data and additional postings made to the subledger. Chapter 4.2.5 demonstrates how the integration of process and financial data may enable an analytical approach to address the completeness of purchases during the period and liabilities at period end.

Process mining provides full transparency over all the different process paths evident in the data. However, due to the multitude of process variations they cannot be evaluated in their entirety in practice. The question how to deal with that in the course of the audit is yet to be answered. Chapter 4.3.1 discusses the emergence of process variations and their root-causes based on case studies conducted in process mining litera-

[19] Cf. IFAC (2021), ISA 200, para. 11(a).

ture and the application of process mining in the audit practice. In Chapter 4.3.2, concurrent evaluation approaches for process variations are analyzed. Considering their shortcomings for the application in an audit of financial statements, alternative approaches are developed to support identifying, assessing and addressing risks of material misstatement resulting from the different ways a process is executed.

Chapter 4.4 examines the modifications of process mining that are necessary in order to increase the efficiency of the technology when dealing with multiple data extraction periods. Chapter 4.4.1 explores the practical challenges when using process mining to compare the audit period's data with the process mining application of the previous audit period or with the procedures performed at an interim date. In developing a concept to compare different types of process data, Chapter 4.4.2 explores how process mining may be enhanced with a comparison functionality.

Chapter 4.5 focuses on the audit methodology related questions and practical challenges that need to be answered when using process mining for assessing the design and operating effectiveness of internal controls. Chapter 4.5.1 describes key developments related to internal controls in process mining theory and audit practice. These developments include the introduction of the Financial Market Integrity Strengthening Act reinforcing the importance of internal control, the divergence of theoretical and empirical research on using process mining to test internal controls and related recurring findings from professional inspections. How the data-driven analysis of a business process in combination with determining the impact of a particular process instance on the financial statements may enhance the auditor's control testing procedures is discussed in Chapter 4.5.2.

Chapter 5 answers the three research questions raised. The key points of criticism for process mining as an audit instrument presented in Chapter 4 as well as the identified solution approaches are summarized to an empirically reasoned criticism. The defined requirements for a data-based process analysis in an audit of financial statements and the empirical insights obtained serve as the basis for a concluding assessment of the appropriateness of process mining as an instrument of the data-driven audit of the future.

Chapter 6 closes with an outlook on further questions and challenges resulting for theory and practice concerned with the implementation of process mining in an audit of financial statements.

Figure 1 illustrates the course of the investigation.

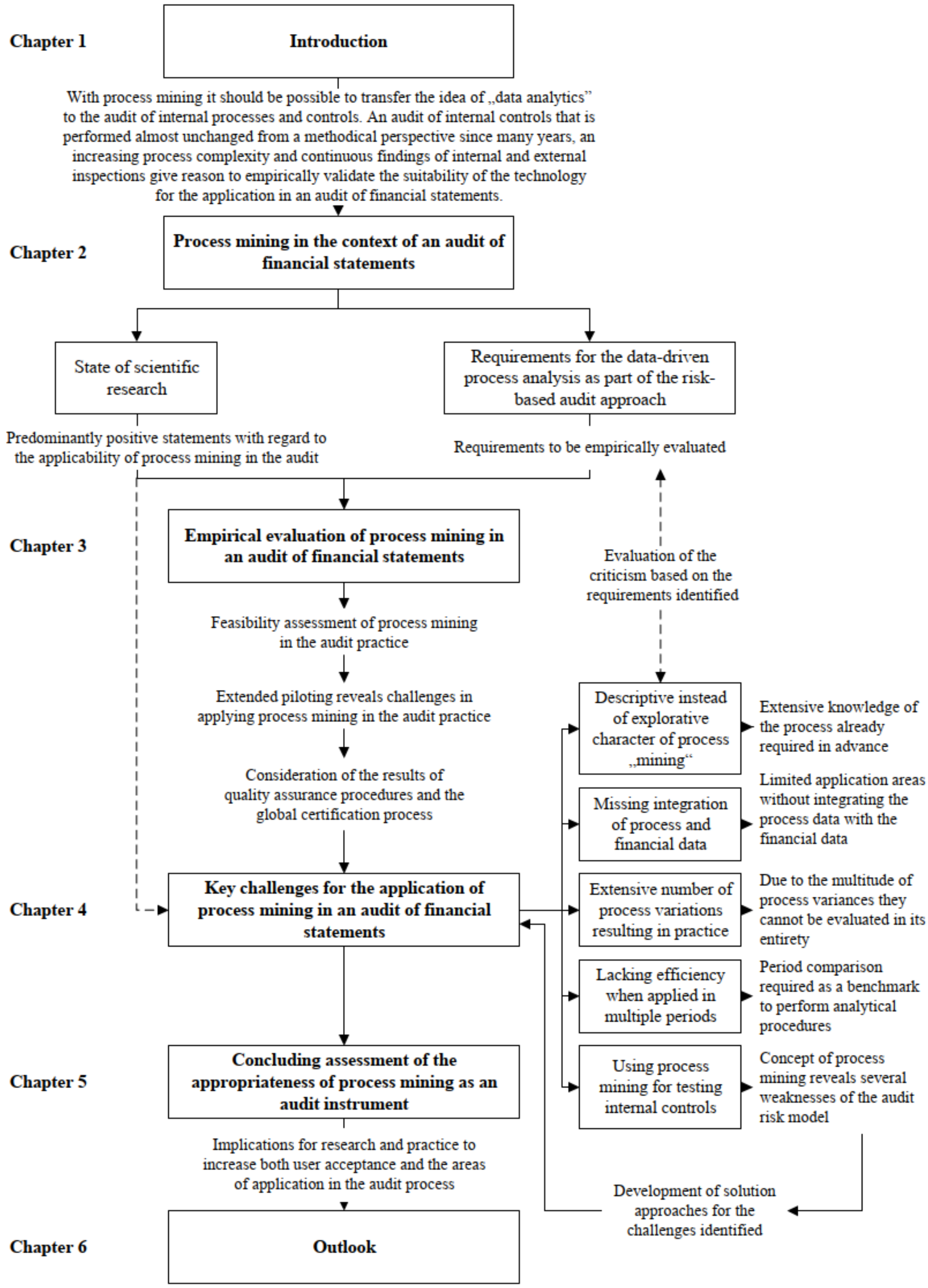


Figure 1: Course of the investigation

2 Process mining in the context of an audit of financial statements

2.1 Process analysis with process mining

Discovering, analyzing and evaluating business processes

The execution of business processes is subject to numerous economic, organizational and regulatory requirements. The economical point of view focuses on the cost-efficient execution of a process whereas regulatory requirements include, for example, the compliance with the requirements of the Sarbanes-Oxley-Act (SOX).[20] Due to these requirements imposed to business processes, companies are interested in obtaining insights in the actual execution of a process, assessing the effectiveness of implemented controls and identifying bottlenecks and improvement potential.[21] These activities may be summarized by the term „Business Process Management" (BPM).[22] At the same time, the auditor of the financial statements is required to obtain an understanding of the entity and the control environment,[23] including business processes and control activities relevant to the financial statements.[24]

Until recently, the analysis and assessment of internal processes and integrated controls has been a manual task. This includes studying the theoretical description of the process in process manuals, inquiries with process owners, the observation and analysis of process outcomes and testing samples of control procedures performed within the process.[25] However, even if there is a common design for a process, the way in which individuals carry out essentially the same process can differ. Frequently, there are significant differences between the process design and the actual execution of a business process. As changing economic conditions may require adjustments or extensions of processes in a short period of time, processes are often historically grown, resulting in an increasing process complexity. In addition, time and cost constraints in the daily business may lead to alternative ways of executing a process that differ from the initial design of the target process flow.[26]

---

20 Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 354.
21 Cf. ibid.
22 Cf. DUMAS, MARLON et al. (2018), p. 1.
23 Cf. IFAC (2021), ISA 315 (Revised 2019), para. 19 and 21.
24 Cf. ibid., para. 25 and 26.
25 Cf. BODENMANN, ANDREAS/MAAS, ANNALOUISE/SCHWAN, MORITZ (2018), p. 41.
26 Cf. ibid.; JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 2; COOK, JONATHAN E./WOLF, ALEXANDER L. (1998a), p. 241.

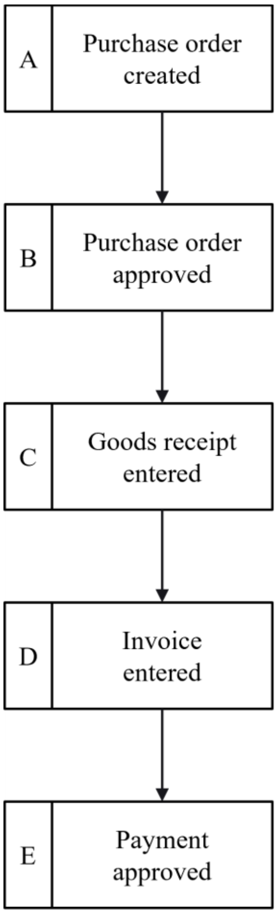
Today's information systems, for example, customer relationship management (CRM) systems or workflow management systems, usually track and continuously record the execution of business processes in so-called "log files".[27] By this, the computer-assisted elements of the process execution leave digital traces in the system.[28] This data does not only include entries made by the system's users but also information related to these user activities that is recorded automatically by the system.[29] Process mining is a technology that extracts process related data from an information system in order to reconstruct the process flow as it actually occurred.[30] By this, process mining has the potential to replace a large portion of the traditional manual and subjective assessment of processes and overcome its limitations. This includes inquiries of process owners who may be influenced by subjective perceptions, observations of employees who may behave differently when they are aware of the observation and reviews of process manuals that may be outdated.

The basic concept of process mining

Process mining is based on the so-called "event log".[31] An event log is a chronological record of activities performed in an information system that is derived from the log files recorded by the system.[32] In general, an event log is related to a separated process, for example, the purchase to pay process of an entity. An "event" is defined as the execution of an activity within a specific process instance ("case"),[33] for example, the payment of a specific invoice. Each event is tracked with an unique ID in a separate line of the event log. Figure 2 shows a simplified presentation of an event log and the related activities in the purchase to pay process.

---

[27] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), pp. 7f.
[28] Cf. BODENMANN, ANDREAS/MAAS, ANNALOUISE/SCHWAN, MORITZ (2018), p. 41.
[29] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2014), p. 1753.
[30] Cf. VAN DER AALST, WIL M. P. (2016), p. 25.
[31] Cf. ibid., p. 10.
[32] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), p. 3. Such log files are frequently used by system administrators or IT staff to analyze errors and their root causes, for example, by reproducing the process steps performed immediately in advance to a system failure.
[33] The terms „process instance" and „case" are used synonymously in scientific literature and in this thesis.



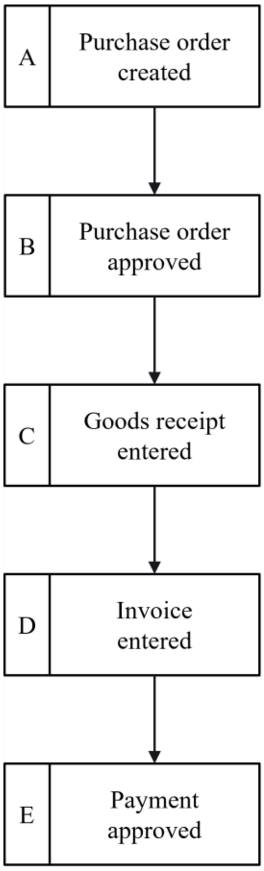
Case	Activity
1	A
2	A
1	B
1	C
3	A
2	C
3	B
2	B
1	D
2	D
2	E
3	C
3	D
1	E
3	D
3	E

Figure 2: Simplified presentation of an event log in purchase to pay

The event log presented in Figure 2 includes three different process instances (for example, invoices). For these cases, the activities A (purchase order created), B (purchase order approved), C (goods receipt entered), D (invoice entered) and E (payment approved) are performed. The expected sequence of these activities is presented at the left side of Figure 2. For each event, the event log contains information about the performed activity and the related process instance, usually referenced by a unique case ID.[34] To perform meaningful analyses, the events within a case need to be sorted. Thus, the event log usually includes the timestamp of the events, i.e., the date and time of the respective execution of an activity.[35] This minimum information in the event log (the event, its timestamp and the related case ID) enables to determine for each process instance which process steps (activities) have been executed in which order.

34 Cf. VAN DER AALST, WIL M. P. (2016), p. 128.
35 Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 5.

Figure 3 illustrates how process mining processes the information of the event log and makes them available for process analyses.



Case	Activity
1	A
2	A
1	B
1	C
3	A
2	C
3	B
2	B
1	D
2	D
2	E
3	C
3	D
1	E
3	D
3	E

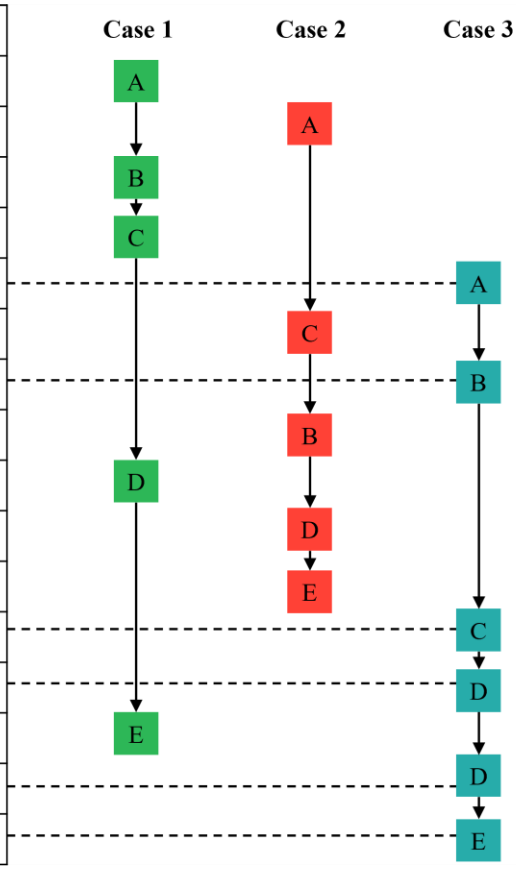


Figure 3: Visualization of the process paths for different process instances

In a first step, the activities performed for each individual process instance are extracted from the event log in chronological order. The visual presentation shows that for invoice 1, all activities have been executed in the expected order. For invoice 2, the goods have been received (activity C) before the related purchase order has been approved (activity B). The third process graph shows that all activities have been executed for invoice 3, however the activity "invoice entered" (activity D) has been performed twice. Process instances that follow the same activities in exactly the same sequence are summarized in a process variation, i.e., one specific process path ("trace"). Thus, a process variation includes all cases for which the process has been executed in exactly the same way. In the example provided, each invoice makes up a separate process variation.

Summarizing all process variations in one process graph results in the visualization of the overall process. The process mining technology uses specific algorithms to group the cases into variations and construct the process visualization.[36] Figure 4 summarizes the basic concept of process mining.[37]

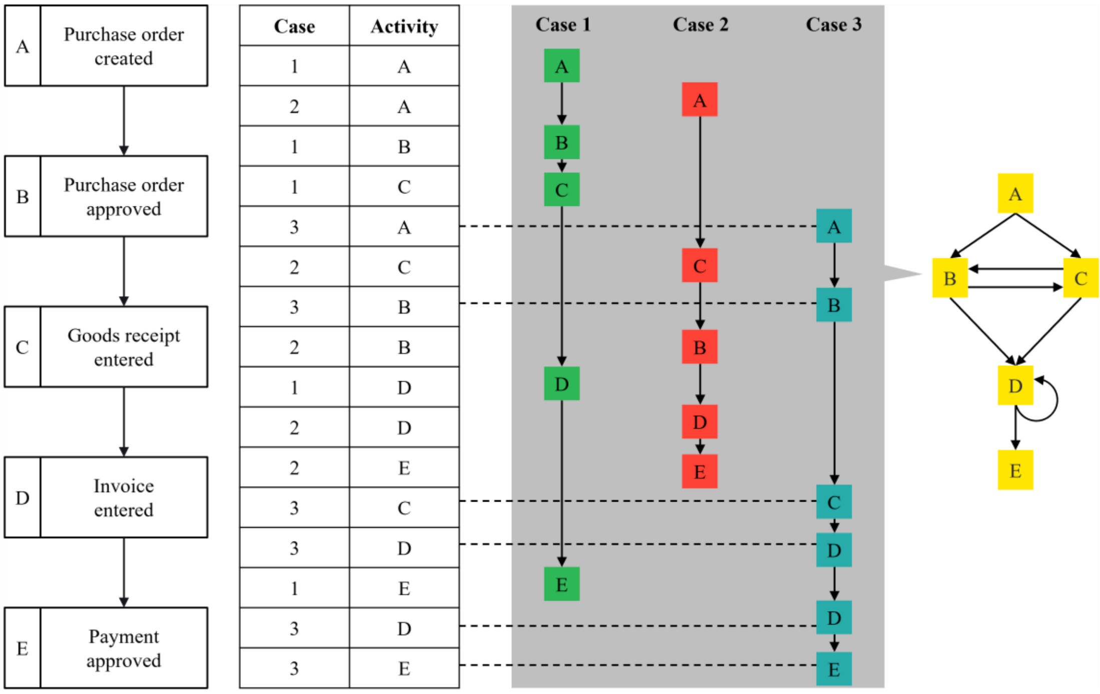


Figure 4: Basic concept of process mining

This is the fundamental concept behind process mining. The actual process flow is reconstructed based on the data stored in the information system and provides a basis for data-driven process analyses. The value of the technology for auditing is especially based on the circumstance that in theory, manually identifying and understanding the target process that has been initially envisioned by the entity is no longer required when using process mining. The auditor obtains fact-based information about the actual process extracted from the system and can evaluate its appropriateness based on the understanding of the entity and its environment.

Appendix I includes a glossary of common process mining terms and their definition.

---

[36] Cf. VAN DER AALST, WIL M. P. (2016), p. 33.

[37] Usually, the overall process graph displays each activity in the process only once. Repetitions of individual activities within a process instance are shown as a loop, i.e., by a looping arrow on the respective activity as displayed for activity D in the process graph of Figure 4.

Basic data model and information content of process mining

Table 1 shows an excerpt of an event log describing the purchase to pay process of an international trading company.[38]

EventID	Case ID	Event time	Activity	User ID	Transaction code	Material number	Document reference
300821903162400010|63484	300821903162400010	24-01-2018 16:59:45	Purchase order created	CKH	ME21N	110151	300821903162400010
300821903162400010|63485	300821903162400010	24-01-2018 17:26:21	Purchase order approved	KXW	ME29N	110151	300821903162400010
300821903162400010|63486	300821903162400010	09-02-2018 00:00:00	Invoice date stated	DIC	MIRO	110151	300510564780420180001
300821903162400010|63487	300821903162400010	13-02-2018 15:43:48	Goods receipt posted	CKH	MIGO_GR	110151	3005000181663201801
300821903162400010|63488	300821903162400010	15-02-2018 22:32:42	Invoice posted	DIC	MIRO	110151	300510564780420180001
300821903162400010|63489	300821903162400010	13-03-2018 08:14:17	Invoice settled by payment run	MTF	F110	110151	300510564780420180001
300821903162400010|63490	300821903162400010	16-03-2018 14:02:16	Bank clearing posted	JHK	FB05	110151	300510564780420180001
300821903162600010|63516	300821903162600010	24-01-2018 17:04:16	Purchase order created	CKH	ME21N	110151	300821903162600010
300821903162600010|63517	300821903162600010	24-01-2018 17:26:00	Purchase order approved	KXW	ME29N	110151	300821903162600010
300821903162600010|63518	300821903162600010	01-03-2018 00:00:00	Invoice date stated	DIC	MIRO	110151	300510567760320180001
300821903162600010|63519	300821903162600010	08-03-2018 21:20:10	Goods receipt posted	CKH	MIGO_GR	110151	3005000272050201801
300821903162600010|63520	300821903162600010	15-03-2018 03:11:55	Invoice posted	DIC	MIRO	110151	300510567760320180001
300821903162600010|63521	300821903162600010	04-04-2018 09:15:21	Invoice settled by payment run	MTF	F110	110151	300510567760320180001
300821903162600010|63522	300821903162600010	11-04-2018 13:25:53	Bank clearing posted	JHK	FB05	110151	300510567760320180001
300821903162800010|63543	300821903162800010	24-01-2018 17:08:18	Purchase order created	CKH	ME21N	110151	300821903162800010
300821903162800010|63544	300821903162800010	24-01-2018 17:25:36	Purchase order approved	KXW	ME29N	110151	300821903162800010
300821903162800010|63545	300821903162800010	25-01-2018 00:00:00	Invoice date stated	DIC	MIRO	110151	300510565263220180001
300821903162800010|63546	300821903162800010	01-02-2018 20:28:08	Goods receipt posted	CKH	MIGO_GR	110151	3005000132253201801
300821903162800010|63547	300821903162800010	22-02-2018 00:40:06	Invoice posted	DIC	MIRO	110151	300510565263220180001
300821903162800010|63548	300821903162800010	27-02-2018 18:12:31	Invoice settled by payment run	MTF	F110	110151	300510565263220180001
300821903162800010|63549	300821903162800010	01-03-2018 12:40:09	Bank clearing posted	JHK	FB05	110151	300510565263220180001

Table 1: The event log as the basis of process mining

Besides the event ID, the case ID, the individual activity names and their timestamps, additional attributes may be recorded in the event log. These usually include the user ID of the individual executing the activity, a reference to the related document as well as process related information like the material group or transaction code. While not all events within the event log need to have the same set of attributes, the same attributes are usually recorded for all events referring to the same activity.[39]

Table 2 shows the corresponding excerpt of the case details.

Case	Case start time	Case end time	Case owner	Vendor	Document currency	PO amount (Local curency)	Invoice amount (Local currency)	PO quantity	Goods receipt quantity	Invoice quantity
300821903162400010	24-01-2018 16:59:45	16-03-2018 14:02:16	ABBB	AIKHH	EUR	361,46	353,98	20,00	20,00	20,00
300821903162600010	24-01-2018 17:04:16	11-04-2018 13:25:53	ABBB	AIKHH	EUR	8.552,60	8.481,95	1,00	1,00	1,00
300821903162800010	24-01-2018 17:08:18	01-03-2018 12:40:09	ABBB	AIKHH	EUR	588,90	576,47	8,00	8,00	8,00
300821903163000010	24-01-2018 17:17:23	01-03-2018 12:40:09	ABBB	AIKHH	EUR	1.523,93	1.491,78	1,00	1,00	1,00
300821903164200010	25-01-2018 22:26:07	25-01-2018 22:26:07	ABBB	RXTF	EUR	28.093,05	0,00	2,00	0,00	0,00
300821903164500010	26-01-2018 02:44:05	08-06-2018 11:06:53	SYP	BPRGM	CHF	6.409,50	6.409,50	2,00	2,00	2,00
300821903164500020	26-01-2018 02:44:05	08-06-2018 11:06:53	SYP	BPRGM	CHF	702,00	702,00	2,00	2,00	2,00
300821903164500030	26-01-2018 02:44:05	08-06-2018 11:06:53	SYP	BPRGM	CHF	241,50	241,50	2,00	2,00	2,00
300821903164500040	26-01-2018 02:44:05	10-04-2018 10:55:19	SYP	BPRGM	CHF	5,17	5,17	10,00	10,00	10,00
300821903164500050	26-01-2018 02:44:05	10-04-2018 10:55:19	SYP	BPRGM	CHF	3,45	3,45	20,00	20,00	20,00
300821903164600010	26-01-2018 02:46:34	10-04-2018 10:55:19	SYP	UQYE	CHF	15.998,55	15.998,56	1,00	1,00	1,00

Table 2: Case details providing information on the attributes stored on case level

The case ID is linking the case information of the case table with their respective events in the event log. The starting and ending times of a case correspond with the timestamps of the first and last event performed for the case. The case table further includes information on the case owner, i.e., the individual that is responsible of the case, and the vendor related to the case. The document currency is included as the

[38] The practical examples provided in this thesis are based on sanitized real-world data derived from the empirical implementation described in Chapter 3.

[39] Cf. VAN DER AALST, WIL M. P. (2016), p. 130.

currency of documents within the process may differ from the local reporting currency of the entity. Further, the case table may include information on the total amount and/or quantity of the documents related to a case, for example, purchase order, invoice and goods receipt.

The event and case tables may usually not be extracted from the IT system in the structured format illustrated in Table 1 and Table 2. Instead, after extracting the raw data, the information is assembled from multiple data sources and transformed into the standardized data format used as input for the process mining application.

To a large extent, the potential of process mining for the audit of internal processes and integrated controls is driven by the nature of the input data. The process mining input data not only includes data entered by the system user („input data“) but also information about these entries that is recorded by the information system automatically and independently from the system user („meta data“).[40] For example, an employee is required to log-in to the system with a unique ID and password and when performing an activity, the system records a timestamp, regardless of whether the employee also enters a date or not.[41] Figure 5 compares the input data and event log data using the example of an invoice document.

40 Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2014), pp. 1751f.

41 Cf. ibid., p. 1754.

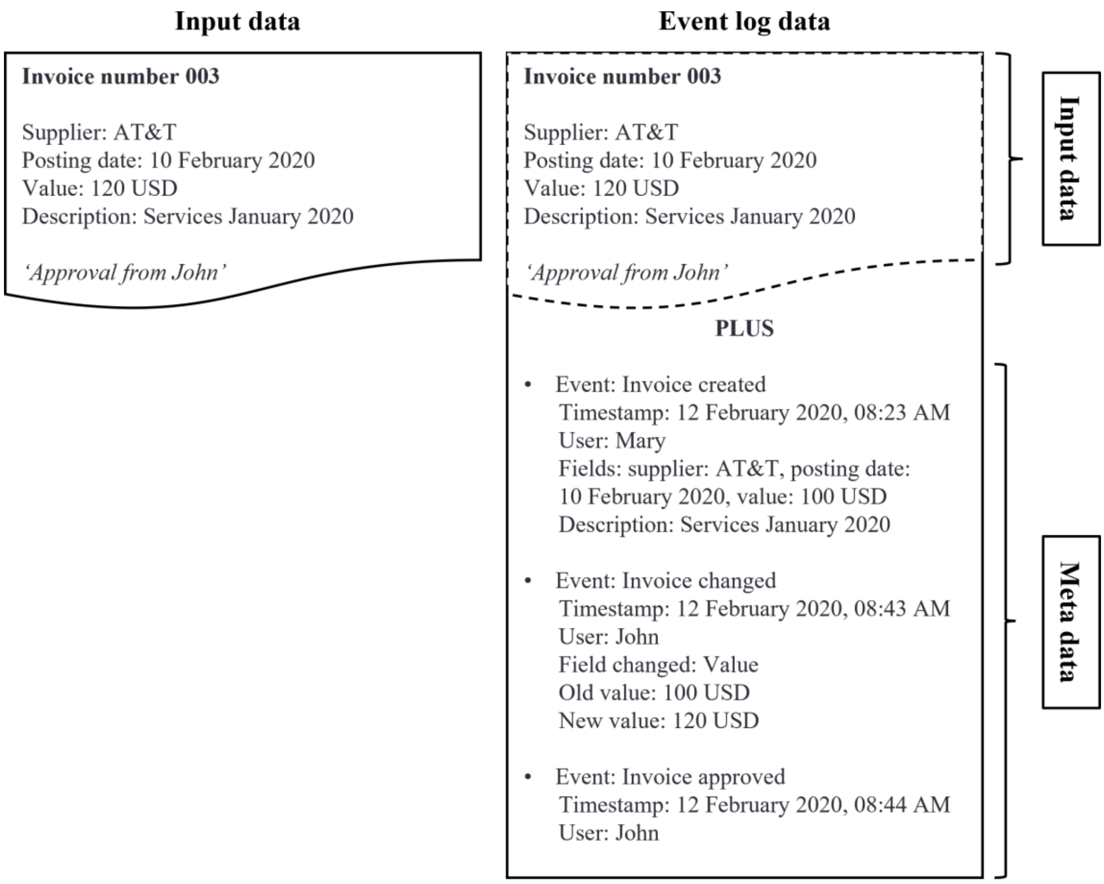


Figure 5: Comparison of input data and event log data of an invoice[42]

The left side of Figure 5 shows the input data, i.e., the data of an invoice that is entered into the information system by an employee of the organization. This data is already used in today's audits of financial statements, for example, in the data-based analysis of the general ledger or as part of the journal entry testing according to ISA 240.[43] All the user-entered information related to the invoice are also included in the event log data of process mining displayed at the right side of Figure 5. However, in addition, the events and information on the related event attributes included in the event log provide contextual information of the transaction that has been entered into the system.[44] In the example provided, besides the transactional data related to the entry of the invoice, the event log includes the user, fields and timestamp information related to the creation, change and approval of the invoice document in the ERP system. This meta data is recorded automatically and independently from the employee whose behavior is subject to the audit and enables reconstructing the history of a particular

[42] Adjusted from Jans, Mieke/Alles, Michael G./Vasarhelyi, Miklos A. (2013), p. 6.
[43] Cf. IFAC (2021), ISA 240, para. 33(a).
[44] Cf. Jans, Mieke/Alles, Michael G./Vasarhelyi, Miklos A. (2013), p. 5.

transaction.[45] For example, instead of relying on inquiries, observations and testing samples of transactions, the user information in the event log related to the creation, change and approval of documents may support the auditor in assessing the appropriateness of the segregation of incompatible duties.

Process mining techniques and perspectives

The objective of process mining is "*to discover, monitor, and improve real (not assumed) processes*"[46]. There are three different process mining techniques distinguished:

- process discovery,
- conformance checking and
- process enhancement.[47]

Figure 6 summarizes the techniques and their different input and output data.

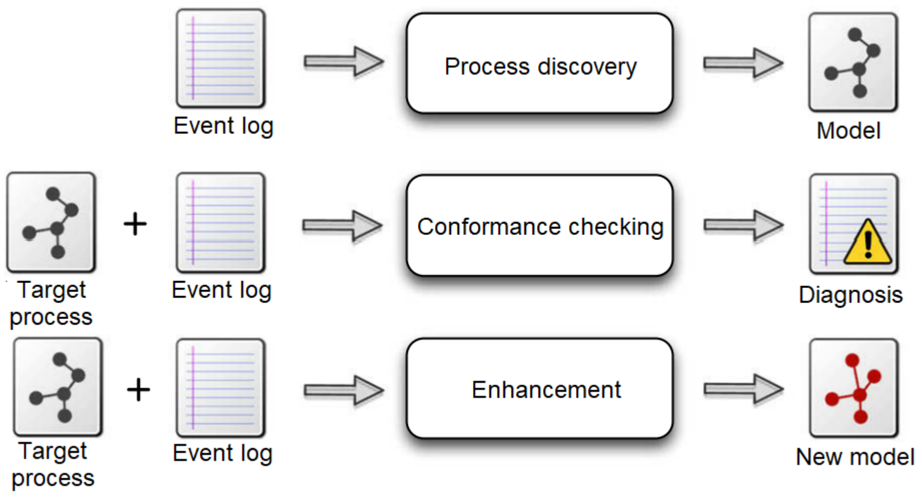


Figure 6: Types of process mining techniques and their input and output data[48]

Process discovery techniques use the event log to construct a process model and do not require any additional a priori information.[49] The process models that are produced and the analyses that can be performed directly depend on the information in the event log.

Conformance checking techniques compare the event log with an existing normative process model in order to determine if the actual process execution as recorded in the

---

45 Cf. ibid., pp. 6f.
46 VAN DER AALST, WIL M. P. et al. (2010), p. 90.
47 Cf. VAN DER AALST, WIL M. P. (2016), p. 33.
48 Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 355.
49 Cf. VAN DER AALST, WIL M. P. (2016), p. 33.

event log conforms with the designed target process.[50] Conformance checking enables to detect and explain deviations between both models, to measure the severity of deviations identified and to determine the overall degree of conformance.[51]

Process enhancement techniques aim to enhance an existing (target) process model by using the information on the actual process in the event log.[52] The model may either be repaired (i.e., modified to better represent the actual process) or extended with additional information.[53]

The different process mining techniques support analyzing different aspects of the process.[54] Thus, in parallel to the three process mining techniques, different process mining perspectives can be identified. The process model illustrated in Figure 4 focuses on the control-flow or process perspective, i.e., the sequence of activities in the process. However, the three different types of process mining are not limited to the control-flow perspective.[55] When using process enhancement techniques to extend the existing process model, additional perspectives can be added based on the information in the event log.[56] Besides the process or control-flow perspective, analyses performed with process mining frequently relate to the organizational perspective, the case perspective and the time perspective.[57] If information on resources is provided in the event log, process discovery techniques may be used to discover resource-related process models, for example, a social network analysis of the interactions of users involved in the process.[58] Conformance checking techniques may then be used to validate existing organizational models.[59] Analyzing roles and responsibilities in the process is referred to as the organizational perspective of the process. The case perspective focuses on analyzing different properties of cases to determine the characteristics of a particular transaction.[60] For example, when investigating cases where goods receipts have been returned, it may be interesting to include the related vendors and materials

---

50 Cf. ibid.; ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 358.
51 Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 358.
52 Cf. VAN DER AALST, WIL M. P. (2016), p. 33.
53 Cf. ibid. A detailed description of the individual process mining techniques is provided by VAN DER AALST, WIL M. P. (2016), pp. 171ff.
54 AILENEI, IRINA et al. (2011) provide an overview of common process mining use cases.
55 Cf. VAN DER AALST, WIL M. P. (2016), p. 34.
56 Cf. ibid.
57 Cf. ibid.; JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 11; ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 355.
58 Cf. VAN DER AALST, WIL M. P. (2016), p. 33.
59 Cf. ibid., p. 34.
60 Cf. ibid.; JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 11.

in the analysis. Similarly, information on the fields that have been changed in a particular purchase order may support identifying the root-cause for purchase orders being rejected. The last category of common process mining perspectives relates to the timestamp information in the process. Analyses include, for example, discovering delays and bottlenecks in the process by investigating throughput times, monitoring process performance, evaluating the adherence to service agreements and measuring the utilization of resources in the process.[61]

2.2 State of scientific research

Computer science as the origin of process mining

Analyzing business processes based on log data stored in information systems is not new to computer science and economics. In the late 1990s, COOK/WOLF described methods for process discovery based on event data in software engineering[62] and introduced first measures to quantify differences between a designed process model and the actual process reconstructed from the data.[63] In the same decade, AGRAWAL/GUNOPULOS/LEYMANN were the first to transfer the idea of process mining to workflow management systems using computer-generated log files to re-construct processes.[64] In the subsequent years, the research focused on developing methods and exploring the use of process mining in the context of workflow management. HERBST and HERBST/KARAGIANNIS introduce an inductive approach to mine sequential and concurrent workflow process models,[65] SCHIMM develops a tool capable of discovering hierarchical models,[66] and WEIJTERS/VAN DER AALST propose a heuristic approach to construct process graphs.[67] An overview of the research conducted on using process mining in the context of workflow management systems until 2002 is

61 Cf. VAN DER AALST, WIL M. P. (2016), p. 34; ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 355.
62 Cf. COOK, JONATHAN E./WOLF, ALEXANDER L. (1998a); COOK, JONATHAN E./WOLF, ALENXANDER L. (1998b).
63 Cf. COOK, JONATHAN E./WOLF, ALEXANDER L. (1999).
64 Cf. AGRAWAL, RAKESH/GUNOPULOS, DIMITROS/LAYMANN, FRANK (1998).
65 Cf. HERBST, JOACHIM (2001); HERBST, JOACHIM/KARAGIANNIS, DIMITRIS (1999); HERBST, JOACHIM/KARAGIANNIS, DIMITRIS (1998).
66 Cf. SCHIMM, GUIDO (2003); SCHIMM, GUIDO (2002).
67 Cf. WEIJTERS, A. J. J. M./VAN DER AALST, WIL M. P. (2003); WEIJTERS, A. J. J. M./VAN DER AALST, WIL M. P. (2001).

provided by VAN DER AALST et al.[68] Based on the research results, the authors summarize available mining approaches to support workflow design and propose a common format for event logs derived from workflow management systems.[69]

Research progress on the three process mining techniques

Over the last two decades, the research on process mining expanded significantly. A large body of scientific publications is available that explores the three process mining techniques,[70] ranging from the development of new or improved algorithms or mining techniques for process discovery over the exploration of methods for testing and measuring process conformance to the evaluation of techniques to enhance the quality of the overall process model.

In analyzing 705 journal articles and conference papers published between 2005 and 2014, MAITA et al. point out that the vast majority of the studies focuses on developing or improving methods for process discovery.[71] Table 3[72] illustrates the variety of studies on mining techniques available for process discovery.

[68] Cf. VAN DER AALST, WIL M. P. et al. (2003).
[69] Cf. ibid.
[70] Cf. Chapter 2.1.
[71] 504 of the analyzed studies deal with process discovery, 259 studies with process conformance and 120 with process enhancement techniques, cf. MAITA, ANA R. C. et al. (2017), pp. 520 and 522.
[72] Adapted and extended from TIWARI, ASHUTOSH/TURNER, CHRIS/MAJEED, BASIM (2008), pp. 8f. and AUGUSTO, ADRIANO et al. (2019a), pp. 689f.

Publication	**Alpha algorithm**	**Genetic miner**	**Heuristic miner**	**Inductive miner**	**Declarative miner**	**Split miner**	**Fuzzy miner**	**Semantic miner**	**Trace clustering**	**Other approaches**
Weijters/van der Aalst (2001)			X							
Weijters/van der Aalst/de Medeiros (2006)			X							
Weijters/Ribeiro (2011)			X							
de Cnuddle/Claes/Poels (2014)			X							
Augusto et al. (2017)			X							
Mannhardt et al. (2017)			X							
de Medeiros et al. (2004)	X									
van der Aalst/ Weijters/Maruster (2004)	X									
Wen et al. (2007)	X									
Guo et al. (2015)	X									
van der Aalst/de Medeiros/Weijters (2005)		X								
de Medeiros (2006)		X								
De Medeiros/Weijters/van der Aalst (2007)		X								
Augusto et al. (2019b)						X				
Maggi/Bose/van der Aalst (2012)					X					
Leemans/Fahland/van der Aalst (2013)				X						
Günther/van der Aalst (2007)							X			
Rozinat (2010)							X			
de Medeiros/van der Aalst (2007)								X		
de Giacomo et al. (2018)								X		
Okoye (2020)								X		
Greco et al. (2006)									X	
Song/Günther/van der Aalst (2008)									X	
de Weerdt et al. (2013)									X	
Neubauer/Fantinato/Peres (2019)									X	
Maruster et al. (2002)										X
Zhang et al. (2003)										X
Gaaloul/Godart (2005)										X
van der Aalst (2005)										X
van der Aalst/de Beer/van Dongen (2005)										X
van Dongen et al. (2005)										X
Goedertier et al. (2009)										X
van der Werf et al. (2009)										X
Huang/Kumar (2012)										X
Buijs/van Dongen/van der Aalst (2014)										X
van Zelst et al. (2017)										X

Table 3: Literature review on algorithms and methods for process discovery

BURATTIN, VAN DER AALST, MAGGI et al. and AUGUSTO et al. provide an overview of the individual technical differences of the numerous algorithms proposed for process discovery.[73] ROZINAT et al. remark that while each mining technique may target specific challenges in process discovery (for example, dealing with noise in logged data, mining concurrent processes and dealing with hidden or duplicate tasks),[74] all of them have their limitations.[75] As a consequence, the quality of the process models discovered with these algorithms varies significantly.[76]

With the vast number of different approaches available that aim to discover a "good" process model from event data, a field of study emerged for evaluating the quality and conformance of the discovered process model.[77] VAN DER AALST points out that the conformance of a discovered process model may be viewed from two angles: (a) the discovered model does not capture the actual process behavior or (b) the real behavior is captured but deviates from the desired process model.[78] Significant contributions to conformance checking with process mining are made by VAN DER AALST, ROZINAT et al., ROZINAT/VAN DER AALST, ADRIANSYAH/SIDOROVA/VAN DONGEN, ADRIANSYAH/VAN DONGEN/VAN DER AALST and DE LEONI/VAN DER AALST[79] and many subsequent publications in this field are based on research conducted by VAN DER AALST.[80] Considering the strengths and weaknesses of the individual process mining algorithms, until today, no best practice has been established and various process mining solutions available on the market are based on different mining techniques.[81]

---

73 Cf. AUGUSTO, ADRIANO et al. (2019a), pp. 689ff.; MAGGI, FABRIZIO MARIA et al. (2018), pp. 4f.; VAN DER AALST, WIL M. P. (2016), pp. 201ff.; BURATTIN, ANDREA (2015), pp. 35ff.

74 An overview on the challenges in process discovery is provided by VAN DER AALST, WIL M. P/WEIJTERS, A. J. M. M. (2004), pp. 241ff.

75 Cf. ROZINAT, ANNE et al. (2008), p. 85.

76 Cf. ibid., p. 84.

77 Cf. MAGGI, FABRIZIO MARIA et al. (2018); VAN DER AALST, WIL M. P. (2016); ROZINAT, ANNE et al. (2008).

78 Cf. VAN DER AALST, WIL M. P. (2016), p. 39.

79 Cf. ibid., pp. 243ff.; DE LEONI, MASSIMILIANO/VAN DER AALST, WIL M. P. (2013); ADRIANSYAH, ARYA/SIDOROVA, NATALIA/VAN DONGEN, BOUDEWIJN F. (2011); ADRIANSYAH, ARYA/VAN DONGEN, BOUDEWIJN F./VAN DER AALST, WIL M. P. (2011a); ADRIANSYAH, ARYA/VAN DONGEN, BOUDEWIJN F./VAN DER AALST, WIL M. P. (2011b); ROZINAT, ANNE et al. (2008); ROZINAT, ANNE/VAN DER AALST, WIL M. P. (2008); ROZINAT, ANNE/VAN DER AALST, WIL M. P. (2006).

80 Cf. VAN BEEST, NICK R. T. P. et al. (2015); GROMOV, VLADIMIR (2014); NGUYEN, HOANG H. et al. (2014); BUIJS, JOOS C. A. M. et al. (2013); RAMEZANI, ELHAM/FAHLAND, DIRK/VAN DER AALST, WIL M. P. (2012); LY, LINH T. et al. (2011).

81 Cf. KERREMANS, MARC (2019), pp. 77ff.; KEBEDE, MUSIE (2015), pp. 25ff.

Although process mining was initially developed by software engineers, different aspects of business processes are analyzed by researchers from various disciplines.[82] A large body of interdisciplinary research emerged over the last two decades that particularly deals with the identification of process enhancement potential. Process mining received increasing attention in healthcare,[83] supply chain and IT service management,[84] human resource management[85] and compliance management in financial institutions.[86] ADRIANSYAH/VAN DONGEN/VAN DER AALST investigate the cost of deviating from the target process in insurance claim handling,[87] VAN DER AALST/DE MEDEIROS deal with detecting security breaches in security networks[88] and DREHER/REIMANN/GRÖGER identify application fields and research gaps for process mining in manufacturing companies.[89] Recent studies further evaluate the use of process mining related to other emerging technologies, such as blockchain data[90] and the internet of things[91]. DOS SANTOS GARCIA et al. explore 1.278 research papers dealing with process mining and identify the predominant application domains healthcare, information and communication technology, manufacturing and education.[92]

While process mining research found its way into many different disciplines, studies on the practical adoption of the technology just emerge.[93] The vast majority of process mining publications between 2000 and 2017 listed at the website of the BPM Center[94] includes theoretical studies aiming at different types of technical improvement only. VOM BROCKE et al. criticize that there is a vacancy of research on how companies integrate the technology into their IT landscape and which kind of effects emerge from this adoption.[95] Based on their proposed research framework, GRISOLD et al. are some of the firsts to explore how process managers understand, evaluate and use process

---

82 Cf. JANS, MIEKE/ALLES, MICHAEL/VASARHELYI, MIKLOS A. (2010), p. 8.
83 Cf. PIKA, ANASTASIA et al. (2020); LEONARDI, GIORGIO et al. (2018); MANS, RONNY S./VAN DER AALST, WIL M. P./VANWERSCH, ROB J. B. (2015); FERNANDEZ-LLATAS, CARLOS et al. (2015); REBUGE, ÁLVARO/FERREIRA, DIOGO R. (2012); YANG, WAN-SHIOU/HWANG, SAN-YIH (2006).
84 Cf. EDGINGTON, THERESA M./RAGHU, T. S./VINZE, AJAY S. (2010); WYNN, MOE T. et al. (2009).
85 Cf. ARIAS, MICHAEL et al. (2018).
86 Cf. BECKER, MICHAEL/BUCHKREMER, RÜDIGER (2019).
87 Cf. ADRIANSYAH, ARYA/VAN DONGEN, BOUDEWIJN F./VAN DER AALST, WIL M. P. (2011a).
88 Cf. VAN DER AALST, WIL M. P./DE MEDEIROS, ALVES K. A. (2005).
89 Cf. DREHER, SIMON/REIMANN, PETER/GRÖGER, CHRISTOPH (2020).
90 Cf. HOBECK, RICHARD et al. (2021); PUFAHL, LUISE et al. (2021); KLINKMÜLLER, CHRISTOPHER et al. (2019); MÜHLBERGER, ROMAN et al. (2019); MENDLING, JAN et al. (2018).
91 Cf. MICHAEL, JUDITH et al. (2019).
92 Cf. DOS SANTOS GARCIA, CLEITON et al. (2019), p. 260.
93 Cf. VOM BROCKE, JAN et al. (2021); GRISOLD, THOMAS et al. (2021); KERREMANS, MARC (2019).
94 Cf. BPM CENTER (2023).
95 Cf. VOM BROCKE, JAN et al. (2021), p. 483.

mining tools, what potentials and risks they perceive when actually applying process mining and how enterprises are generating value from the technology.[96]

Process mining research in accounting and auditing

Despite the interdisciplinary research progress on process mining, applying the concept of process mining to accounting and auditing is a comparatively young research stream.[97]

Under the headline "Auditing 2.0", VAN DER AALST et al. predict that process mining will significantly change the audit profession, as the technology shall enable to move away from sample-based and ex-post testing of internal processes and controls to continuous monitoring and auditing of the entire population while the process is still running.[98] However, the authors focus on the technical characteristics and functionality of process mining rather than the specific ways of application during an audit. Since then, numerous contributions have been made by JANS and JANS et al., providing examples on the value process mining might add to the audit and conducting a case study to identify anomalous transactions not identified by the entity's internal auditors using traditional audit procedures.[99] A large opportunity is attributed to the contextual meta information of activities and transaction that is logged independently from the user performing the activity or entering the transaction into the system (and thus, independently from the auditee).[100]

Predominantly, the research on process mining continues to focus on workflow management systems. However, the related extraction methods and mining algorithms are of minor interest to auditing as they are not applicable to accounting processes recorded in an entity's ERP system. GEHRKE/MÜLLER-WICKOP and WERNER/GEHRKE/NÜTTGENS are the firsts to consider mining financial information by introducing "financial process mining".[101] Financial process mining uses the open item

---

[96] Cf. GRISOLD, THOMAS et al. (2021), pp. 369f.

[97] Cf. WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021); CHIU, TIFFANY/BROWN-LIBURD, HELEN/VASARHELYI, MIKLOS A. (2019); JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013); WERNER, MICHAEL/GEHRKE, NICK/NÜTTGENS, MARKUS (2012).

[98] Cf. VAN DER AALST, WIL M. P. et al. (2010), p. 90.

[99] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2014), p. 1751; JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2012), p. 1.

[100] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), pp. 14f.; ALLES, MICHAEL G./JANS, MIEKE/VASARHELYI, MIKLOS A. (2011), p. 3; JANS, MIEKE/DEPAIRE, BENOIT/VANHOOF, KOEN (2011), pp. 32f; JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), pp. 4f.; Chapter 2.1.

[101] Cf. WERNER, MICHAEL/GEHRKE, NICK/NÜTTGENS, MARKUS (2012); GEHRKE, NICK/MÜLLER-WICKOP, NIELS (2010).

accounting functionality of ERP systems to link documents and items and visualize dependencies.[102] The authors argue that with financial process mining, the manual modeling of these processes becomes obsolete.[103] However, the approach is limited to accounts with open item accounting on the one hand and no longer considers the process data on the other hand. Removing the process perspective from process mining simultaneously removes the potential benefits to the auditor's understanding of the critical path[104] of a transaction, including the potential redundance of manual walkthrough procedures. MÜLLER-WICKOP/SCHULTZ present an algorithm that adds timestamp information to process instances mined with financial process mining to enable the integration into a sequential event log based on process data.[105] However, no details on the practical feasibility of analyzing and interpreting such an event log are provided. For example, the question how to analyze an event log that is no longer based on a single process instance (such as the purchase order for which the approval event is performed) but on multiple process instances (such as both the purchase order and an open item cleared by a payment run) is yet to be answered.[106] A significant contribution is made by MOCHTY,[107] who first explores the intersection of mining process and accounting data by systematically evaluating how a meaningful combination of the event log with financial data used on every audit - the general ledger of an entity - may enhance auditing internal controls over financial reporting.

The first study considering an initial methodical integration of process mining into the audit process is provided by WERNER/WIESE/MAAS,[108] who investigate how process mining may be systematically embedded into a financial statement audit as required by the ISA.

---

[102] Cf. WERNER, MICHAEL (2017), p. 1; GEHRKE, NICK/MÜLLER-WICKOP, NIELS (2010), p. 3.
[103] Cf. WERNER, MICHAEL (2017), p. 1.
[104] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 18(b). The "critical path" describes the flow of a transaction through the system, i.e., the initiation, processing, recording and the reporting of the transaction in the financial statements.
[105] Cf. MÜLLER-WICKOP, NIELS/SCHULTZ, MARTIN (2013).
[106] Further details on the practical implications of the chosen process instance are provided in Chapter 4.1.4.
[107] Cf. MOCHTY, LUDWIG (2015).
[108] Cf. WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021).

Table 4 summarizes relevant research papers exploring the use of process mining in internal or external auditing while providing an overview on suggested audit procedures and fields of application identified by the respective authors.[109]

Publication	Audit procedures that may be supported by process mining					Fields of application and analyses identified				
	Understand the process	Risk identification	Tests of controls	Analytical procedures	Tests of details	Variation/outlier analysis	Sequence or timing of activities	Segregation of duties	Resource analysis	Missing approval or authorization
van der Aalst (2010)	X	X				X				
Jans (2012)	X	X								
Jans (2019)	X	X	X	X	X				X	
Jans/Alles/Vasarhelyi (2010)	X	X	X	X	X	X	X	X	X	X
Alles/Jans/Vasarhelyi (2011)	X	X	X	X	X	X			X	
Jans/Alles/Vasarhelyi (2012)	X	X	X	X	X	X	X	X	X	X
Jans/Alles/Vasarhelyi (2013)	X	X	X	X	X	X	X	X	X	X
Jans/Alles/Vasarhelyi (2014)	X	X	X	X	X	X	X	X	X	X
Jans/Depaire/Vanhoof (2011)	X	X	X	X	X	X	X	X	X	X
AICPA (2017)	X	X	X	X		X	X	X		X
Chiu (2018)	X	X	X	X	X	X	X	X	X	X
Chiu/Brown-Liburd/Vasarhelyi (2019)	X	X	X		X		X			
Mochty (2015)	X	X	X	X	X	X			X	
Werner (2016)	X	X	X			X		X	X	X
Werner (2017)	X	X	X			X	X	X	X	X
Werner/Gehrke (2019)	X	X	X		X			X	X	X
Werner/Wiese/Maas (2021)	X	X	X	X	X	X	X		X	X

Table 4: Literature review on audit related research on process mining

All authors emphasize the potential of process mining to add value to an audit of financial statements, especially to the evaluation of business processes and related internal controls. There is wide agreement that the technology has the potential to support the auditor in understanding the process and identifying and assessing related risks of material misstatement. While tests of controls are most commonly identified as supported audit procedures, most authors agree that process mining may further support analytical procedures and tests of details of individual process instances and by this may be useful in all phases of an audit. Exemplary analyses frequently referred to in-

[109] To enable comparability, the analysis is limited to research related to this thesis that is exploring the added value and fields of application of process mining in auditing. Predominantly technical publications or studies focusing on mining ERP data or financial data only (as outlined earlier in this chapter) are not considered.

clude (1) the identification of risks of material misstatement and anomalous transactions through the investigation of individual process variations or (2) the sequence of individual activities, (3) the evaluation of the segregation of incompatible duties and (4) appropriate authorization of transactions as common tests of controls and (5) resource analyses to understand roles and responsibilities and to identify potential collusion between users involved in the process or between users and other actors such as suppliers.

Overall, there is theoretical consensus in related research that process mining adds value to auditing and may be used to increase both quality and efficiency of audit procedures throughout all phases of the audit.

However, available research lacks empirical validation and evaluation. Many publications neither do include a practical example nor are validated with real-life data. Some studies indicate that a case study has been performed, however, aforesaid case studies are limited to a single and relatively small dataset, while five of the studies listed in Table 4 refer to the same dataset from an European bank.[110] Even when analyzing the real-world data of the European bank and researching in full cooperation with the internal auditors of the case company,[111] none of the authors performs a holistic investigation of outliers or anomalies identified to conclude if they result in a material misstatement, represent a control exception or if the analysis performed actually results in "false positive" transactions. Instead, the authors note that the detailed investigation is part of the follow-up procedures that would need to be performed by the auditor.[112] As a result, the criticism raised by VOM BROCKE et al. and GRISOLD et al.[113] regarding a missing consideration and evaluation of the actual application of process mining also holds true for research related to accounting and auditing.

Although all Big Four audit firms invested in process mining in the past,[114] until now, the technology did neither revolutionize the audit as predicted by VAN DER AALST[115] nor established itself as a common instrument for auditing internal processes and controls. This thesis aims to contribute to concurrent scientific research on process mining by providing (a) a detailed methodical integration of the technology into the ISA,

110 Additional details on these studies and the case study are provided in Chapter 4.3.1.
111 Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2012), p. 4.
112 Cf. ibid.; CHIU, TIFFANY (2018), pp. 22f.
113 Refer to the previous section of this chapter covering the interdisciplinary process mining research.
114 Cf. DELOITTE (2020); EY (2019); PWC (2019); KPMG (2018).
115 Cf. VAN DER AALST, WIL M. P. et al. (2010), p. 90.

(b) insights into the actual application of process mining in the audit practice and (c) a discussion of practical key challenges and solution approaches to support a broader adoption of the technology in the audit profession.

2.3 Requirements for the data-driven process analysis as part of the risk-based audit approach

2.3.1 Fundamentals of the risk-based audit approach

The audit risk model

The auditor has to conclude with "*reasonable assurance*"[116] whether the financial statements are free from material[117] misstatement.[118] Thus, an audit of financial statements is performed risk-based, following the concept of the audit risk model.

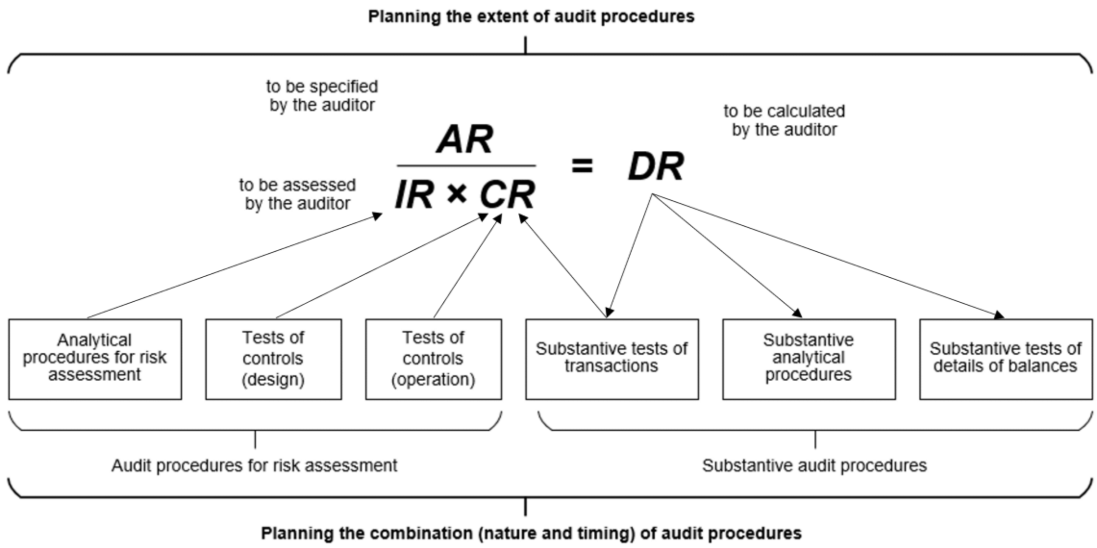


Figure 7: Audit risk model[119]

The acceptable audit risk (AR) is the probability that the auditor expresses a positive opinion on the financial statements although a material misstatement[120] exists.[121] In

[116] IFAC (2021), ISA 200, para. 5.
[117] A misstatement (individual or aggregated with other misstatements) is material if it is expected to qualitatively or quantitatively influence the decisions made by the stakeholders of the financial statements, cf. ibid., para. 6.
[118] Cf. ibid., para. 11(a).
[119] Cf. ELDER, RANDAL J. et al. (2020), p. 375; WIESE, MICHAEL (2013), p. 17.
[120] In the following, the term "material misstatement" is referring to a misstatement that is material individually or when aggregated with other misstatements.
[121] Cf. IFAC (2021), ISA 200, para. 13(c).

practice, the audit risk is defined by the auditor in advance of the audit.[122] The inherent risk (IR) describes the probability that material misstatements exist in the financial statements, without consideration of any related internal controls.[123] The control risk (CR) represents the likelihood that a material misstatement exists in the financial statements that is not prevented, or detected and corrected, by the entity's internal controls.[124] Multiplying the inherent risk and the control risk in the audit risk model results in the risk of material misstatement, i.e., the probability that material misstatements exist in the financial statements.[125] Dividing the acceptable audit risk by the risk of material misstatement leads to the detection risk (DR). The detection risk is the risk that material misstatements exist in the financial statements that are not prevented, or detected and corrected, by the entity's internal controls and that are not identified by the auditor.[126]

When the audit risk is kept constant, the formula of the audit risk model shows a direct relationship between the risks of material misstatement and the detection risk.[127] If the assessment of the risks of material misstatement increases, the acceptable detection risk decreases and the extent of audit evidence to be obtained by the auditor (1-DR) increases.[128] In this scenario, the extent of substantive audit procedures needs to be increased in order to obtain sufficient appropriate audit evidence to address the increased risks of material misstatement.[129] That is, the actual "audit need" of an entity (that is unknown to the auditor) is dependent of the extent of material misstatements (unknown to the auditor as well) existing in the financial statements.[130]

Besides determining the extent of audit procedures required, the audit risk model is used to plan the nature and timing of the audit procedures in order to obtain sufficient appropriate audit evidence.[131] Audit procedures can be performed as risk assessment

---

122 Due to limitations inherent to the audit, the audit risk cannot be reduced to zero in order to achieve absolute audit assurance, cf. IFAC (2021), ISA 200, para. 5. In audit practice, the audit risk is usually defined at five to ten percent, cf. WIESE, MICHAEL (2013), p. 18. Thus, the complement of the audit risk (1-AR) indicates that the audit practice interprets the „reasonable assurance" as an audit assurance between 90 and 95 percent.
123 Cf. IFAC (2021), ISA 200, para. 13(n)(i).
124 Cf. ibid., para. 13(n)(ii).
125 Cf. ibid., para. 13(n).
126 Cf. ibid., para. 13(e).
127 Cf. ibid., para. 47.
128 Cf. ibid.; WIESE, MICHAEL (2013), p. 18.
129 Cf. IFAC (2021), ISA 315 (Revised 2019), para. 7.
130 Cf. WIESE, MICHAEL (2013), p. 20.
131 Cf. ibid., p. 17; IFAC (2021), ISA 315 (Revised 2019), para. 13(a) and (b).

procedures or as substantive procedures.[132] Figure 7 illustrates the relation of the audit procedures to the individual components of the audit risk model.

Risk assessment procedures include analytical procedures to assess inherent risk and procedures to assess the design and operating effectiveness of internal controls (and thus, ultimately assess control risk).[133] Based on the assessment of the risks of material misstatement, substantive procedures are planned to respond to the identified risks. By this, the detection risk is kept at an acceptable level and the audit risk formula balanced.[134] Substantive procedures include substantive tests of transactions, substantive tests of details of balances and analytical procedures at the assertion level.[135] Thus, detection risk can be divided into the test of details risk (TDR) and the analytical procedures risk (APR), as illustrated in Figure 8.[136]

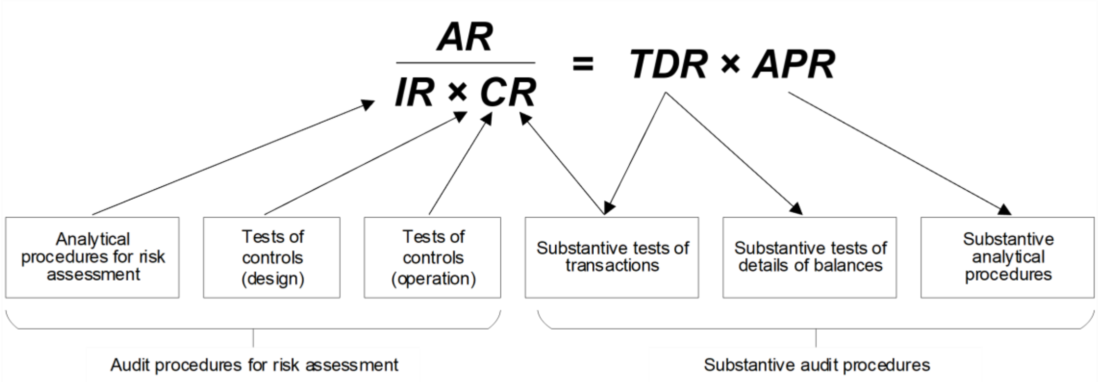


Figure 8: Components of detection risk[137]

As a result of the relationships between the individual components of the audit risk model, assessing risks and determining the "audit need" of an entity is a dynamic and iterative process.[138] It requires continuously revisiting the audit strategy, i.e., identifying an effective and efficient combination of audit procedures as a response to

---

[132] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 13; IFAC (2021), ISA 330, para. 4; WIESE, MICHAEL (2013), p. 19.
[133] Cf. ELDER, RANDAL J. et al. (2020), p. 375; WIESE, MICHAEL (2013), p. 19.
[134] Cf. ELDER, RANDAL J. et al. (2020), p. 375.
[135] Cf. ibid., p. 376.
[136] Cf. WIESE, MICHAEL (2013), p. 20; GÄRTNER, MICHAEL (1994), p. 54.
[137] Cf. WIESE, MICHAEL (2013), p. 20.
[138] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 7; IFAC (2021), ISA 500, para. 6; IFAC (2021), ISA 300, para. 10.

the identified risks of material misstatement.[139] An audit strategy is effective and efficient if the auditor performs neither more (over-auditing) nor less (under-auditing) procedures than required to meet the audit objective.[140]

The concept of materiality and disaggregating the financial statements into smaller segments

The audit is performed using the concept of materiality.[141] As the audit opinion is provided on the financial statements as a whole, the auditor is not required to detect misstatements that are not material to the financial statements.[142] According to ISA 200, "*misstatements, including omissions, are considered to be material if, individually or in the aggregate, they could reasonably be expected to influence the economic decisions of users taken on the basis of the financial statements*"[143]. Determining materiality is a matter of professional judgment and is influenced by the auditor's perception of the information needs of the users of the financial statements.[144] As the audit has to be performed in such a way that material misstatements, whether due to fraud or error, are detected with reasonable assurance,[145] a relationship between the interests of the users of the financial statements and the amount of audit evidence to accumulate is established.[146] Materiality depends on the users of financial statements on the one hand and influences the detection risk on the other hand. The lower the materiality level, the more potential misstatements are material and the higher the risk for the auditor to not detect such a misstatement.

Figure 9 illustrates the auditor's steps in applying the concept of materiality.

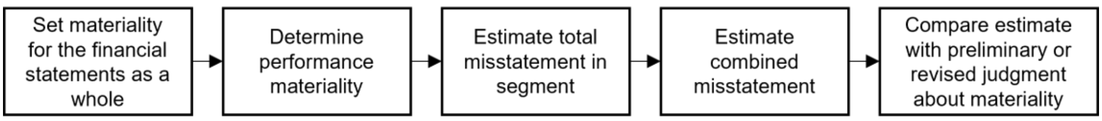


Figure 9: Steps in applying materiality[147]

Materiality is first determined for the overall financial statements and set to the maximum amount for which the auditor believes a misstatement would still not affect the

[139] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 7; IFAC (2021), ISA 500, para. 6.
[140] Cf. MOCHTY, LUDWIG/WIESE, MICHAEL (2012), p. 8.
[141] Cf. IFAC (2021), ISA 320, para. 5 and 8.
[142] Cf. IFAC (2021), ISA 200, para. 6.
[143] Ibid.
[144] Cf. IFAC (2021), ISA 320, para. 4.
[145] Cf. ibid., para. 11(a).
[146] Cf. ELDER, RANDAL J. et al. (2020), p. 211.
[147] Cf. ibid., p. 210, own presentation.

economic decisions of the users of the financial statements.[148] In a second step, a certain percentage of materiality ("performance materiality") is allocated to individual segments of the audit.[149] The financial statements are disaggregated into smaller segments to reduce the complexity of the audit and support assigning audit procedures to different members of the audit team.[150] It is subject to the auditor's professional judgment how to disaggregate the financial statements. For example, each account (or account class) of the balance sheet and income statement could be treated as individual segment. However, this approach does not consider the relationships between the accounts (for example, trade receivables and sales).[151] A method proven of value in the audit practice is to assign closely related types of transactions, the account balances resulting from these transactions and related disclosures to the same segment.[152] Figure 10 illustrates the accounts and types of transactions that are affected by the class of transactions "purchase to pay".

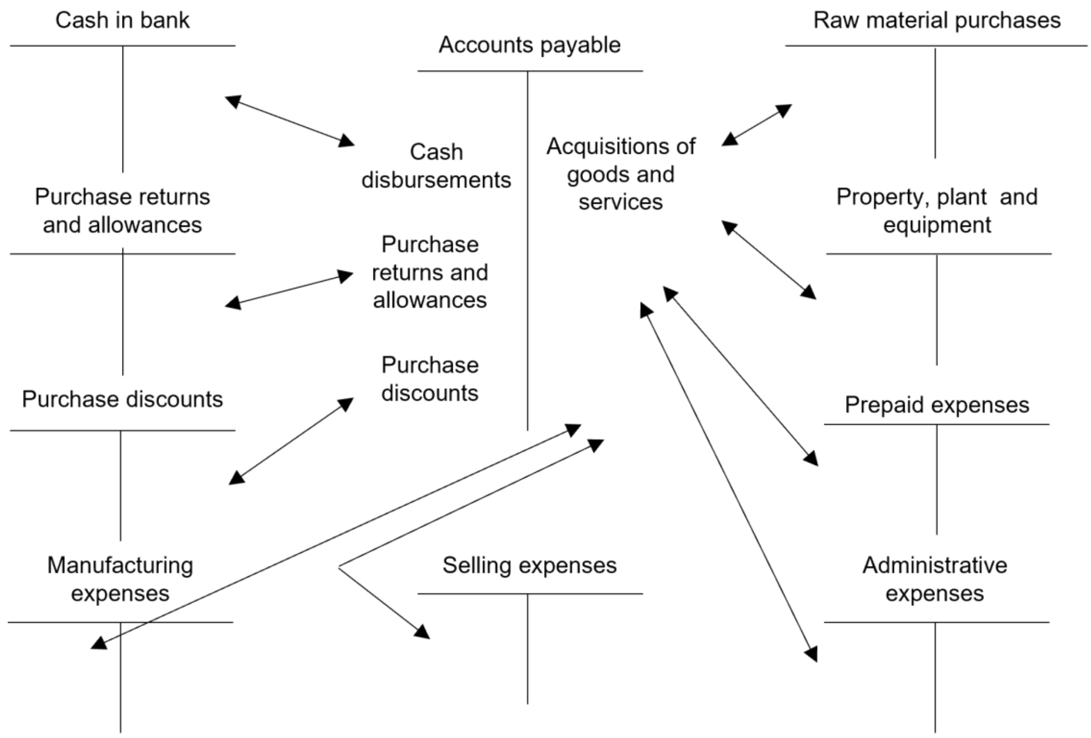


Figure 10: Accounts and types of transactions in the purchase to pay process[153]

[148] Cf. ELDER, RANDAL J. et al. (2020), pp. 210f.; IFAC (2021), ISA 320, para. 10.
[149] Cf. IFAC (2021), ISA 320, para. 11; ELDER, RANDAL J. et al. (2020), pp. 210f. The performance materiality for individual segments of the audit is usually set at 50 to 75 percent of the overall materiality and may vary for different account balances, classes of transactions and disclosures, depending on the interests of the financial statements' users, cf. ELDER, RANDAL J. et al. (2020), p. 214.
[150] Cf. ibid., p. 129.
[151] Cf. ibid.
[152] Cf. ibid.
[153] Cf. ibid., p. 566.

For example, the acquisition of goods and services, cash disbursements and purchase returns are types of transactions that cause the accounts payable to increase or decrease. To effectively manage the complexity of the audit, auditor's typically treat the individual segments separately to the extent possible in practice.[154] Once the audit of the individual segments is completed and the relationships with other segments have been assessed, the results are aggregated to an overall estimate for the combined misstatement of the financial statements.[155] As a final step in applying the concept of materiality, the estimate of the combined misstatement is compared to the judgment about materiality to determine if the financial statements as a whole are materially misstated. If a material misstatement is detected and not corrected by the entity, the auditor is required to issue a qualified or adverse audit opinion, depending on the nature and materiality of the misstatement.[156]

Determining the audit objectives

In applying the audit risk model to determine if the financial statements are free from material misstatement, the auditor determines if the recorded information correctly reflects the economic events that occurred during the accounting period. Figure 11 summarizes the flow of information from the business event to the auditee's financial statements.

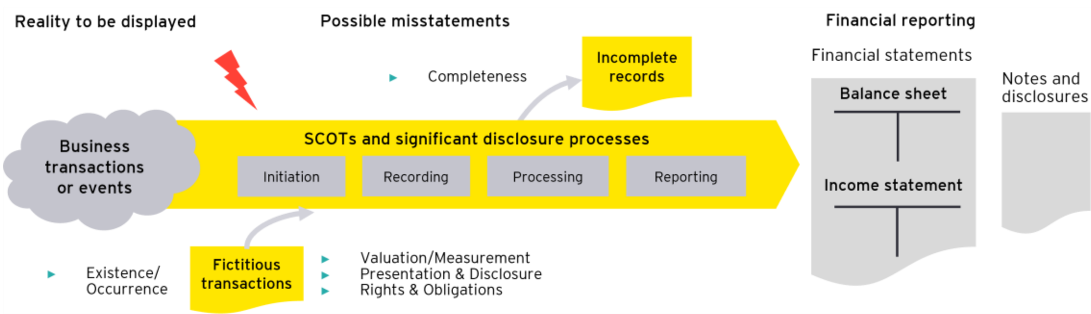


Figure 11: Flow of information from the business event to the financial statements

The ending balance of an account within the financial statements results from the beginning balance at the start of the audit period that is increased or decreased by transactions throughout the period. In other words, the difference between the audited financial statements of the prior period and the financial statements of the audit period

[154] Cf. Elder, Randal J. et al. (2020), p. 133.
[155] Cf. ibid., p. 210.
[156] Cf. ibid.

is the business activity of the auditee throughout the audit period, including the initiation, recording, processing and reporting of transactions within a significant class of transactions (SCOT).[157] From a bookkeeping perspective, the business activity is represented by transactions recorded as journal entries that cause additions or disposals on the balance sheet or income statement accounts. Assuming that the beginning balance has been audited in the prior period and thus is considered as correctly stated, if the auditor was absolutely sure that the classes of transactions affecting the account are correctly stated, he or she could conclude that the audit period's ending balance is stated correctly as well.[158] However, as it is impractical to obtain absolute assurance, the audit of the individual segments of the financial statements is not limited to tests of the transactions that are making up the ending balances, but the auditor obtains some combination of assurance for each individual class of transactions and for the related account balances and disclosures as of period end in order to increase the overall assurance.[159]

As indicated in Figure 11, the flow of information from the business event to the financial statements might be disrupted by misstatements. For example, if a purchase that occurred during the audit period is not recorded in the subledger, the records are incomplete and both inventory and accounts payable at period end are understated. As such, several audit objectives must be met to conclude that a class of transactions, account balance or a related disclosure is free from material misstatement.[160] ISA 315 refers to these audit objectives as assertions, as they are derived from management's implicit or explicit assertions about the recognition, measurement, presentation and disclosure of the financial statements.[161]

Assertions about classes of transactions and events and related disclosures include the following:

- Occurrence: Recorded or disclosed transactions have actually occurred.
- Completeness: Transactions that should have been recorded have actually been recorded and disclosures that should have been included in the financial statements have been included.

[157] Cf. IFAC (2021), ISA 315, para. 25(a)(i).
[158] Cf. ELDER, RANDAL J. et al. (2020), p. 134.
[159] Cf. ibid., pp. 133f.; IFAC (2021), ISA 330, para. 18.
[160] Cf. ELDER, RANDAL J. et al. (2020), p. 134.
[161] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 12(a) and A189.

- Accuracy: Transactions have been recorded using the appropriate data, including appropriate amounts, and the related disclosures have been appropriately measured and described.
- Cutoff: Transactions have been recorded in the correct period.
- Classification: Transactions have been recorded using the appropriate accounts.
- Presentation: Transactions are aggregated or disaggregated appropriately and related disclosures are relevant and clearly described.[162]

Assertions about account balances and related disclosures include:

- Existence: Assets, equity interests and liabilities exist.
- Rights and obligations: The entity holds or controls the rights to assets and liabilities are the entity's obligations.
- Completeness: All assets, equity interests and liabilities that should have been recorded have actually been recorded and all disclosures that should have been included in the financial statements have been included.
- Accuracy, valuation and allocation: Assets, equity interests and liabilities as well as any valuation or allocation adjustments have been recorded at appropriate amounts and the related disclosures have been appropriately measured and described.
- Classification: Assets, equity interests and liabilities have been recorded using the appropriate accounts.
- Presentation: Assets, equity interests and liabilities are aggregated or disaggregated appropriately and related disclosures are relevant and clearly described.[163]

These general audit objectives are adjusted by the auditor as necessary in the circumstances and transformed into specific audit objectives for a given class of transactions or financial statement account.[164] For example, the general assertion about the occurrence of transactions may be transformed into the specific assertion for cash disbursements in the purchase to pay cycle stating that recorded cash disbursements are for goods and services actually received by the entity.[165] As indicated in Figure 11, the

---

[162] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A190(a).
[163] Cf. ibid., para. A190(b).
[164] Cf. ELDER, RANDAL J. et al. (2020), p. 137.
[165] Cf. ibid., p. 576.

assertions are used by the auditor to identify the different types of potential misstatements that may occur within a significant class of transactions or account balance and for which sufficient and appropriate audit evidence needs to be obtained.[166] An assertion that is related to a risk of material misstatement is a relevant assertion for the audit.[167] Consequently, the auditor needs to design appropriate audit procedures to address the inherent risks identified at the assertion level.

2.3.2 The different phases of the audit process

The auditor needs to set the audit objectives and determine the evidence to obtain for meeting these objectives. In doing so, the auditor follows a dedicated process that ensures all audit objectives are specified and met by accumulating sufficient and appropriate audit evidence.[168] The different phases of the audit process include (1) audit planning, (2) risk identification and risk assessment, (3) addressing identified risks and (4) concluding the audit and reporting.

Audit planning

The audit needs to be planned in order to be performed in an effective manner.[169] An appropriately planned audit supports the auditor, for example, in devoting attention to important areas of the audit, in identifying and resolving any issues on a timely basis and in selecting engagement team members with appropriate levels of capabilities to respond to anticipated risks.[170]

Audit planning involves preliminary engagement activities, establishing the overall audit strategy and developing the audit plan.[171] As part of the preliminary engagement activities, the auditor (a) decides if a new entity is accepted or an existing entity relationship is continued,[172] (b) evaluates the compliance with relevant ethical requirements according to ISA 220, including independence,[173] and (c) establishes the terms of the engagement according to ISA 210, including why the entity wants or needs an

[166] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A1.
[167] Cf. ibid., para. A188.
[168] Cf. ELDER, RANDAL J. et al. (2020), p. 142.
[169] Cf. IFAC (2021), ISA 300, para. 4.
[170] Cf. ibid., para. 2.
[171] Cf. ibid., para. 2 and 6.
[172] Cf. ibid., para. 6(a); IFAC (2021), ISA 220, para. 12 and 13.
[173] Cf. IFAC (2021), ISA 300, para. 6(b); IFAC (2021), ISA 220, para. 9 to 11.

audit.[174] The preliminary engagement activities are usually completed prior to performing other significant activities[175] to ensure any significant costs that incur can be recovered.[176] The audit strategy defines the scope, timing and direction of the audit.[177] Factors that are considered in determining the audit strategy include the characteristics of the engagement that define its scope, the reporting objectives to plan the timing and the communication requirements of the audit, the assignment and direction of resources as well as knowledge obtained from preliminary engagement activities, the previous audit of the entity and the auditor's other audit engagements.[178] Based on the audit strategy, the auditor develops an audit plan that includes information on the nature, timing and extent of the planned risk assessment procedures according to ISA 315 and the nature, timing and extent of audit procedures performed in accordance with ISA 330 in response to the identified risks.[179]

Audit planning is not an isolated phase at the beginning of the audit but is continuous and iterative in nature.[180] For example, new risks of material misstatement identified as part of executing the planned audit procedures may impact the audit plan and result in consequential changes to the overall audit strategy.[181]

Risk identification and risk assessment

The risk identification and risk assessment phase of the audit is regulated in ISA 315.[182] The auditor's responsibility is to identify and assess the risks of material misstatement at the overall financial statement level and at the assertion level of classes of transactions, account balances and related disclosures. The risk assessment provides a basis for designing and implementing responses to address the identified risks according to ISA 330.[183] Figure 12 illustrates the risk identification and risk assessment process according to ISA 315.

---

[174] Cf. IFAC (2021), ISA 300, para. 6(c); IFAC (2021), ISA 210, para. 9 to 13; ELDER, RANDAL J. et al. (2020), p. 197.
[175] Cf. IFAC (2021), ISA 300, para. A7.
[176] Cf. ELDER, RANDAL J. et al. (2020), p. 197.
[177] Cf. IFAC (2021), ISA 300, para. 7.
[178] Cf. IFAC (2021), ISA 300, para. 8 and the related Appendix to the ISA.
[179] Cf. ibid., para 9.
[180] Cf. ibid., para. A2.
[181] Cf. ibid., para. 10 and A10.
[182] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 1.
[183] Cf. ibid., para. 11 and 28. Risks of material misstatement according to ISA 315 (Revised 2019) include both those due to error and due to fraud. Due to the significance of risks of material misstatement due to fraud, additional requirements are regulated in ISA 240, cf. IFAC (2021), ISA 315 (Revised 2019), para. 6.

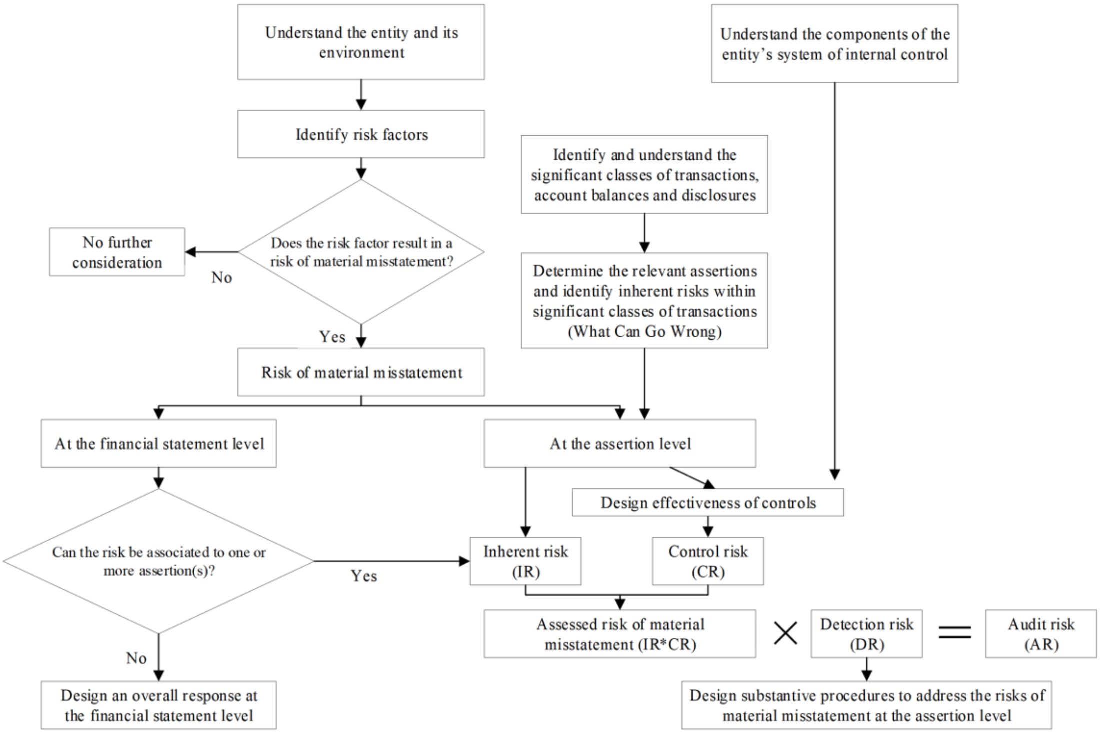


Figure 12: The auditor's risk identification and risk assessment process

The auditor's risk assessment is based on an understanding of the entity and its environment. In obtaining an understanding of the entity's business, the auditor obtains information on (i) the organizational structure and the business model, (ii) industry, regulatory and other external factors, (iii) the factors used to assess the entity's financial performance and (iv) the entity's accounting policies and the applicable financial reporting framework.[184] The auditor's understanding of the entity further includes understanding the entity's system of internal control. As part of understanding the individual components of internal controls,[185] the auditor understands for each significant class of transactions, account balance and related disclosure how transactions and events are initiated, recorded, processed through the entity's information system and reported in the financial statements.[186] To obtain the understanding of the entity and its environment, including the entity's system of internal control, the auditor performs risk assessment procedures. Risk assessment procedures include inquiries of appropriate personnel, observations of business practices, inspections of documents and analytical procedures.[187]

[184] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 9.
[185] Cf. ibid., para. 21 to 26.
[186] Cf. ibid., para. 25(a).
[187] Cf. ibid., para. 14. Risk assessment procedures are discussed in more detail in Chapter 2.3.3.

The understanding obtained enables the auditor to identify inherent risk factors and determine if a risk factor gives rise to an inherent risk at the overall financial statement level or at the assertion level of a class of transactions, account balance or related disclosure.[188] According to ISA 315, the inherent risk is first identified and assessed at the overall financial statement level.[189] In a second step, the auditor determines how inherent risks affect the assertion level of the individual classes of transactions, account balances and disclosures.[190] This approach supports the auditor in determining the relevant audit objectives for significant classes of transactions, account balances and disclosures identified.[191] The auditor further uses professional judgment to determine if any risk represents a significant risk that requires special audit consideration provided the combination of the likelihood with that a misstatement may occur and the magnitude of the potential misstatement should the misstatement occur.[192]

An inherent risk that is related to one or more assertions is referred to as "What could go wrong" in practice. Based on the understanding obtained on the entity's system of internal control, the auditor identifies whether management designed and implemented controls addressing the identified inherent risks by preventing, or detecting and correcting, related material misstatements. As part of assessing the design effectiveness of controls, the auditor usually performs a "walkthrough" through the process by selecting one or a few transactions and performing risk assessment procedures to follow the transaction through the process, i.e., to understand the initiation, recording, processing and reporting of the transaction and any supporting accounting relevant control activities.[193] In assessing the design effectiveness of relevant controls, the auditor performs a preliminary assessment of control risk.[194] If designed effectively, the auditor may determine to test the operating effectiveness of controls in order to reduce control risk and thus the extent of further substantive audit procedures required to keep audit risk at an acceptable low level.[195]

Inherent risk, control risk and the combined risk of material misstatement are assessed for each relevant assertion for each segment of the audit.[196] Based on the results of the

---

188 Cf. IFAC (2021), ISA 315 (Revised 2019), para 28.
189 Cf. ibid.
190 Cf. ibid.
191 Cf. ibid., para. 29.
192 Cf. ibid., para. A10; ELDER, RANDAL J. et al. (2020), p. 241.
193 Cf. ELDER, RANDAL J. et al. (2020), p. 376.
194 Cf. IFAC (2021), ISA 315 (Revised 2019), para 34.
195 Cf. IFAC (2021), ISA 330, para. 8.
196 Cf. ELDER, RANDAL J. et al. (2020), p. 253.

risk assessment, the auditor revisits the audit plan and overall audit strategy by determining the nature, timing and extent of further procedures (tests of controls and/or substantive procedures) necessary to obtain sufficient appropriate audit evidence for each relevant assertion.[197] Similarly, if a risk of material misstatement is identified on the financial statement level, the auditor is required to design an overall response at the financial statement level according to ISA 330.[198]

Addressing identified risks

The auditor is required to design and implement appropriate responses to the assessed risks of material misstatement in order to obtain sufficient appropriate audit evidence.[199] If material misstatements are identified, the planned audit procedures need to be adjusted. The continuous adjustment of the overall audit strategy and the audit plan corresponds to a stepwise approximation of the extent of audit procedures with the given (but unknown) entity-specific audit need.[200] Procedures to address identified risks of material misstatement may be distinguished into responses at the overall financial statement level and responses at the assertion level.[201] Risks of material misstatement at the financial statement level relate to the overall financial reporting and may potentially impact multiple audit objectives.[202] The auditor's overall responses to these risks include maintaining professional skepticism, integrating elements of unpredictability into the audit process and including experienced employees in the audit team.[203]

Audit procedures performed in response to the risks identified at the assertion level include:

- tests of controls,
- substantive analytical procedures,
- substantive tests of transactions and
- substantive tests of details of balances.[204]

[197] Cf. IFAC (2021), ISA 200, para. A38.
[198] Cf. IFAC (2021), ISA 330, para. 5.
[199] Cf. ibid., para. 3.
[200] Cf. WIESE, MICHAEL (2013), p. 20.
[201] Cf. IFAC (2021), ISA 330, para. 5 and 6.
[202] Cf. IFAC (2021), ISA 200, para. A38.
[203] Cf. IFAC (2021), ISA 330, para. A1.
[204] Cf. ibid., para. 4; Figure 7; The procedures are described in more detail in Chapter 2.3.4.

The risks of material misstatement at the assertion level of a class of transactions and related disclosures are addressed by tests of controls, substantive analytical procedures and substantive tests of transactions, while the risks of material misstatement at the assertion level of an account balance and related disclosures are addressed by substantive analytical procedures and tests of details of balances. As illustrated in Figure 7, tests of controls support the assessment of control risk while substantive analytical procedures and tests of details of balances are performed to address detection risk.[205] Substantive tests of transactions test the operating effectiveness of controls and the monetary amount of transactions and thus affect both control risk and detection risk.[206]

Irrespective of the effectiveness of internal controls, the auditor must perform substantive procedures for each material class of transactions, account balance and disclosure.[207] However, if controls for a segment of the financial statements are designed and operating effectively, the control risk decreases to a value below 100 percent and the planned acceptable detection risk may be increased while keeping audit risk at an acceptable low level. Consequently, the extent of further substantive procedures may be reduced. Figure 13 shows the relationship between tests of controls and substantive procedures at different effectiveness levels of internal controls.

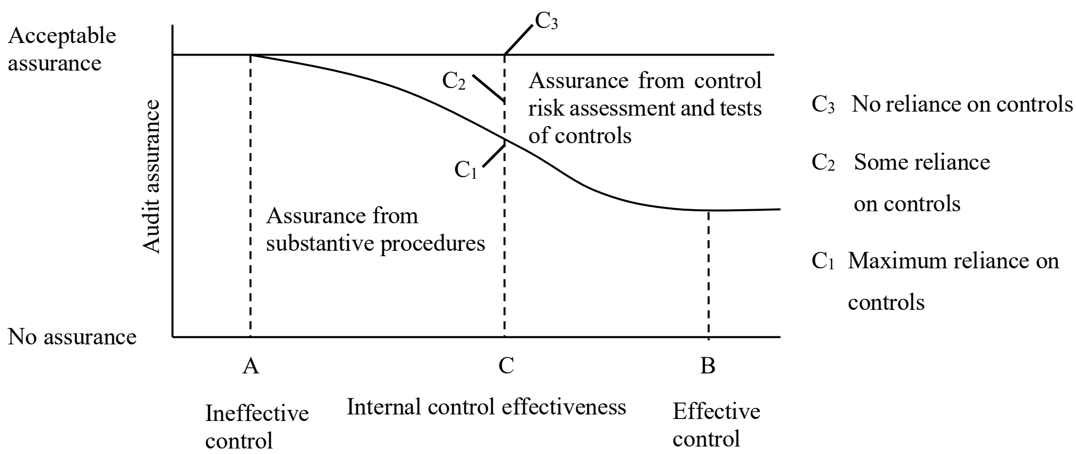


Figure 13: Audit assurance from substantive tests and tests of controls[208]

The area above the curve represents the maximum audit assurance that may be achieved through assessing the design and operating effectiveness of controls.[209] The control risk at the left of point A is 100 percent. In this scenario, following the audit

[205] Cf. ELDER, RANDAL J. et al. (2020), p. 375.
[206] Cf. ibid.
[207] Cf. IFAC (2021), ISA 330, para. 18.
[208] Cf. ELDER, RANDAL J. et al. (2020), p. 384.
[209] Cf. ibid., p. 383.

risk model, the risk of material misstatement is equal to the assessed inherent risk, as the auditor either assessed the controls as not designed effectively or determined to follow a substantive audit strategy.[210] At the right of point B, control risk may not be further reduced, because the risk may not be zero even if no control deficiencies are identified.[211] Even if controls are operating effectively, there is a risk that they may be overridden by management or circumvented through collusion between multiple individuals.[212] Assuming that the auditor's assessment of internal control is represented by point C, the greater the reliance on internal controls, the greater the extent of tests of controls required to support the assessment of control risk. For example, at point C_1, the auditor plans with maximum reliance on internal controls to reduce the extent of substantive procedures.[213] At point C_2, the auditor obtains C_3 - C_2 assurance from tests of controls and C - C_2 assurance from substantive testing.[214] However, if the auditor confirmed the design effectiveness of a control but the test of control reveals that the control is not operating effectively, the cost of the testing procedures represent sunk cost. In this scenario, in contrast to the auditor's initial expectation based on the assessment of the design effectiveness, the test of control does not reduce control risk, i.e., the risk is set to the maximum level C_3 and the extent of substantive procedures necessary to obtain sufficient appropriate audit evidence increases.[215]

Completing the audit and reporting

In the last phase of an audit, the auditor reviews the results of the audit, accumulates final audit evidence in performing procedures for completing the audit and issues the audit report. Procedures include, in particular:

- reassessing materiality,[216]
- performing an overall analytical review of the financial statements,[217]
- revisiting the risk assessment,[218]

---

210 For example, if the auditor assesses the risk of material misstatement as lower because of the characteristics of the significant class of transactions without considering internal controls, the auditor may determine that sufficient appropriate audit evidence may be obtained from substantive procedures alone, i.e., without performing tests of controls, cf. IFAC (2021), ISA 330, para. A10.
211 Cf. ELDER, RANDAL J. et al. (2020), p. 383.
212 Cf. IFAC (2021), ISA 200, para. A39; IFAC (2021), ISA 315 (Revised 2019), para. A47.
213 Cf. ELDER, RANDAL J. et al. (2020), p. 384.
214 Cf. ibid.
215 Cf. ibid.
216 Cf. IFAC (2021), ISA 320, para. 12 and 13; IFAC (2021), ISA 450, para. 10.
217 Cf. IFAC (2021), ISA 520, para. 3(b) and 6.
218 Cf. IFAC (2021), ISA 315 (Revised 2019), para. 7 and 37; IFAC (2021), ISA 330, para. 25 and 26.

- performing subsequent event procedures[219] and
- concluding the audit and issuing the audit opinion.[220]

Throughout the audit and in particular at its conclusion, the auditor assesses whether the amount determined for materiality, including the amounts allocated to individual segments of the audit, remains appropriate considering changes to factors or conditions about the entity or its environment.[221] If materiality needs to be revised, the auditor determines whether the nature and extent of procedures remains appropriate or if further audit procedures need to be performed.[222]

The auditor performs an overall analytical review of the financial statements near the end of the audit using analytical procedures[223] to identify balances, trends, relationships or other matters that are unusual or unexpected or inconsistent with the results of procedures performed during the audit. The overall analytical review is intended to corroborate conclusions formed during the audit of individual segments of the financial statements[224] and assists the auditor in identifying any risks of material misstatement not previously identified.[225]

ISA 315 and ISA 330 require the auditor to evaluate before concluding the audit whether the risk assessment remains appropriate and whether sufficient appropriate audit evidence to address the assessed risks has been obtained.[226] Changes to the risk assessment may be required, for example, if (a) new information is obtained that is inconsistent with the audit evidence underlying the risk assessment, (b) the auditor becomes aware of significant changes in the entity's system of internal control, (c) control deficiencies or exceptions in tests of controls have been identified or (d) the overall analytical review indicates a new risk of material misstatement.[227] If the risk assessment needs to be revised and/or the nature and extent of procedures performed is not sufficient, the auditor shall attempt to obtain further audit evidence and qualify

[219] Cf. IFAC (2021), ISA 560, para. 4.
[220] Cf. IFAC (2021), ISA 700 (Revised), para. 6.
[221] Cf. IFAC (2021), ISA 320, para. 12.
[222] Cf. ibid., para. 13.
[223] Cf. IFAC (2021), ISA 520, para. 3(b) and 6.
[224] Cf. ibid., para. A17.
[225] Cf. ibid., para. A18.
[226] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 7 and 37; IFAC (2021), ISA 330, para. 25 and 26.
[227] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 37 and A236; IFAC (2021), ISA 330, para. A60.

or disclaim the opinion on the financial statements in case sufficient appropriate audit evidence cannot be obtained.[228]

Many financial reporting frameworks explicitly refer to events that occur or facts that have been discovered after the date of the financial statements, as they may or may not need to be considered retrospectively in the financial statements.[229] In completing the audit, the auditor needs to obtain sufficient appropriate audit evidence that all events between the financial statement date and the date of the auditor's report that need to be considered in the financial statements have been identified.[230] These procedures are referred to as subsequent event procedures.

Finally, based on the conclusions drawn from the audit evidence obtained, the auditor (a) determines whether reasonable assurance has been obtained on whether the financial statements as a whole are free from material misstatement, (b) forms an opinion on the financial statements and (c) prepares and issues the audit report.[231]

The use of data analytics throughout the audit process

In the recent past, audit procedures summarized under the concept "(audit) data analytics" receive increasing attention in audit theory and audit practice. The AICPA defines audit data analytics as "*the science and art of discovering and analyzing patterns, identifying anomalies, and extracting other useful information in data underlying or related to the subject matter of an audit through analysis, modeling, and visualizing for the purpose of planning or performing the audit*"[232].

Data analytics may be used to support all types of audit procedures and by this are relevant in all phases of the audit.[233] An overview of the contemporary use of data analytics in the audit, including examples of specific procedures that may be supported in each phase of the audit, is provided by WERNER/WIESE/MAAS.[234] However, although data analyses are used to support all types of audit procedures, they are predominantly applied to financial data used to support analytical procedures or substantive tests.[235] The application of data analytics to analyze business processes,

228 Cf. IFAC (2021), ISA 330, para. 27.
229 Cf. IFAC (2021), ISA 560, para. 2.
230 Cf. ibid., para. 6.
231 Cf. IFAC (2021), ISA 700, para. 6 and 11. Further information on the requirements and components of the audit report are provided in ISA 700.
232 AICPA (2017), p. 2; BRYNES, PAUL et al. (2017), p. 5.
233 Cf. ELDER, RANDAL J. et al. (2020), p. 380; AICPA (2017), para. 1.01.
234 Cf. WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021), p. 7.
235 Cf. FINANCIAL REPORTING COUNCIL (2017), pp. 12f.

identify controls relevant to the audit and perform tests of controls is rather limited. If data analytics are applied in the context of the entity's system of internal control, they are usually limited to an isolated control that has already been identified, for example, by facilitating the understanding of the control activity through reperformance or by supporting the sample selection for the specific control.[236]

As process mining focuses on extracting process related data from an information system to reconstruct the process flow,[237] the technology may support all tasks that deal with analyzing business processes. Hence, process mining has the potential to expand the use of data analytics in the audit beyond the analysis of primarily financial data to the analysis of process data. The technology may especially support audit procedures performed at the assertion level of a significant class of transactions and related account balances and disclosures. This includes procedures performed in obtaining an understanding of a process, i.e., a significant class of transactions, as well as the identification and evaluation of integrated control activities performed in processing audit relevant information.

Consequently, the process mining technology is especially relevant in the risk identification and risk assessment phase according to ISA 315 and the phase of addressing the assessed risks of material misstatement according to ISA 330. In the following, process mining is theoretically embedded into the audit process by analyzing how the technology may contribute to meeting the regulatory requirements of relevant ISA.

[236] Cf. WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021), p. 5.
[237] Cf. Chapter 2.1.

2.3.3 Requirements for process mining to support risk identification and risk assessment

Figure 14 summarizes the auditor’s activities during risk identification and risk assessment that may be supported by process mining.

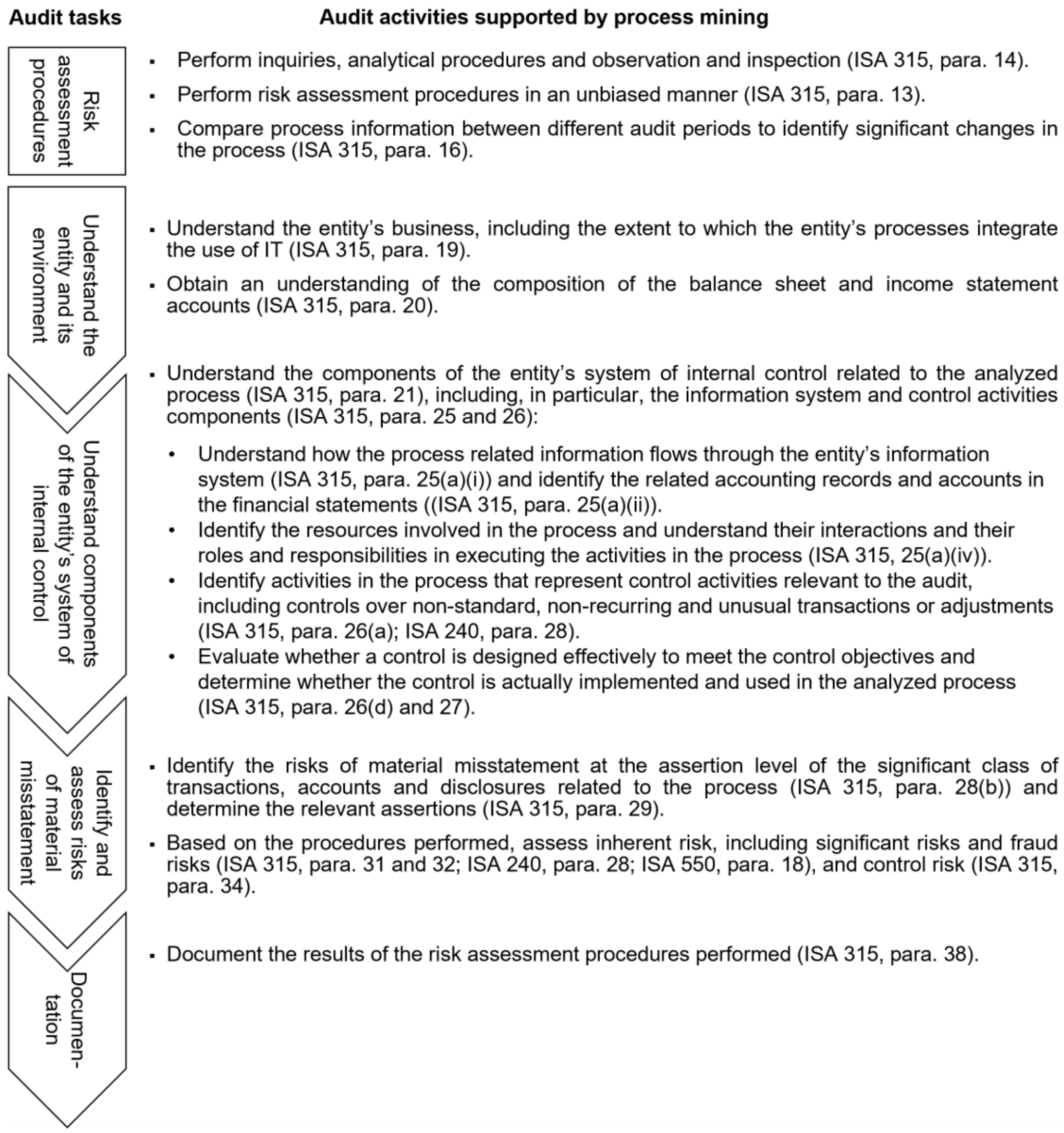


Figure 14: Risk assessment procedures that may be supported by process mining[238]

The following sections provide details on how process mining may support related activities for each audit task.

[238] Adapted from the initial version of this thesis’ author in WERNER, MICHAEL/WIESE, MICHAEL/ MAAS, ANNALOUISE (2021), p. 8.

Risk assessment procedures

Perform inquiries, analytical procedures and observation and inspection (ISA 315, para. 14)

Risk assessment procedures according to ISA 315 include inquiries of individuals within the entity, analytical procedures and observation and inspection[239] and may be supported by automated tools and techniques used on large volumes of data.[240]

Inquiry according to ISA 315 supports establishing the basis for the identification and assessment of risks.[241] In this phase of the audit, inquiries of management, those responsible for financial reporting and other appropriate employees within the entity are used to obtain an understanding of the entity and the different components of the entity's system of internal control in order to identify and assess risks. Despite being no inquiry procedure by nature, process mining can support directing inquiries to risk areas identified from analyzing the data or to aspects of the process that are not evident in the data. This includes manual process steps that may be relevant to the audit but are not included in the data extracted from the IT system.

Analytical procedures are "*evaluations of financial information through analysis of plausible relationships among both financial and non-financial data*"[242]. They include the investigation of identified trends, patterns or relationships that are inconsistent with other information obtained or that significantly differ from the auditor's expectation.[243] Analytical procedures are applied in all phases of the audit. As part of the risk assessment, they help to understand the entity's business, sensitize the auditor for aspects of the entity the auditor was not aware of and support identifying and assessing risks of material misstatement.[244] Analytical procedures may be performed by using automated tools and techniques and performing data analytics.[245] In providing examples for those analytical procedures, ISA 315 explicitly mentions applying visualization techniques to data extracted from the entity's information system.[246] In this regard, process mining can be interpreted as analytical procedure that supports understanding an entity's busi-

---

[239] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 14.
[240] Cf. ibid., para. A21 and A31.
[241] Cf. ibid., para. A22 and A23.
[242] IFAC (2021), ISA 520, para. 4.
[243] Cf. ibid.
[244] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A28; ELDER, RANDAL J. et al. (2020), p. 167.
[245] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A31.
[246] Cf. ibid.

ness process and identifying risks of material misstatement by helping to identify inconsistencies and unusual transactions or events in data relevant to the financial statements.[247] Exemplary analytical procedures supported by process mining include (a) the analysis of users and activities to understand roles and responsibilities in the process,[248] (b) the analysis of the sequence of activities based on their timestamp information and (c) the identification and analysis of process deviations that may give rise to a risk of material misstatement.

Observation and inspection may support understanding the entity's business and corroborate or contradict information obtained through inquiry.[249]

An observation according to ISA 500 includes looking at processes or procedures performed by other individuals, such as the performance of control activities.[250] ISA 315 concretizes that the auditor may be able, for example, to observe segregation of duties[251] and highlights that observation and inspection may be performed using automated tools and techniques.[252] In this regard, process mining may be interpreted as a "digital observation"[253] of user activities in the process, as it provides reliable information about the execution of a business process, including control activities and resources involved. As process mining may be performed using data from the entire audit period, the technology removes the inherent limitation of the audit evidence obtained through observation to the point in time the observation takes place.[254] In addition, an observation performed with process mining is not biased by the circumstance that individuals may act differently when they are aware of the auditor's observation. However, an observation with process mining is limited to those parts of the process that are supported by IT and covered by the event log.

Inspection includes the examination of documents and records, either physically or in electronic format.[255] When applied as a test of control, inspection may involve, for

[247] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A31 in combination with para. 27.
[248] Also see JANS, MIEKE/ALLES, MICHAEL/VASARHELYI, MIKLOS A. (2014), p. 1755.
[249] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A32.
[250] Cf. IFAC (2021), ISA 500, para. A21.
[251] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A33.
[252] Cf. ibid., para. A35.
[253] Also see JANS, MIEKE/ALLES, MICHAEL/VASARHELYI, MIKLOS A. (2013), p. 14. The authors compare process mining with a surveillance camera.
[254] Cf. IFAC (2021), ISA 500, para. A21.
[255] Cf. ibid., para. A18.

example, the examination of records for evidence of authorization.[256] In current process mining solutions, determining the process instance is equivalent to determining the document[257] whose processing is analyzed throughout the activities performed in the process. That is, process mining principally supports (digitally) inspecting the documents related to a case. For example, when analyzing the purchase to pay process, the activity "purchase order approved" that is representing the click of the approval button in the ERP system by a particular employee may be interpreted as an inspection of the electronic purchase order document for evidence of authorization.

Perform risk assessment procedures in an unbiased manner (ISA 315, para. 13)

ISA 315 requires the auditor to design and perform the risk assessment procedures in an unbiased manner.[258] As process mining reconstructs the process flow from the IT system's log files as it actually occurred, the technology has the potential to overcome the bias that may be related to the traditional means of obtaining an understanding of a significant class of transactions. For example, risk assessment procedures may include observations and inspection of the employee's behavior, such as the adherence to the segregation of incompatible duties.[259] However, as mentioned in the previous section, an employee might act differently if there is awareness that the performed activities are observed by the auditor. Similarly, inquiries of employees responsible for processing transactions that may support the auditor in determining if accounting policies are appropriately applied[260] are subject to the subjectivity of the respective individual. Instead of relying on observations, inquiries and the study of potentially outdated process manuals, the process flowchart derived with process mining is based on data that has been automatically recorded. This not only includes data entered by the auditee but additional meta information about the data entered into the system.[261] The process graph may corroborate or contradict the information obtained from inquiries and observations, helping the auditor in exercising professional skepticism.[262] By this, the process mining technology contributes to an unbiased identification and assessment of the risks of material misstatement related to the process.

---

[256] Cf. IFAC (2021), ISA 500, para. A18.
[257] Cf. JANS, MIEKE (2019), p. 61.
[258] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 13.
[259] Cf. ibid., para. A33 and A34.
[260] Cf. ibid., para. A23.
[261] Cf. Chapter 2.1.
[262] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 13 and A14.

Compare process information between different audit periods to identify significant changes in the process (ISA 315, para. 16)

Information relevant to the auditor's risk assessment may include the auditor's previous experience with the entity, including information from previous audit procedures performed.[263] To determine whether changes have occurred between both audit periods, the application materials of ISA 315 specifically include the walkthrough of relevant systems as appropriate audit procedure.[264] The timing of the audit procedures supported by process mining is determined directly by the particular time or period for which the data has been extracted. Consequently, changes to the prior period's process may be identified by comparing the information between the process mining analyzer of the previous audit period and the analyzer of the current audit period. By this, process mining may help the auditor in identifying significant changes of business practices, including changes in the entity's system of internal control.[265]

Understand the entity and its environment

Understand the entity's business, including the extent to which the entity's processes integrate the use of IT (ISA 315, para. 19)

The revised version of ISA 315 explicitly requires the auditor to understand the entity's business model and the extent to which the business model integrates the use of IT.[266] According to ISA 315, the auditor may use automated tools and techniques as part of the procedures to understand the use of IT in order to understand flows of transactions and their processing.[267] As a result of these procedures, the auditor may obtain information about the entity's organizational structure and business partners.[268]

Although current process mining solutions are limited to the analysis of a single business process only, elements of the analyzed process may contribute to understanding the overall business model and the extent of IT integration. For example, vendors (or customers or other external parties) included in the event log may help the auditor in analyzing the vendor structure and changes therein, either throughout the period or when compared with the previous period, and ultimately understanding with whom the

[263] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A39.
[264] Cf. ibid., para. A41.
[265] Cf. ibid., para. A40.
[266] Cf. ibid., para. 19(a)(i).
[267] Cf. ibid., para. A57.
[268] Cf. ibid.

entity conducts business. In understanding the purchase to pay process, price and material related information helps the auditor to understand the types of products or materials purchased and to identify vendors the entity is highly dependent of, for example, because they account for more than ten percent of the total invoice volume processed or because they are the only supplier for a material critical for the entity's production process. The degree of automation in the process (as evidenced, for example, by the length of the throughput times between activities) helps to identify automated flows of transactions. At the same time, any expected activities or aspects of the process that are not included in the discovered process model may point the auditor to manual parts of the business process that are relevant to the audit but not supported by IT (at least not by the information system from which the log files have been extracted). The user information in the event log may further support understanding the entity's organizational structure, including roles and responsibilities of users involved in the process.

Obtain an understanding of the composition of the balance sheet and income statement accounts (ISA 315, para. 20)

As process mining has not been developed for accounting and auditing in the first place, traditional process mining techniques do not consider financial data but focus on the process perspective, i.e., the case and event information stored in the event log.

However, as will be demonstrated in Chapter 4.2 of this thesis, process mining may be adjusted to consider the general ledger data in addition to the event data, including the balance sheet and income statement account movements. The financial data supports the auditor in understanding which financial accounts are involved in the process, including the nature and volume of the transactions related to cases in the process. By this, process mining may support meeting the ISA requirement of obtaining an understanding of the balance sheet and income statement accounts.[269]

[269] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 20.

Understand the components of the entity's system of internal control

Understand the components of the entity's system of internal control related to the analyzed process (ISA 315, para. 21), including, in particular, the information system and control activities components (ISA 315, para. 25 and 26)

ISA 315 requires the auditor to understand the entity's system of internal control.[270] Components of internal control include the control environment, the entity's risk assessment process, the entity's process to monitor the system of internal control, the information system and communication and control activities.[271]

Figure 15 illustrates the extent to which the individual components of internal control operate at the entity level and on the transaction or process level.

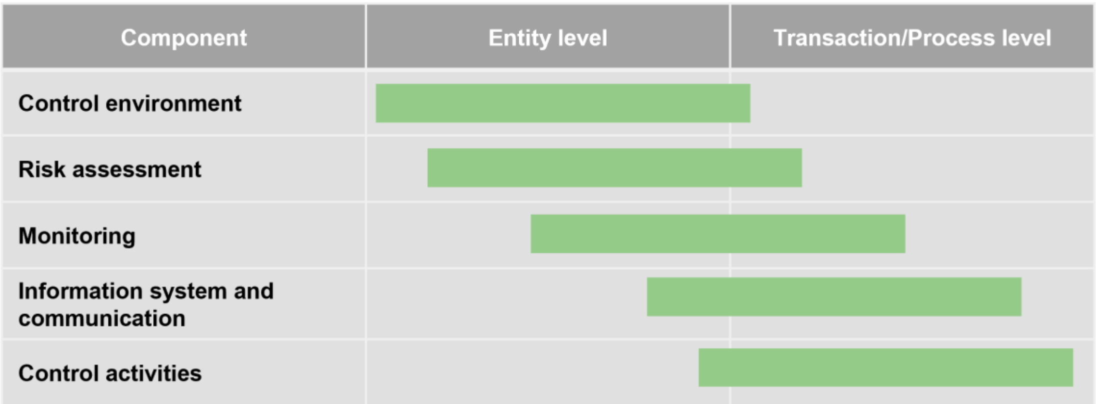


Figure 15: Organizational level of the COSO components of the system of internal control

The control environment, the entity's process to monitor the system of internal control and the entity's risk assessment process are components pervasive to the organization and established at the level of the entity. That is, they are not limited to specific classes of transactions, accounts, disclosures and related assertions and hence may not be analyzed sufficiently using data from a specific business process only. Process mining is particularly relevant to the information system and the control activities components as these are primarily established at the transaction level and affect the individual business processes. The following sections further explore how process mining may support related ISA requirements.

[270] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 21ff.
[271] Cf. ibid., para. 12(m). A definition and more detailed information of the individual components of an entity's system of internal control is provided by the COSO framework, cf. COSO (1994).

Understand how the process related information flows through the entity's information system (ISA 315, para. 25(a)(i)) and identify the related accounting records and accounts in the financial statements ((ISA 315, para. 25(a)(ii))

In its application materials, ISA 315 substantiates that the auditor may use automated tools and techniques to access or download the data stored in the entity's information system to understand the processing of transactions.[272] By reconstructing the process model from the data extracted from the entity's information system and visualizing how the process instances are processed through the system,[273] process mining inherently supports understanding how the process information flows through the system as required by ISA 315.[274] The process graph supports reviewing the process flow and the sequence of activities performed within the process. Process mining applications typically deconstruct the different ways the process is executed into individual process variations.[275] The auditor may determine if the major process paths represent typical workflows in the business of the entity and determine whether these align with the understanding of the process and any documentation obtained from the auditee. In reviewing individual process variations, the auditor may further identify non-routine process instances outside the normal course of the business[276] and determine if the sequence of activities gives rise to risks of material misstatement.

The flow of transactions according to ISA 315 involves the initiation, recording, processing and reporting of transactions in the general ledger and the entity's financial statements.[277] The first activity included in a case may provide information about the initiation of a transaction (for example, the purchase order creation). Process mining further includes processing activities (for example, the purchase order approval), correction activities (for example, the purchase order modification) and activities for recording transactions into the system (for example, the entering of an invoice).

Understanding the reporting of transactions in the general ledger requires information on the related accounting records and accounts in the financial statements.[278]

[272] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A137.
[273] Cf. Chapter 2.1.
[274] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 25(a)(i).
[275] Cf. Chapter 2.1.
[276] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A158; IFAC (2021), ISA 240, para 33(c).
[277] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 25(a)(i)a.
[278] Cf. ibid., para. 25(a)(ii).

Accounting records in an accounting system include a reference to the document relating to the transaction (for example, a purchase order or an invoice). As will be shown in Chapter 4.2 of this thesis, in using these references, a relationship between the process information and the accounts in the financial statements may be established. By this, the process data can be linked to the general ledger account movements enabling to analyze the reporting of transactions by the activities included in the process.

ISA 315 further states that the "*analysis of complete or large sets of transactions may also result in the identification of variations from the normal, or expected, processing procedures for these transactions, which may result in the identification of risks of material misstatement*"[279]. This is effectively achieved through the analysis of process variations[280] identified with process mining.

To support an initial understanding of the process, process mining applications usually provide an overview on the data, including a summary of the key information of the process, a visualization of the process flow reconstructed from the event log[281] and information on the nature and frequency of activities that are executed in the process. Appendix II introduces the "Overview" dashboard of the process mining application for purchase to pay from one of the Big Four audit firms that will be discussed in greater detail later in this thesis. The appendix further describes how the individual components of this dashboard support understanding the flow of the process related information through the entity's purchase to pay process extracted from the information system.

Identify the resources involved in the process and understand their interactions and their roles and responsibilities in executing the activities in the process (ISA 315, 25(a)(iv))

Figure 16 illustrates how process mining may support both identifying the resources involved in the process and understanding their roles and responsibilities.

[279] IFAC (2021), ISA 315 (Revised 2019), para. A137.
[280] Cf. Chapter 2.1.
[281] Cf. Table 1.

User ID	Events	%
IKM	42,967	31.1%
GMM	37,380	27.05%
JEN	27,253	19.72%
DHM	6,708	4.86%
JGP	5,937	4.3%
BWW	5,304	3.84%
NMG	4,239	3.07%
GNO	3,272	2.37%
EDL	2,652	1.92%
CMM	2,453	1.78%
Total	138,165	100%

Activity type	▼ Events	▼ %
Invoice settlement	40,532	29.34%
Invoice posting	34,661	25.09%
Payment posting	33,160	24%
Goods receipt posting	13,096	9.48%
Invoice creation	8,638	6.25%
Payment block modification	4,215	3.05%
Purchase order approval	3,082	2.23%
Purchase order creation	621	0.45%
Purchase order rejection	76	0.06%
Purchase order modification	54	0.04%
Invoice modification	30	0.02%
Total	138,165	100%

Figure 16: Identifying the users involved in the process and understanding their roles and responsibilities

Using the information on the users executing activities that is stored in the event log, the table at the top of Figure 16 lists the absolute and relative number of events processed by user. Analyzing the distributions of events or cases among users supports the auditor in understanding which users are involved in the process and identifying key users with a high degree of involvement. The table at the bottom of Figure 16 shows the distribution of events over the activities in the process. Both tables may be used in conjunction to understand roles and responsibilities in executing activities in the process. The auditor may select a specific user from the table at the top and the analysis updates to only display this user's activities. Similarly, when selecting a specific activity from the table at the bottom, the process data is filtered to only show the users who performed this activity.

VAN DER AALST demonstrates how techniques from social network analysis may be integrated in process mining to understand the interactions between resources, identify key resources in the process and analyze work handovers between users or departments.[282]

The information on the entity's resources obtained with process mining may further be used to corroborate other information obtained during the audit, for example, a list of user accounts in the ERP system, organizational charts and hierarchies for individual functions or departments executing relevant control activities, or job descriptions of key personnel.[283]

Identify activities in the process that represent control activities relevant to the audit, including controls over non-standard, non-recurring and unusual transactions or adjustments (ISA 315, para. 26(a); ISA 240, para. 28)

The activities in the event log and the understanding of the flow of transactions, the vendor or customer structure and the resources involved in the process support the auditor in identifying those activities that represent control activities relevant to financial reporting. As process mining is based on the population of transactions within the significant class of transactions, the technology supports identifying controls for routine transactions as well as non-standard, non-recurring or unusual transactions or adjustments.[284]

---

282 Cf. VAN DER AALST, WIL M. P. (2016), pp. 282ff.

283 Information on the frequency of activities, throughput times and reversals or repetitions of activities may further help assessing the skills and competence of users involved in the process.

284 Cf. IFAC (2021), ISA 315 (Revised 2019), para. 26(a)(ii).

For example, when analyzing the activities included in the purchase to pay process (Figure 17), the auditor may identify the control activity “purchase order approved”.

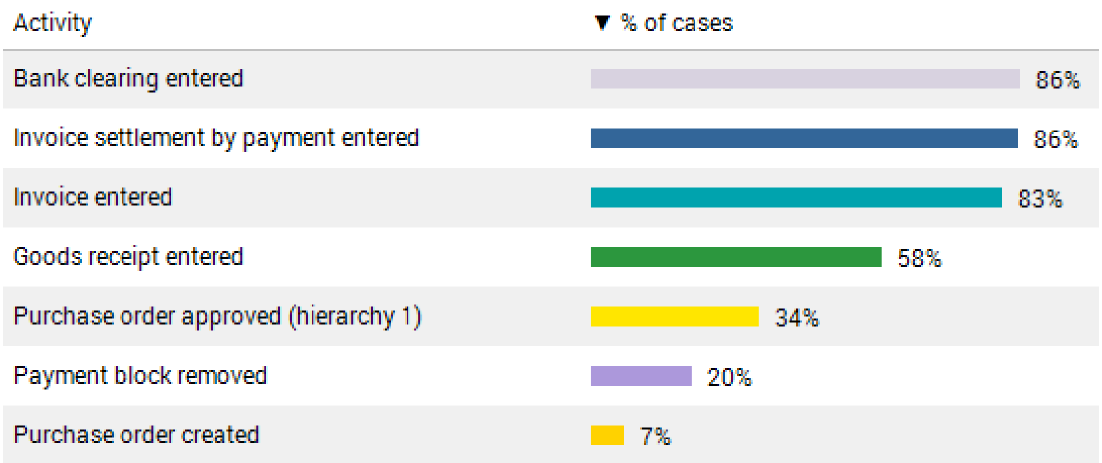


Figure 17: Excerpt of activities within the purchase to pay process

Similarly, the entity may make use of a business approval for invoices and the auditor may identify the approval of invoice modifications by an authorized individual as control activity relevant to financial reporting, as it addresses the risk that invoice adjustments for purchases goods or services are not properly recorded.

Process mining does not only support identifying the control activities but also their characteristics, represented by the individual events and their attributes in the event log. For example, according to the definition of the “purchase order approved” activity, the timestamp of the related events might reflect the date and time when an individual clicked the related approval button in the ERP system. Consequently, the control activity is an IT dependent manual control. Meta data logged by the system[285] further supports evaluating the timing and frequency of the control’s execution. For example, by analyzing the distribution of related events over time, the auditor may confirm that the control activity is performed many times per day. By considering the related purchase order amount and comparing the users performing the approval activity[286] with the entity’s hierarchy of authorization, the auditor may determine if the approval activity is performed by an authorized individual.

Using the data from the event log, additional analyses may be set up for control activities that are not represented by an activity in the event log but supported by the information system and the attributes of cases and events in the event log. For example,

285 Cf. Chapter 2.1.
286 Cf. Figure 16.

control activities relevant to financial reporting usually involve assessing whether appropriate segregation of duties that are incompatible in nature is maintained over the course of the audit period.[287] Management's approval of rights to perform duties might permit an individual to authorize a purchase order and authorize the receipt of the goods, the invoice and the payment of the purchase. This lack of segregation of duties may lead to an increased risk that fraudulent transactions may occur or that errors may be made in processing. Figure 18 illustrates how process mining may support analyzing segregation of duties.

Users with cases not in line with the four-eyes principle

Four eyes activities: Invoice entered, Purchase order approved

User ID	Cases	%
GNO	5	100%
Total	5	100%

Cases not in line with the four-eyes principle

Purchasing vendor | Material group | + | PO amount

Purchasing vendor	Material group	PO amount	%
EALP	OLC11	312,744	64%
BUJQL	DVW11	179,165	36%
Total		491,909	100%

Figure 18: Using process mining to analyze the segregation of incompatible duties

In the example provided, the activities "purchase order approved" and "invoice entered" have been determined to be incompatible in nature. The table at the top shows there are five cases where the same user "GNO" approved the purchase order and entered the related invoice. The table at the bottom enables to further analyze the characteristics of the five cases, and determine, for example, whether they are concentrating on a specific vendor or material group and the total purchase order and invoice amount involved. As smaller entities may frequently lack the resources to maintain an appropriate segregation of duties, management may have implemented alternative control activities that mitigate related risks of material misstatement.[288] Based on the results of the segregation of duties analysis, the auditor may filter the entire process

[287] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A33 and Appendix 3, para. 20.
[288] Cf. ibid., Appendix 3, para. 20.

mining application to those cases where expected segregation of duties is not warranted and identify if the cases' process flow gives rise to a risk of material misstatement.

Further examples of controls that may be relevant in the purchase to pay process and may be identified with process mining include:

- the match of quantity and/or price information between the purchase order, goods receipt and invoice documents (three-way-match),
- approval activities in case of a mismatch (for example, payment block activities),
- the automatic use of the goods received/invoice received (GR/IR) account (especially in SAP),
- approvals for documents (for example, purchase order or invoice approvals),
- the validation and approval of vendor master data changes,
- the prevention of duplicate payments by requiring an invoice prior to payment
- and the automatic posting of bank clearing after the payment run is initiated.

Evaluate whether a control is designed effectively to meet the control objectives and determine whether the control is actually implemented and used in the analyzed process (ISA 315, para. 26(d) and 27)

ISA 315 requires the auditor to perform risk assessment procedures to obtain audit evidence about the design and implementation of identified controls.[289] Evaluating the design effectiveness involves considering if the control, individually or combined with other controls, may effectively prevent, or detect and correct, material misstatements.[290] The implementation of a control may be confirmed by determining that the control activity exists and the entity is using it.[291] To obtain audit evidence on the design effectiveness and the implementation of identified controls, the auditor traditionally performs a walkthrough by following the processing of an individual transaction along the critical path of the process using inquiry, observation and inspection

[289] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 26(d) and A177.
[290] Cf. ibid., para. A175.
[291] Cf. ibid., para. A176.

procedures.[292] With process mining, the procedures to identify control activities relevant to the audit, to determine if they are actually implemented and to evaluate if they are designed effectively to meet the control objectives may be performed concurrently.

If the execution of a control activity is logged by the information system and included in the event log (and thus, in the list of activities in the process mining analyzer presented in Figure 17), the auditor may conclude that the control is existing and has been performed for at least one transaction. As such, using the information on the percentage of cases for which a control activity has been performed,[293] the auditor may not only identify the control activity but also determine that the entity actually makes use of the control, i.e., the control is implemented.

The lines between evaluating the design of a control and testing its operating effectiveness[294] may blur with process mining.

For example, in evaluating if an approval control is designed effectively to prevent, or detect and correct, a material misstatement, the auditor may filter the cases using the process graph in order to determine if there are any cases for which the approval has not been performed. If the approval activity has been performed for all cases that the auditor expects to run through the control activity, the auditor may determine that the control is designed effectively. Similarly, if the auditor identifies that management established a control for the segregation of the "purchase order creation" and "purchase order approval" activity, and no cases are identified that violate the segregation between these duties, the auditor may conclude that the control is designed effectively.[295]

However, as process mining is based on the population of transactions rather than an individual process instance selected for walkthrough procedures, the auditor may likely encounter exceptions, i.e., cases that are expected to be subject to a control but actually did not pass through the control. In the example provided in Figure 18 on the segregation of duties between the purchase order approval and the invoice entry, five cases are identified for which the segregation of duties is not warranted. Similarly, there might be cases for which an approval of the purchase order is not performed. These exceptions will presumably not cause the auditor to conclude that the control is

[292] Also see MOCHTY, LUDWIG (2015), pp. 34f.
[293] Cf. Figure 17.
[294] Tests of controls for assessing the operating effectiveness of a control are discussed in Chapter 2.3.4.
[295] In these scenarios where no exceptions are noted the auditor might further conclude that the control is operating effectively, cf. IFAC (2021), ISA 315 (Revised 2019), para. A19; Chapter 2.3.4.

not designed effectively, especially when considering that the traditional evaluation of the design effectiveness is based on the walkthrough using a single process instance only. Instead, the auditor may investigate the characteristics of these exceptions. For example, an obvious reason for a missing purchase order approval may be that the purchase order has just been created and is not further processed, or that the purchase order has been rejected. The five cases for which the segregation of duties is not warranted might have run through a mitigating control activity, for example, an invoice approval by another authorized individual. The auditor might also determine that the invoice amount related to these cases is not material and conclude that there is no remaining risk of material misstatement.

The examples demonstrate that process mining may support both testing the design and operation of a control, however, the distinction between evaluating the design and the operating effectiveness may blur. The auditor traditionally determines the design effectiveness of controls before testing its operating effectiveness because the cost of tests of controls are sunk cost if the control is not effective.[296] However, as testing the operating effectiveness of controls with process mining does not involve selecting a sample of transactions for which the documentation needs to be requested from the auditee and evaluated by the auditor, the cost of tests of controls are reduced. Consequently, and in accordance with ISA 315[297] and ISA 330[298], it may be feasible for the auditor to jointly test the design and operation of a control.

Identify and assess risks of material misstatement

Identify the risks of material misstatement at the assertion level (ISA 315, para. 28(b)) and assess inherent risk, including significant risks and fraud risks (ISA 315, para. 31 and 32; ISA 240, para. 28; ISA 550, para. 18), and control risk (ISA 315, para. 34)

Irrespective of whether process mining is used as analytical procedure, observation or inspection, the information obtained on the key elements of the process supports the auditor in identifying and assessing risks of material misstatement for relevant assertions of significant classes of transactions, account balances and disclosures related to the analyzed process.[299]

---

[296] Cf. Figure 13.
[297] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A19.
[298] Cf. IFAC (2021), ISA 330, para. A22.
[299] Also see WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021), p. 9.

For example, in analyzing the different ways the process is executed to understand the flows of transactions, the auditor may consider the number of cases processed through an individual variation compared to the total number of cases. A high frequency indicates a routine variation subject to different risks than non-routine variations. The total invoice amount processed through an individual variation compared to the number of cases may help the auditor to identify non-routine processing involving a higher risk, for example, high value transactions being processed separate to the routine transactions. Risks of material misstatement may also be identified when analyzing the activity sequence within the process, for example, when a case in purchase to pay starts with a payment activity before the receipt on an invoice. Information on the number and invoice volume of cases following a specific process path may further support the auditor in assessing the likelihood and magnitude of potential misstatements and determining whether a risk is a significant risk.[300]

While obtaining an understanding of the key elements of the process, including relevant internal controls, the auditor continuously updates the documentation on the risks of material misstatement and evaluates the impact on the risk assessment. This is to ensure that appropriate audit procedures are designed to address the assessed risks.

Documentation

Document the results of the risk assessment procedures performed (ISA 315, para. 38)

Automated tools and techniques used in an audit of financial statements need to support incorporating the results of the procedures performed into the audit documentation. As of today, to the author's best knowledge, no process mining solution supports documenting the audit procedures performed and the conclusions drawn directly within the application. However, most process mining solutions available on the market support exporting the analyses in a file format that may be integrated into the auditor's documentation of the risk assessment.

2.3.4 Requirements for process mining to support addressing identified risks

Procedures to address identified risks of material misstatement may be distinguished into responses at the overall financial statement level and responses at the assertion level. Risks of material misstatement at the overall financial statement level relate to

[300] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 32.

multiple audit objectives.[301] As process mining focuses on a particular business process, it is particularly relevant to the auditor's procedures performed in response to the risks identified at the assertion level of individual classes of transactions, account balances and related disclosures. Figure 19 summarizes the audit procedures responsive to the assessed risks of material misstatement that may be supported by process mining.

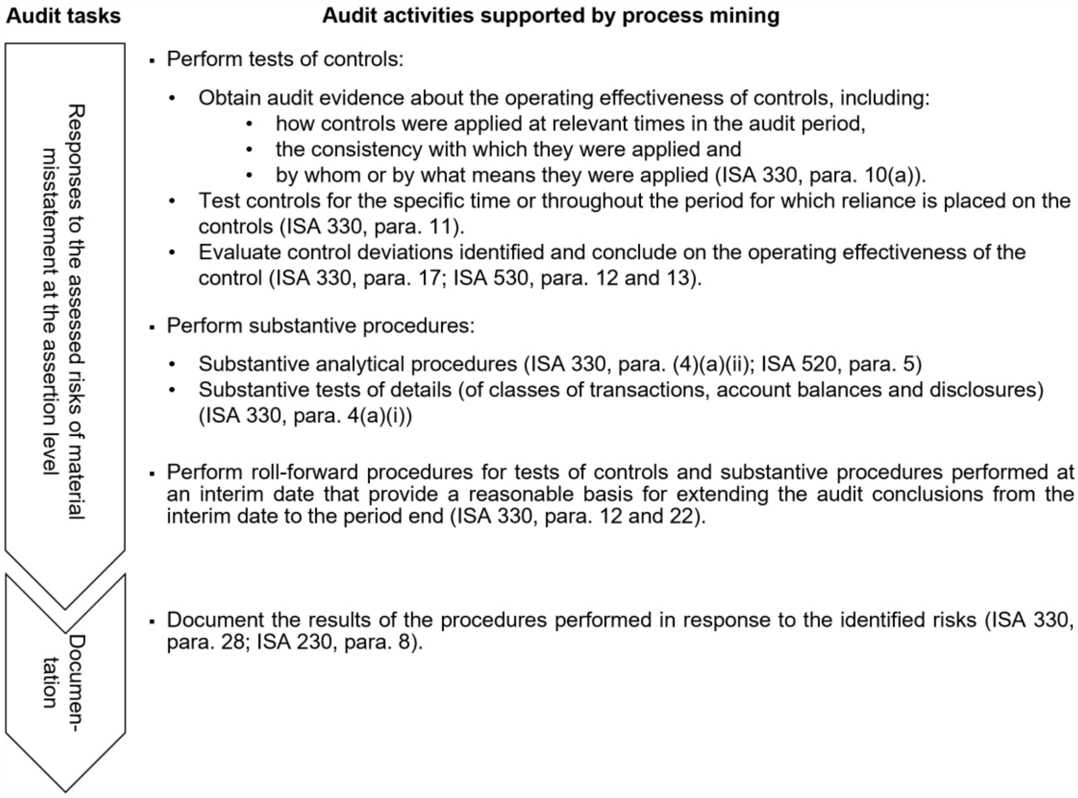


Figure 19: Procedures responsive to the assessed risks that may be supported by process mining[302]

The following sections provide details on how process mining may support related activities for each audit task.

[301] Cf. Chapter 2.3.2.

[302] Adapted from the initial version of this thesis' author in WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021), p. 8.

Perform tests of controls

Obtain audit evidence about the operating effectiveness of controls, including (i) how controls were applied at relevant times in the audit period, (ii) the consistency with which they were applied and (iii) by whom or by what means they were applied (ISA 330, para. 10(a))

If a control relevant to financial reporting is designed effectively, the auditor may test the control to determine if the control is operating effectively to prevent, or detect and correct, a material misstatement.[303]

Tests of controls consist of inquiry in combination with other procedures that, depending on the nature of the control, may involve the use of CAATs.[304] To a large extent, the evidence on the operation of a control that may be obtained with process mining does not differ from the information obtained in today's audits. For example, as part of understanding the process and performing walkthrough procedures in a traditional audit, the auditor may have identified the control activity that a purchase order is signed by an authorized individual prior to release to the vendor. When testing the operating effectiveness of the control activity, the auditor may determine to select a sample of purchasing transactions throughout the audit period to determine if the related purchase order includes a signature from an authorized individual. As organizations continue to digitize, the related purchase order document will presumably no longer be a physical document with a signature that is reviewed by the auditor and checked against the hierarchy of authorization. Instead, the auditor will determine the control activity to be IT-dependent manual, as the approval is performed electronically within the information system. As a result, the evidence of the test of control will include screenshots or printouts from the ERP system, evidencing the approval of the electronic purchase order document in the system.[305] With process mining, the evidence of the approval of the electronic purchase order document is represented by the "purchase order approved" event. The timestamp information enables to determine if the approval took place prior to the release of the purchase order, while the user performing the event may be compared against the hierarchy of authorization. As process mining uses the ERP log files as input, the conclusion on the control's operation

---

303 Cf. IFAC (2021), ISA 330, para. 8.
304 Cf. ibid., para. A27.
305 Also see JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), p. 19.

is using the same information as today – but based on the population of transactions instead of a sample.

As the same types of audit procedures are performed in evaluating the design and operating effectiveness of controls, ISA 330 acknowledges that tests of controls may be performed concurrently with the assessment of the design effectiveness.[306] As discussed in the previous chapter, this may especially apply to procedures performed with process mining, where the identification and evaluation of controls is based on the population of transactions rather than an individual transaction (walkthrough procedures) or a sample of transactions (test of control) only.

According to ISA 330, testing a control should provide information on:

- how the control was applied throughout the audit period,
- the consistency with which it was applied and
- by whom or by what means it was applied.[307]

To a large extent, the auditor will already have obtained this information when using process mining to identify the control and determine its implementation and design effectiveness.

For example, using process mining to identify if segregation of duties was maintained[308] simultaneously involves determining how the control was applied, i.e., by ensuring that the respective incompatible activities are executed by different authorized individuals. The distribution of related events over time helps to determine the consistency with which the control was applied throughout the audit period. The auditor may further filter to cases that are expected to pass through a particular control activity to determine if the control has been performed consistently. If the event log includes the events' start and end time, the auditor may determine the duration of control activities as an additional indicator for its consistency over time. Information on the individual performing the control is provided by the user information in the event log.[309] The unique technical definition of each activity, together with the additional event attributes and meta-data recorded for each event (for example, the timestamp of the activity and the document fields affected by the activity) provide further information on the means the control activity has been applied with. For example, the user

[306] Cf. IFAC (2021), ISA 330, para. 10(a), A21 and A22.
[307] Cf. ibid., para. 10(a).
[308] Cf. Figure 18.
[309] Cf. Figure 16.

data and information on the duration of activities may help the auditor to evaluate automated control activities that are expected to be performed by a specific (batch) user within little or no throughput time.

Test controls for the specific time or throughout the period for which reliance is placed on the controls (ISA 330, para. 11)

Controls are usually tested using a random sample of transactions subject to the control. If tests of controls are performed at an interim date, they are updated as of period end to extend the reliance to the period between the initial tests of controls and the balance sheet date.[310] With process mining, the timing of tests of controls is determined directly by the particular period for which the data is extracted from the system and loaded into the process mining application. By this, process mining supports obtaining evidence on a control's effectiveness based on the entire period the auditor intends to rely on the control. The timestamp information included in the event log further helps to identify if substantially different controls are used at different times throughout the period under audit that would need to be considered separately by the auditor.[311]

Evaluate control deviations identified and conclude on the operating effectiveness of the control (ISA 330, para. 17; ISA 530, para. 12 and 13)

Control deviations result if a control is not operating as designed. For example, if a purchase order meets qualitative or quantitative approval criteria and thus, according to the control design, should have been approved by an authorized individual prior to release to the vendor but it has not been approved, this represents a control deviation. While a control might still operate effectively if some control deviations occur,[312] it is the auditor's responsibility to determine if the deviation causes the test of control to no longer provide an appropriate basis for the operating effectiveness of the control.[313] In this case, additional tests of controls that might mitigate the related risk of material misstatement (mitigating controls) or, alternatively, substantive procedures are required to address any related risks of material misstatement.[314] In the example provided, an approval of the invoice received from the vendor prior to payment might mitigate the risk resulting from the unauthorized purchase order. With process mining,

[310] Cf. IFAC (2021), ISA 330, para. 12.
[311] Cf. ibid., para. A20.
[312] Cf. ibid., para. 12.
[313] Cf. ibid., para. 17.
[314] Cf. ibid.

the auditor may filter the deviations of the purchase order approval control accordingly to identify and test mitigating control activities along the flow of transactions.

The evaluation of whether a deviation causes the test of control to no longer provide an appropriate basis for reliance on the control is performed based on specific inquiries to understand the nature and cause of the deviation and its potential impact.[315] In analyzing control deviations, the auditor may identify that the deviation represents an anomaly that is not representative for the population,[316] or determine that the deviation is systematic, as the control exceptions have a common feature and need to be addressed with additional audit procedures.[317] Rather than relying on inquiries, process mining supports evaluating control deviations based on the information recorded in the entity's information system. The case and event attributes help to identify common characteristics of cases not passing through a specific control activity and support assessing if an exception is systematic. For example, using the timestamp information of the event log or the information on the vendor related to a case, respectively, the auditor may identify that a control activity has not been performed during a particular period or for cases related to a specific vendor.

To design an audit sample for a test of control, the auditor assesses the expected rate of deviation,[318] as the sample size increases with the number of expected control deviations[319] and the comparison of the expected deviation rate with the detected rate may indicate if the auditor can rely on the control to reduce the risk of material misstatement at the assertion level to the level assessed by the auditor.[320] By determining the nature and volume of control deviations based on data, process mining supports a more robust assessment of control deficiencies. As the technology considers the population of transactions instead of a sample, the auditor may further determine if control exceptions exceed acceptable tolerances, resulting in the failure of the control to prevent, or detect and correct, a material misstatement. When identifying control exceptions, the auditor does not extend the sample size but assesses whether the monetary amount of the affected transactions is material to the financial statements. By this, process mining may directly show if a control is sufficient to reduce the risk

---

[315] Cf. IFAC (2021), ISA 330, para. 17; IFAC (2021), ISA 530, para. 12.
[316] Cf. ISA 530, para. 5(e) and 13.
[317] Cf. ibid., para. A17.
[318] Cf. ibid., para. A7.
[319] Cf. IFAC (2021), ISA 330, para. 46.
[320] Cf. ibid., para. A41.

of material misstatement at the assertion level to the level assessed by the auditor. As testing controls with process mining removes the need for a sample-based approach, a presumption about the exception rate becomes obsolete. In general, the limitations inherent to sampling techniques no longer apply, including the sampling risk[321] defined in ISA 530, i.e., the risk that the audit conclusion achieved based on a sample differs from the conclusion that is drawn based on the entire population.

Perform substantive procedures

If the auditor determines that relevant controls are operating effectively, the transactions that pass through the expected control activities in the process may be tested following a "rely on controls" audit strategy. Consequently, the acceptable remaining detection risk is increased and the extent of substantive testing for these transactions can be reduced significantly.[322] Within process mining, the auditor may filter, summarize and evaluate all cases for which required control activities have not been performed. Based on this summary, the auditor may determine the timing and extent of further substantive testing. For the population of remaining transactions, the base sample size can be reduced as materiality does not change. For example, if 90 percent of the case amount of transactions pass through the expected control activities, the base sample size is reduced to 10 percent compared to the sample size needed without controls testing. In these situations, the auditor may determine that a combination of analyses performed with process mining and corroborating analytical procedures will be sufficient to hold audit risk at five percent. Depending on the evaluation of control deviations, or in case the auditor determined to not rely on controls, the auditor may select specific cases for further investigation and perform additional substantive procedures.

ISA 330 acknowledges that the use of CAATs to perform procedures in response to assessed risks enables a more extensive testing of electronic transactions and accounting files and may support to select a sample of transactions, to sort specific transactions or to test a population instead of a sample only.[323]

[321] Cf. IFAC (2021), ISA 530, para. 5(c).
[322] Cf. Figure 13.
[323] Cf. IFAC (2021), ISA 330, para. A16.

Perform substantive analytical procedures (ISA 330, para. (4)(a)(ii); ISA 520, para. 5)

Analytical procedures may be used as substantive tests to obtain audit evidence about particular assertions related to account balances or classes of transactions.

Data analytics are widely used in the audit practice to support analytical procedures,[324] be it for the application during risk assessment[325] or when used as substantive analytical procedures for risk response. Used as substantive analytical procedure, they support analyzing large datasets in a number of disaggregated ways, usually providing more detailed insights from the procedures performed than analytical procedures used as risk assessment procedures.[326] Auditors frequently use visualization software to identify and evaluate unusual fluctuations[327] and to communicate findings from substantive analytical procedures.[328] In analyzing the definition of analytical procedures, JANS/ALLES/VASARHELYI conclude that process mining expands the evidential domain for analytical procedures from just the outcome of a business process (transactions posted as journal entries to account balances) to the business process that gives rise to these transactions.[329]

The use of process mining as a substantive analytical procedure in line with ISA 520[330] may include, for example:

- connecting the financial information of transactions to process-related nonfinancial information,
- comparing the development of key performance indicators (KPIs) of the process[331] and activities over time against the auditor's expected development,
- performing consistency checks, for example, on the timing of activities and their sequence,
- analyzing the resulting process variations and comparing individual variations to identify unusual transaction processing and

[324] Cf. ELDER, RANDAL J. et al. (2020), p. 380.
[325] Cf. Chapter 2.3.4.
[326] Cf. ELDER, RANDAL J. et al. (2020), p. 380.
[327] Cf. IFAC (2021), ISA 520, para. 4.
[328] Cf. ELDER, RANDAL J. et al. (2020), p. 380.
[329] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 10.
[330] Cf. IFAC (2021), ISA 520, para. 4.
[331] Cf. Figure 80.

- performing predictive analyses, for example, on the expected paths of cases based on their characteristics.

Perform substantive tests of details of classes of transactions, account balances and disclosures (ISA 330, para. 4(a)(i))

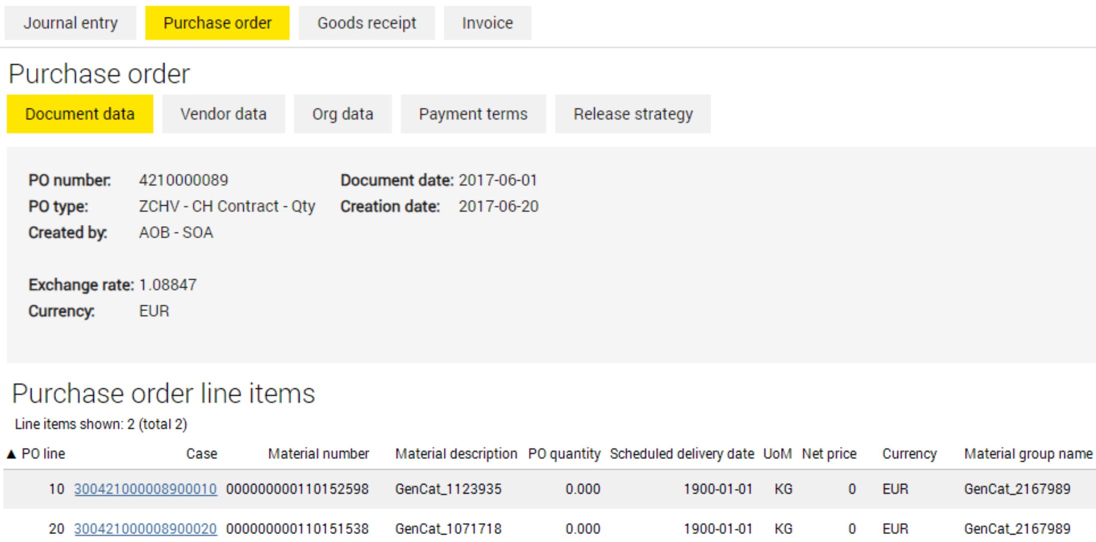
Substantive tests of details may be differentiated into tests of details of account balances and substantive tests of transactions.[332] Tests of details of balances are performed to test for monetary misstatements in the balances of the financial statements and address the risks of material misstatement identified at the assertion level of account balances and related disclosures.[333] Substantive tests of transactions are performed to test transactions and related disclosures for monetary misstatements and address the risks of material misstatement identified at the assertion level of classes of transactions and events and related disclosures.[334]

Audit evidence for substantive tests of details is derived from inspecting documents related to the transaction, inquiring responsible personnel, reperforming a related control activity or recalculating monetary amounts associated with the transaction to verify its mathematical accuracy.[335] As inspection and reperformance are involved in both tests of controls and substantive tests of transactions, both audit procedures are frequently performed at the same time, referred to as dual purpose test.[336] For example, while verifying the evidence of approval by an authorized individual on a document related to a transaction, the auditor may simultaneously determine the monetary correctness of the document and the related journal entry.[337]

As illustrated in Figure 5, the event log data includes information on the documents related to a particular case. The document information may be enhanced with detailed line item information for each document extracted from the system. Figure 20 shows an example of how the process mining input data for purchase to pay may be transformed and made available to the auditor for analysis.

---

[332] Cf. Chapter 2.3.1.
[333] Cf. ELDER, RANDAL J. et al. (2020), p. 379.
[334] Cf. ibid., p. 377.
[335] Cf. ibid., p. 381.
[336] Cf. ibid., p. 378; IFAC (2021), ISA 330, para. A23.
[337] Cf. IFAC (2021), ISA 330, para. A23.



Journal entry | Purchase order | Goods receipt | Invoice

Purchase order

Document data | Vendor data | Org data | Payment terms | Release strategy

PO number: 4210000089
PO type: ZCHV - CH Contract - Qty
Created by: AOB - SOA
Document date: 2017-06-01
Creation date: 2017-06-20

Exchange rate: 1.08847
Currency: EUR

Purchase order line items

Line items shown: 2 (total 2)

▲ PO line	Case	Material number	Material description	PO quantity	Scheduled delivery date	UoM	Net price	Currency	Material group name
10	300421000008900010	000000000110152598	GenCat_1123935	0.000	1900-01-01	KG	0	EUR	GenCat_2167989
20	300421000008900020	000000000110151538	GenCat_1071718	0.000	1900-01-01	KG	0	EUR	GenCat_2167989

Figure 20: Document details of an individual case in process mining

The figure shows a particular purchase order document of the process, including the header information, the line item information and a reference to the specific process instance related to each purchase order line item. Similar information may be included for the goods receipts and invoices in the process as well as the accounting documents, i.e., the individual journal entries related to a case.

By this, process mining may be used to select specific cases and supplement related transactions with supporting evidence (for example, the purchase orders for invoices recorded). Selected transactions may include both cases selected as key items as part of other analyses performed with process mining or sampled items. In this scenario, process mining supports performing substantive tests of details. The testing results are reflected in the calculation of the acceptable remaining detection risk through an increased key item coverage or the calculated remaining representative sample size.

Perform roll-forward procedures

Perform roll-forward procedures for tests of controls and substantive procedures performed at an interim date that provide a reasonable basis for extending the audit conclusions from the interim date to the period end (ISA 330, para. 12 and 22)

If tests of controls or substantive procedures have been performed at an interim date throughout the audit period, ISA 330 requires obtaining additional audit evidence to extend the conclusions drawn at the interim date to the balance sheet date.[338] Additional evidence for the operating effectiveness of controls over the period between the interim date and the balance sheet date is usually obtained by extending tests of controls over the remaining period.[339] The auditor may further determine that it is effective to perform substantive procedures at an interim date and perform substantive analytical procedures or tests of details for the roll-forward period to identify significant unusual transactions in the roll-forward period, other amounts that appear unusual or changes (or the absence of expected changes) in the composition of a class of transactions or account balance.[340]

The period for which the data is extracted from the information system and transformed and loaded into the process mining application can be flexibly defined by the auditor. As such, the auditor may determine to set-up a process mining application with the data available at the interim date and then compare the data with the process mining analyzer as of period end. Some process mining solutions further support saving the analyses performed, including any filters applied to the data, in order to repeat them at a subsequent date.[341] By this, tests of controls and substantive procedures may be efficiently re-performed on an extended data population and controls and process variations that have been evaluated at an interim date may be compared to the period between the interim and the balance sheet date.

Process mining may further support identifying if controls relevant to financial reporting were subject to significant changes throughout the roll-forward period.[342] For example, changes in the information system may result in additional or less activities in the event log, changes in the process will lead to new process variations and changes in key personnel may be identified by comparing the user information related to the events.

[338] The procedures performed to extend the audit conclusions are referred to as „roll-forward" procedures in practice, whereas the period between the interim date and the balance sheet date is referred to as "roll-forward period", cf. IFAC (2021), ISA 330, para. 12 and 22.
[339] Cf. ibid., para. A34.
[340] Cf. ibid., para. A55 and A57.
[341] Saving analyses performed is provided, for example, by the „Favorite" functionality of UiPath's process mining solution, cf. UiPath (2021).
[342] Cf. IFAC (2021), ISA 330, para. A33.

Documentation

Document the results of the procedures performed in response to the identified risks (ISA 330, para. 28; ISA 230, para. 8)

Analyses performed using automated tools or techniques form part of the auditor's documentation.[343] With regard to documentation of the procedures performed, process mining does not differ from any other automated tool or technique used in the audit. Most process mining solutions available on the market support exporting the analyses performed, including the filters applied. If any cases have been selected for further testing outside the process mining application, the auditor may drill-down to the detailed case information for integration into the documentation of the procedures performed in response to the assessed risks.

[343] Cf. IFAC (2021), ISA 230, para. A3.

2.3.5 Interim conclusion

Process mining can be methodically integrated into various procedures throughout the audit process to support meeting the objectives of the auditing standards. Table 5 summarizes relevant ISA requirements and the key mechanisms substantiating process mining's theoretical suitability to meet these requirements.

ISA requirement	Key mechanism substantiating process mining's theoretical suitability
ISA 315, para. 14 (a)	Despite being no inquiry procedure by nature, process mining can support directing inquiries to risk areas identified from analyzing the data or to aspects of the process that are not evident in the data.
ISA 315, para. 14 (b)	Process mining can be interpreted as analytical procedure that supports understanding an entity's business process and identifying risks of material misstatement by helping to identify inconsistencies and unusual transactions or events in data relevant to the financial statements.
ISA 315, para. 14 (c)	Process mining may be interpreted as a "digital observation" of user activities in the process as it includes reliable information about the execution of a business process based on log files of the ERP system. It further removes the inherent limitation of the audit evidence to the point in time the observation takes place.
ISA 315, para. 13	Instead of relying on observations, inquiries and the study of potentially outdated process manuals, the process flowchart derived with process mining is based on data that has been automatically recorded.
ISA 315, para. 16	The timing of the audit procedures supported by process mining is determined directly by the period for which the data has been extracted. Changes to the prior period's process may be identified by comparing the information between the process mining analyzers.
ISA 315, para. 19	Elements of the analyzed process may contribute to understanding the entity's overall business model and the extent of IT integration.
ISA 315, para. 20	When enhanced with general ledger data, process mining supports understanding which financial accounts are involved in the process, including the nature and volume of the transactions related to the cases processed in the analyzer.
ISA 315, para. 25(a)(i) and (ii)	By reconstructing the process model from the data extracted from an entity's information system and visualizing how the process instances are processed through the system, process mining inherently supports understanding how the process information flows through the system.
ISA 315, para. 25(a)(iv)	Identifying the resources involved in the process and understanding their interactions and roles and responsibilities is supported by the user information of the event log.
ISA 315, para. 26(a); ISA 240, para. 28	The activities in the event log and the obtained understanding of the process support the auditor in identifying those activities that represent control activities relevant to financial reporting. As based on the population of transactions, the technology supports identifying controls for both routine transactions and non-standard, non-recurring or unusual transactions or adjustments.
ISA 315, para. 26(d) and 27	With process mining, the procedures to determine if a control is implemented and to evaluate the design effectiveness may be performed concurrently by investigating the characteristics of cases that do (not) pass the control activity.
ISA 315, para. 28(b), 29, 31 and 32; ISA 240, para. 28; ISA 550, para. 18	The information obtained on the key elements of the process supports the auditor in identifying and assessing risks of material misstatement for related assertions of significant classes of transactions, account balances and disclosures.
ISA 315, para. 38	The analyses performed may be exported in a file format that may be integrated into the auditor's documentation of the risk assessment.

ISA requirement	Key mechanism substantiating process mining's theoretical suitability
ISA 330, para. 10(a)	Process mining may support testing the operating effectiveness of controls as it provides information on how the control was applied throughout the audit period, the consistency with which it was applied and by whom or by what means it was applied.
ISA 330, para. 11	The timing of tests of controls is determined directly by the particular period for which the data is extracted from the system and loaded into the process mining application.
ISA 330, para. 17; ISA 530, para. 12 and 13	Rather than relying on inquiries, process mining supports evaluating control deviations based on the information recorded in the entity's information system. The case and event attributes help to identify common characteristics of cases not passing through a specific control activity and support assessing if an exception is systematic. By determining the nature and volume of control deviations based on data, process mining supports a more robust assessment of control deficiencies. As the technology considers the population of transactions instead of a sample, the auditor may further determine if control exceptions exceed acceptable tolerances.
ISA 330, para. 4(a)(ii); ISA 520, para. 5	Process mining may serve as a substantive analytical procedure, for example, by (a) connecting the financial information to process-related nonfinancial information, (b) comparing the development of KPIs and events over time against the auditor's expected development, (c) performing consistency checks, for example, on the timing of activities and their sequence, (d) analyzing the resulting process variations and comparing individual variations to identify unusual transaction processing and by (e) performing predictive analytics on the expected path of cases based on their characteristics.
ISA 330, para. 4(a)(i)	Process mining may be used to select specific cases and supplement related transactions with supporting evidence (for example, related documents). Selected transactions may include both cases selected as key items as part of other analyses performed with process mining or sampled items.
ISA 330, para. 12 and 22	Tests of controls and substantive procedures performed with process mining may be efficiently re-performed on an extended data population and controls and process variations that have been evaluated at an interim date or in the previous audit period may be compared to the roll-forward period or the balance sheet date, respectively.
ISA 330, para. 28; ISA 230, para. 8	The auditor may export the analyses performed and drill-down to the detailed case information for integration into the documentation of the procedures performed in response to the assessed risks.

Table 5: Key mechanisms substantiating process mining's theoretical suitability to meet relevant ISA requirements

In summary, from a theoretical point of view, the positive statements made in scientific research about the suitability of process mining as an audit instrument can be supported.

Besides supporting the design and execution of audit procedures in line with the ISA, the process mining technology has the potential to increase the quality and efficiency of the audit and its relevance to stakeholders. Process mining may enhance the quality of the audit, as the nature of evidence obtained is more robust compared to the evidence from traditional procedures. With process mining, the bias that is inherent to inquiry and observation procedures is reduced. Bias is further reduced through the use of meta information about the entries made in the ERP system that is recorded independently from the user entering the data. As process mining is based on the total population of transactions, conclusions are more precise compared to judgments or projections based on relatively small samples only. In parallel, process mining may enhance the efficiency of the audit, as analyzing the entire population of transactions may reduce the extent of substantive tests of details significantly. Many manual steps of a traditional audit (for example, walkthrough procedures) can potentially be automated, enabling the auditor to save time and focus on those areas of the audit subject to a higher risk of material misstatement. Due to the increased precision compared to traditional conclusions reached based on samples, the technology may further increase the usefulness of audit results for decision making by management and those charged with governance, for example, for deciding on investments and priorities for improving internal controls.

The following Chapter 3 complements the theoretical discussion of how process mining may be embedded into the audit with an empirical evaluation of the implementation of the technology into the audit practice of one of the Big Four audit firms.

3 Empirical evaluation of process mining in an audit of financial statements

3.1 Feasibility assessment in Germany, Switzerland and Austria conducted in 2017

3.1.1 Scope and design of the feasibility assessment

Scope and timing of the initial piloting project

The case audit firm's project of exploring the practical use of process mining in an audit of financial statements started back in early 2017, when the company introduced process mining to the Assurance practice in Germany, Switzerland and Austria for an initial feasibility assessment. The objective of the first piloting was to obtain insights on the overall suitability of the technology to support an audit of financial statements. Process mining has been piloted on six audit engagements, whereas one of the audit teams tested the analyzer using data from twelve different entities. The pilot teams performed the feasibility assessment in parallel to their traditional audit procedures, that is, due to the piloting state of the project the information obtained with process mining has not been used as audit evidence and was not incorporated into the audit documentation. Table 6 summarizes the engagement information.

Entity	Sector	Process investigated	ERP system
[Entity 1]	Retail/Consumer goods	Purchase to pay	SAP
[Entity 2]	Retail/Consumer goods	Purchase to pay	SAP
[Entity 3]	Middle market	Purchase to pay	SAP
[Entity 4]	Industry products	Purchase to pay (twelve entities), Order to cash (one entity)	SAP
[Entity 5]	Chemicals	Purchase to pay	SAP
[Entity 6]	Automotive	Purchase to pay	SAP

Table 6: Scope of the audit firm's feasibility assessment conducted in 2017

All pilot teams analyzed the purchase to pay process in the audit period from 1 January 2017 to 31 December 2017 and one of the teams used a second application to investigate the order to cash process. In total, 18 process mining applications have been deployed for piloting in five different industries across Germany, Switzerland and Austria.

Process mining application used for initial piloting

The process mining application used for the audit firm's feasibility assessment was based on the "ApplicationOne" as of February 2017 from the software vendor UiPath (former ProcessGold). Figure 21 shows the overview dashboard[344] of the analyzer, summarizing KPIs of the process, listing the activities in the dataset and illustrating different distributions of cases, events and variations in the process.

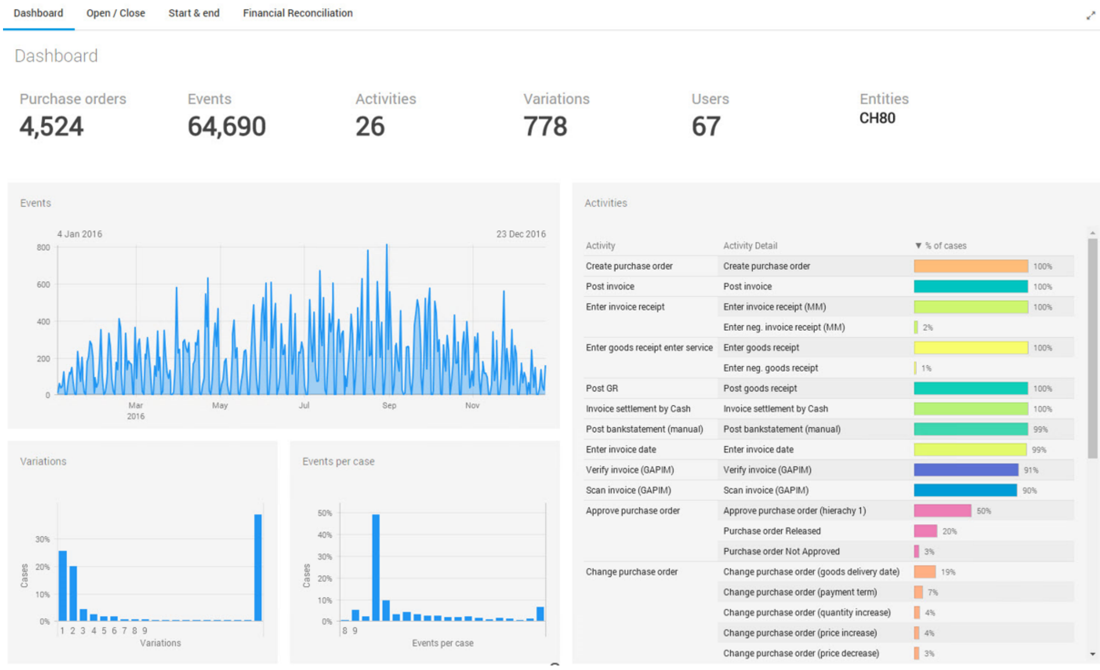


Figure 21: Overview dashboard of the process mining application in 2017

A detailed description of the interface functionalities, the navigation options, the default menu pages, dashboards and analyses provided in UiPath's process mining solution is available in the software vendor's tool guidance.[345] Until today, the case audit firm has extensively customized the application to the requirements of a financial statement audit. For the scope of the first feasibility assessment in 2017, however, only basic modifications of UiPath's process mining software have been conducted.[346]

[344] UiPath's process mining application consists of different menu pages that contain multiple analysis tabs, referred to as "dashboards". Each dashboard contains multiple charts and tables supporting the analysis of the process data.
[345] Cf. UIPATH (2021).
[346] As in 2017 only a first feasibility assessment has been conducted and the vendor's process mining application has not been modified significantly, the application is not presented in detail in this chapter. Refer to UIPATH (2021) for an illustrated description of the default analyses provided in UiPath's "ApplicationOne".

Key modifications include:

(1) the integration of a dashboard supporting the reconciliation of the process data to the general ledger accounts of the auditee,

(2) the integration of a dashboard providing details on the documents related to a case,

(3) the incorporation of quantitative information on the purchase order or sales order amounts related to a case and

(4) the removal of a predefined set of business rules from the application that was intended to flag non-conforming cases for further analysis.

As the subject of the audit are the financial statements of the auditee,[347] the data used by the auditor within any automated tool or technique needs to be reconcilable to the financial statements. Current process mining applications are limited to the analysis of a single process at a time that has as few interrelations with other processes as possible.[348] However, many account balances are affected from multiple classes of transactions. For example, the activity on the bank account is not limited to the payment of purchased goods and services as the account is used in most of the entity's other significant classes of transactions as well. Hence, it is important for the auditor to identify which part of the transaction volume recorded on an account is covered by the analyzed process data and which part is not. Information on transactions flows or (parts of) account balances that are not covered by the analyzed process supports the auditor in identifying a significant class of transactions that has not been previously identified.[349] If there are significant transaction volumes or account balances relating to the process that are not covered by the process data, further audit procedures to identify, assess and address related risks of material misstatement need to be designed.

Figure 22 illustrates the initial implementation of an analysis supporting the reconciliation between the process data and the general ledger data.

347 Cf. IFAC (2021), ISA 200, para. 11(a).
348 Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 359.
349 Cf. IFAC (2021), ISA 315 (Revised 2019), para. A203.

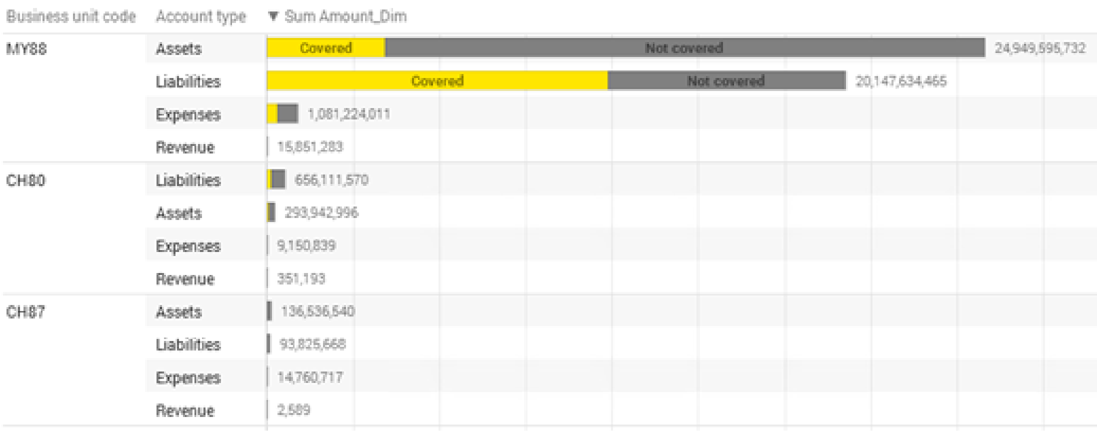


Figure 22: Prototype supporting to reconcile the process and the financial data

The dashboard shows the total movement of the general ledger accounts in the audit period by account type for different business units considered in the dataset. The analysis may be adjusted to show the total transaction movements by financial statement account class or account, respectively. For each account type, account class or individual general ledger account, the bar chart indicates the part of the transaction volume that is covered (colored yellow) or not covered (colored gray) by the process data. The connection between the journal entry data and the cases is established by using the references between the accounting documents in the ERP system and the documents related to cases in the process.[350]

Analyzing flows of transactions using process mining frequently leads to the identification of outliers, for example, cases that are not following an expected process path or cases that for which key activities have not been performed. The UiPath process mining application supports further investigation of these cases by enabling the drill-down to the detailed case and event input data. However, an advantage of applying process mining on data extracted from an ERP system is that in addition to the log files, the information about the documents related to a case is available. Figure 23 illustrates how the header and line item information of different documents related to a case have been incorporated into the process mining application.

[350] Details on the integration of both data sources are provided in Chapter 4.2.2.

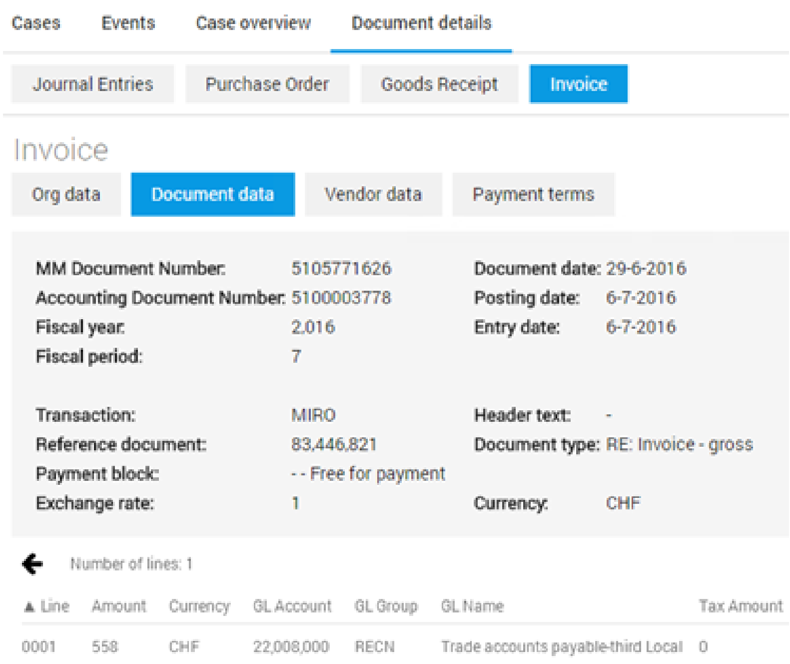


Figure 23: Header and line item information of the documents related to a case

At the top of the dashboard, the auditor may select a document for further investigation. The gray area in the middle of the dashboard consolidates the header information of the document with additional relevant meta-data extracted from the ERP system, for example, regarding the vendor, supplier or payment terms related to the document. At the bottom of the dashboard, the line items of the document are listed. The document information may support the auditor in designing and performing additional audit procedures on exceptional cases identified and facilitates conducting inquiries with the auditee.[351]

Besides the integration of the two new dashboards, quantitative case information has been added to the application. As a case has been determined to represent a purchase order line item in the purchase to pay process and a sales order line item in the order to cash process, the analyses have been customized to provide information not only on the number and percentage of cases but also on the related purchase order or sales order amount.

Lastly, UiPath's process mining solution includes tags representing business rules that are defined for the individual dataset when setting up the analyzer. These tags are used to flag those cases for further analysis that are, for example, approved and paid by the same user or that represent a violation against a specific service level agreement

[351] Cf. Chapter 2.3.4.

(SLA). Figure 24 illustrates a set of predefined tags together with the total number of cases flagged for each tag.

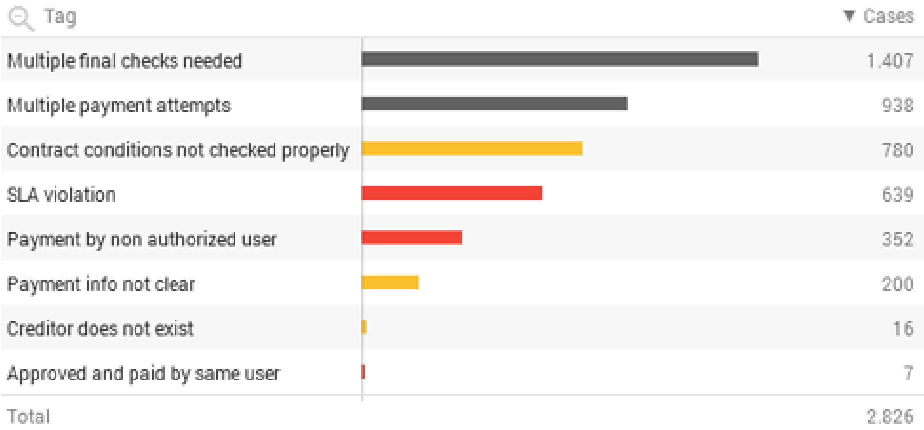


Figure 24: Tag analysis in UiPath's process mining application[352]

When selecting the cases within a specific category, the analyzer is filtered to this subgroup of cases, supporting a more detailed investigation of the related activity flows in the process graph. An important aspect in exploring if process mining may facilitate the auditor's process understanding is the data-driven identification of risks of material misstatement and control activities.[353] However, to identify business practices, business rules or control activities that may be incorporated as a tag into the analyzer, prior knowledge of the entity and an understanding of the process is required already in advance to the analysis. Instead of limiting the analysis to specific predefined business rules only, the subgroups of cases that are not following routine paths of the process or are not passing certain control activities may be identified concurrently with other audit procedures. For example, whether the same user approved and paid a specific case may be identified when analyzing the segregation of duties between the related approval and payment activities. Similarly, the auditor may identify payments by non-authorized users by investigating the user's involvement in each payment-related activity. If the information on authorized users is instead required already during the setup process of the analyzer, the reliability of the audit evidence obtained by investigating cases related to the tag depends on the completeness and accuracy of this input data. A new authorized individual in the current period that is not considered in the definition of the tag "payment made by non-authorized user"

[352] Cf. UiPath (2021), navigate to "Menu Conformance".
[353] Cf. Chapter 2.3.3.

would lead to false positive results, whereas employees considered but no longer authorized would lead to false negative results. To support an effective and efficient audit of financial statements, data analytics should not provide any false negative results and as few false positive results as possible.[354] Consequently, the custom predefinition of tags has been removed from the analyzer.

Delivery and support process for the audit teams

While the design of the analyses within the application has not been changed significantly for the first piloting, more work needed to be devoted to the process of extracting the raw data from the entity's ERP system, transforming the data into a data model suitable for process mining and loading the input data into the process mining application. Most of the process mining case studies described in current research are based on data derived from workflow-engine-based systems that provide dedicated functionalities to support data extraction for process mining.[355] However, there is only a limited number of case studies based on other systems, such as ERP systems or related preprocessing applications.[356] Furthermore, a "best practice" for assembling the relevant input data from the source systems and building the event log for process mining is not yet established.[357]

The ETL process[358] and the delivery of the process mining application for each pilot engagement was facilitated by a team of domain experts, including ERP specialists, IT experts, data engineers and data scientists. These technical experts supported the audit teams in identifying and extracting the raw data, creating a customized process mining data model for each specific entity and transforming the input data into an event log format that may be loaded into the process mining application.

Besides the process mining application, the audit teams received the software vendor's official end-user guidance explaining the functionalities available. With each engagement team, either a workshop or dedicated coaching sessions have been conducted, covering an introduction to process mining, support in familiarizing with the application and a discussion of audit procedures that may be supported by the solution. Further

354 Cf. WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021), p. 13.
355 Cf. IEEE TASK FORCE ON PROCESS MINING (2012).
356 Cf. AGUIRRE, SANTIAGO/PARRA, CARLOS/SEPULDEVA, MARCOS (2017), p. 103; Chapter 2.2.
357 Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 6.
358 The process of extracting, transforming and loading the data from different data sources into a database or another data repository is referred to as ETL process.

technical and methodical execution support was provided on demand during the feasibility assessment.

3.1.2 Audit team feedback and results from quality assurance procedures

Feedback obtained from the initial feasibility assessment

Feedback interviews have been conducted with all pilot engagements in February and March 2018. Feedback has been obtained on:

(1) the process of extracting, transforming and loading the data (ETL process),
(2) the integration of the procedures performed with process mining into the audit approach,
(3) the tool handling and tool performance,
(4) the analyses provided in the analyzer and
(5) any additional comments from the audit teams.

The feedback obtained within each category summarizes as follows:

At (1): Feedback on the ETL process

- All audit teams confirmed that an extensive involvement of IT experts was required in communicating the data requirements to the auditee and in the process of extracting, transforming and loading the process mining input data.
- Many audit teams noted that even when supported by domain specialists, a detailed understanding of the entity's process, including a solid understanding of the IT landscape relevant to the process is needed to appropriately specify the data requirements. Some teams concluded that in order to determine the IT applications and the activities to consider in the event log, the involvement of experienced members of the engagement team is required already in the setup process of the application.
- The ETL process has been performed manually to a large extent. The technical expert team reported that frequently, many iterations with both the auditees and the audit teams have been necessary due to errors in extracting and transforming the data. One audit team reported that an incomplete data extraction performed by the entity lead to significant delay in the delivery of the final application. Another team noted that due to the size of the dataset, the observation period has been reduced retrospectively from twelve to six months.

- For four audit teams, the scope of the data extraction needed to be expanded as the entity made use of customized SAP workflows not considered in the initial data extraction. One of the teams reported that resource restrictions at both the audit team and the technical team prevented a retrospective implementation of these custom activities on a timely basis.

At (2): Feedback on the integration into the audit approach

- All engagement teams agreed that the transparency achieved with process mining increases the quality of the process understanding. Teams especially mentioned the potential of the technology to support risk identification and risk assessment procedures and evaluating the control activity component of the entity's system of internal control.
- As process mining includes the initiating activities of the process, teams further highlighted that with process mining it should be possible to analyze the entire critical path[359] of a significant class of transactions according to ISA 315 using data analytics.
- The pilot teams noted that they have been able to identify key controls in the process and evidence their operation. Common controls mentioned by the engagement teams analyzing the purchase to pay process included the approval of a purchase order by an authorized individual, the use of the GR/IR account for processing material-related transactions, the automatic recording of the bank clearing after a payment run has been performed and the segregation of duties between specific activities in the process.
- Frequently, the documentation of the understanding of the process obtained in prior audit periods has been extended as the traditional walkthrough procedures only covered the most frequent flow of transactions in the process. As a result, two teams reported that they were able conduct more informed inquiries with the entity, focusing on manual aspects of the process that are not evident in the data and asking more specific questions related to identified risks of material misstatement. On the other hand, teams stated that the audit efficiency decreased when compared to the traditional audit approach.
- One audit team stated that with process mining, more precise audit conclusions are expected as the technology is based on the entire population of transactions

[359] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 25(a).

rather than a sample only. At the same time, however, the expected increase in audit quality needs to be contrasted to the cost and effort involved to set up the analyzer and design and perform the audit procedures.

- All audit teams highlighted that additional methodical guidance on how to audit a process using process mining is required, especially regarding the conformance of the technology with the requirements of the ISA.
- In particular, the pilot teams faced challenges in determining the nature and extent of audit procedures necessary to evaluate the high number of resulting process variations. Provided the multitude of the variations, the dashboard supporting the comparison of two different process traces was not perceived as an appropriate instrument to conclude on the appropriateness of the individual process executions. Some teams further raised the question whether the evaluation of process variations is subject to professional judgment only or whether a best practice may be established, for example, evaluating the top ten variations or a certain percentage of variations.
- In testing the operating effectiveness of controls, the auditor uses a sample-based approach according to ISA 530. As the sample deviation rate represents the projected deviation rate for the entire population, the auditor and does not project misstatements identified to the population as a whole.[360] One audit team was unsure how to evaluate control exceptions identified with process mining based on the total population of cases in the process, especially if in parallel the control has been tested with the traditional sample-based approach without any exceptions identified.
- As the documents related to a case are available in the process mining application, one of the pilot teams suggested to further explore the use of process mining not only for testing internal controls but also to support substantive procedures.
- Audit teams missed an efficient and effective way of documenting the audit procedures performed with process mining to meet regulatory requirements, including the audit trail of analyses performed and the rationale for filtering subpopulations of data.

[360] Cf. IFAC (2021), ISA 530, para. 14 and A20.

At (3): Feedback on the tool handling and tool performance

- All audit teams emphasized the high performance and quick response times of the application.
- For most of the audit teams, navigating through the application was perceived intuitive. Some teams requested a separate training session on the general application handling.
- One audit team fed back that the menu pages and dashboards should be renamed to better align with terms commonly used in an audit of financial statements.

At (4): Feedback on the analyses provided in the analyzer

- All audit teams pointed out that they would benefit from additional technical guidance on the functionalities provided in the analyzer and dedicated coaching sessions that facilitate customizing the dashboards and interpreting the analyses.
- The reconciliation of the process data to the financial statements was not intuitive. One team missed an export functionality for the financial data to facilitate reconciliation with the auditee's trial balance. Three teams remarked that the link between the cases in the analyzer and the general ledger account balances was not clear. Another team noted that the coverage of the trade payables related accounts by the process data was very low as the process mining application did not consider the auditee's customized workflow for invoices that are not related to a purchase order.

At (5): Additional feedback received from the audit teams

- In the application used for the feasibility assessment, the process instance has been determined to be a purchase order line item (or a sales order line item in the order to cash process, respectively). Consequently, in UiPath's process mining application, the related case amount is the purchase order line item amount. The initial engagement team feedback confirmed that for future versions of the application, the invoice amount related to the cases should be considered as an additional metric to analyze. Additionally, one team requested the integration of a functionality that supports applying user-defined thresholds to the analyses to consider, for example, materiality thresholds or tolerances specified for the execution of specific controls.

- Provided the complexity of the overall process graph, one team proposed to integrate a functionality that supports coloring or grouping specific activities within the process graph, for example, key control activities or the set of activities relating to different documents within the process.
- To enable the adoption of process mining on additional audit engagements, one audit team asked for the availability of process mining for entities using ERP systems from Oracle and Microsoft Navision.
- All six pilot engagements acknowledged the potential of process mining for the audit and committed to continue participating in piloting for the next audit period.

Feedback evaluation and results of quality assurance procedures

Overall, the first empirical feedback obtained on the suitability of process mining to support an audit of financial statements was positive. The pilot teams especially highlighted the benefits process mining may provide to their process understanding, including risk assessment and walkthrough procedures. Feedback further confirmed the potential of the technology for extending the use of data analytics in the audit to the identification and testing of internal controls.

However, several larger action items have been identified that needed to be addressed before continuing to explore the application of process mining in the audit. On the one hand, these action items resulted from the audit team feedback that has been evaluated and prioritized accordingly for further consideration in upcoming releases of the application. On the other hand, challenges have been identified as part of the quality assurance procedures conducted by the author of this thesis throughout the piloting process. The quality assurance procedures included both (a) a detailed review of each individual process mining application prepared by the technical expert team and delivered to the audit team and (b) an investigation of the design and execution of the procedures performed by each audit team using process mining. The objective of the quality assurance procedures was to ensure appropriate functionality and performance of each process mining analyzer and to avoid systematic errors when analyzing the data and interpreting the results of the analyses.

The main action items identified from the feasibility assessment included:

- the standardization and automation of the ETL process and the data model,
- the development of methodical and technical guidance,

- the enhancement of existing functionalities and analyses and
- the design and implementation of additional functionalities and analyses.

Standardization and automation of the ETL process and the data model

The process of data extraction, transformation and loading was a manual task during the initial piloting. The feedback includes occasions where most of the time planned for auditing the purchase to pay process was consumed for data extraction and preparation. In some instances, the scope of the data export needed to be extended retrospectively after setting up and validating the analyzer, leading to significant delay in the delivery of the final application. Furthermore, due to the extensive resource involvement in the manual data preparation, the pilot teams received different levels of support from the technical development team. Customized workflows in the ERP system and preprocessing activities performed in other IT applications have only been incorporated into the analyzer for some of the audit teams. Due to the extensive prerequisites related to data preparation and transformation, some audit teams challenged the benefit process mining may provide to the audit of internal processes and controls.

Quality assurance procedures conducted as part of the feasibility assessment further revealed that the description and labeling of the same activities in the same process differs across engagements, aggravating the comparison of the activities and the overall process across multiple process mining applications. Varying technical descriptions of activities further impact the reliability of the data and interpretation of the actual procedure that is performed from a business perspective. For example, the initiating activity of the purchase to pay process was labeled with “approve purchase order”, “purchase order approval” or “purchase order approved” across different engagements. As the timestamp that is recorded upon the click on the approval button in the ERP system reflects the end of the approval process, the uniform label “purchase order approved” would facilitate interpretation from a business perspective.

Based on the results of the first piloting, decisions have been made to standardize and automate the ETL process and the data model to the largest extent possible. The objective was to facilitate both the technical support structures and later scalability of the solution to a larger number of regions and audit engagements. As a prerequisite for standardization, the first process mining application was limited to the analysis of the purchase to pay process only. In order to develop and maintain a common process mining data model, the scope of the piloting was further limited to entities primarily using the ERP system SAP to process transactions in purchasing. As part of the standardization, a set of standard activities performed in the default configuration of the purchase to pay process in SAP has been developed. The repository of standard activities was tested, validated and maintained by the central development team and includes a unique label and description, a standardized technical definition and information on the related data source for each activity. In a second step, the set of standard activities can be adjusted and extended by "custom" activities based on the SAP configuration, customized workflows and relevant preprocessing IT applications of the specific entity. The final decision on the activities to include for an individual audit engagement is then made by the auditor based on the understanding of the process, previous experiences with the entity and the auditor's professional judgment. Over time, new audit teams may be supported in adopting the technology by a repository of custom activities already tested and validated for entities with similar IT configurations.

Development of methodical and technical guidance

All audit teams stated that additional technical and methodical guidance is required that facilitates using the technology to perform audit procedures. Technical guidance involves not only guidance on the ETL process and related roles, responsibilities and data requirements but also explanations of the functionalities, analyses and customization options provided in the process mining analyzer. Methodical guidance was requested regarding the integration of process mining into the audit approach. In particular, teams faced challenges in understanding and evaluating resulting process variations and evaluating control exceptions identified from the total population of transactions. Further, provided the detailed insights in the process execution for each individual process instance, several audit teams noted that further guidance on the nature and extent of documenting the audit procedures performed is necessary. In response to the feedback obtained, the author of this thesis prepared guidance for auditing with process mining, including both methodical and technical guidance on the application of process mining in an audit of financial statements. Besides the guide intended for use as reference book, the application of process mining in the subsequent period has been supported with a template supporting the documentation of procedures performed, additional technical and methodical training, coaching and experience-sharing sessions.

Enhancement of existing process mining functionalities and analyses

From the feedback received on the analyses provided in the process mining application, the redesign of the financial reconciliation, the integration of the invoice amount related to a case as additional case metric as well as the enhancement of the export functionalities facilitating documentation have been prioritized for the next version of the analyzer.[361]

Design and implementation of additional analyses

In exploring the use of process mining to support audit procedures, the pilot teams were flexible in adjusting the analyses provided in the process mining application. The technical expert team supported the audit teams in customizing the dashboards to their engagement as necessary, for example, by adding additional case or event attributes to existing analyses, adjusting the default filter functionalities supported by the solution or designing and implementing new analyses. However, quality assurance procedures revealed several downsides of customizing the application individually for each specific entity.

Frequently, the newly incorporated analyses were redundant to already existing functionalities, as the same perspective of the data could have been achieved by adjusting existing tables in the application using the default filter and selection options provided.

A customized analysis may further lead to unexpected results for the audit team. For example, for one audit team, an analysis of the invoice types received in different business unit has been implemented, considering the number of vendors related to each invoice type and business unit.

[361] Details on major modifications made to the process mining application are provided in Chapter 3.2.1.

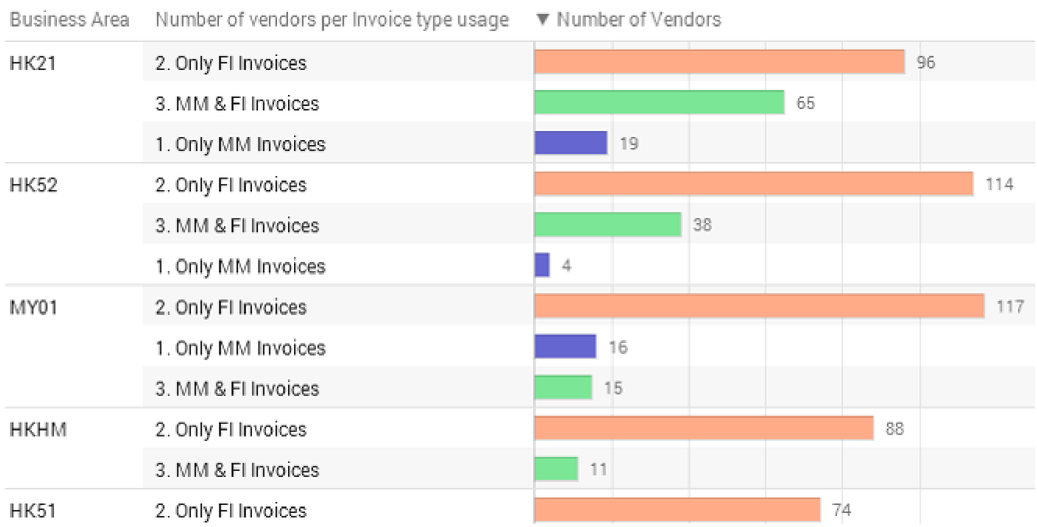


Figure 25: Custom analysis of the types of invoices received

In Figure 25, “MM Invoices” is referring to transactions processed using the SAP material management (SAP MM) module. These transactions include a goods receipt and are related to a purchase order. “FI Invoices” are processed using the SAP financial invoice (SAP FI) module and are related to direct purchases that do not contain a purchase order or a goods receipt. In performing the analysis, the audit team raised the issue that the data could not be filtered in order to investigate, for example, the cases related to a specific invoice type, vendor or business unit in the process graph. A link between the custom analysis and the process graph may not be established because the process instance used in the audit firm’s process mining application is a purchase order line item whereas financial invoices are not related to a purchase order. As such, transactions processed using the SAP FI module are not included in the event log and thus not covered by the process data.

Lastly, if the information obtained with process mining is used as audit evidence, additional quality assurance procedures with regard to the technical design and implementation of the custom analyses need to be performed, as they are not considered in any centralized procedures ensuring the appropriate technical functionality of the process mining application.

As a result, decisions have been made to standardize the set of dashboards and analyses provided in the process mining application to ensure that different applications meet the same quality criteria. New functionalities for the standard application may be suggested by the audit teams as part of the piloting projects or regular user acceptance testing performed for new releases of the application.

Missing process mining functionalities to support auditing the purchase to pay process that have been identified as part of the feasibility assessment include:[362]

- analyses supporting the understanding of the vendor structure and changes therein according to ISA 315,[363]
- analyses facilitating the evaluation of process variations and
- additional analyses related to common control activities typically evaluated by the auditor as part of auditing the purchase to pay process.

3.2 Piloting in Europe, Middle East and Africa conducted in 2018

3.2.1 Scope and design of the piloting project

Scope and timing of the piloting project

After the introduction of process mining to the audit practice, the piloting has been extended to additional countries in Europe, Middle East and Africa. In the period from September 2018 until June 2019, the process mining analyzer has been tested on 19 audit engagements in eleven different countries using data from 1 January 2018 to 31 December 2018. Most pilot teams applied process mining at both an interim date and at period end, resulting in a total number of 35 process mining analyzers that have been deployed for piloting. The objective of the second piloting project was to obtain further empirical evidence on the suitability of the technology to support an audit of financial statements and to validate the changes made to the process mining application. Similar to the feasibility assessment, due to the piloting state of the project the audit teams performed the analyses in parallel to their traditional audit procedures and did not use the results obtained as audit evidence.

Besides the continuing standardization and automatization of the ETL process, the audit firm's delivery model of the process mining applications to the audit teams did not change significantly compared to the feasibility assessment. In advance to the delivery of the process mining application using the individual entity's data, all teams have been provided with guidance on using process mining, a template to document the procedures performed and access to a central process mining application using sanitized data to support obtaining an overall understanding of the technology.

[362] The design and implementation of these functionalities is described in Chapter 3.2.1.
[363] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A57, A68 and Appendix 1; Chapter 2.3.3.

Modifications of the process mining application

After the first piloting project, the process mining analyzer has been continuously enhanced to meet the requirements of the audit and support the auditor in effectively performing audit procedures related to the purchase to pay process. The adjustments of the analyzer were considering the feedback obtained from the first piloting project and the quality assurance procedures performed. Figure 26 shows the overview dashboard of the process mining application used for the extended piloting in 2018.

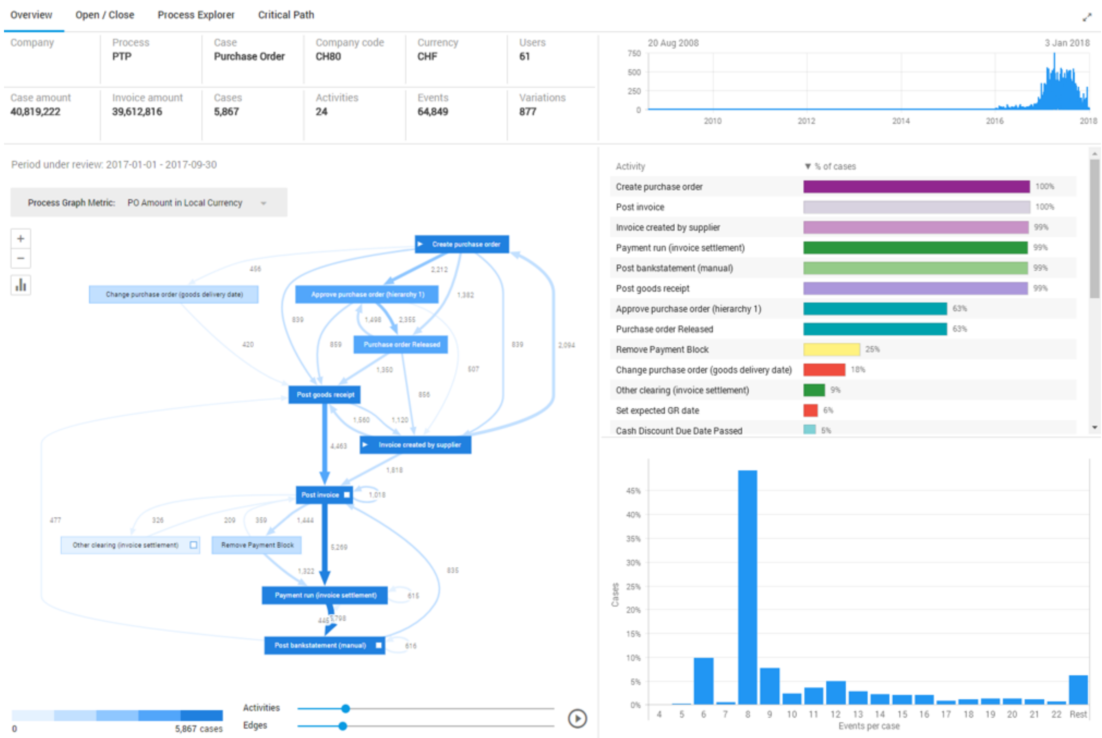


Figure 26: Overview dashboard of the process mining application in 2018

Key modifications of the analyzer include:

(1) the re-design of the reconciliation between the process data and the general ledger data,

(2) the integration of a three-way-match analysis between the purchase order, goods receipt and invoice documents included in the process data,

(3) the implementation of a dashboard facilitating the analysis of process variations and

(4) the integration of vendor-related analysis, including vendor master data.

At (1): Re-design of the analyses related to the financial reconciliation

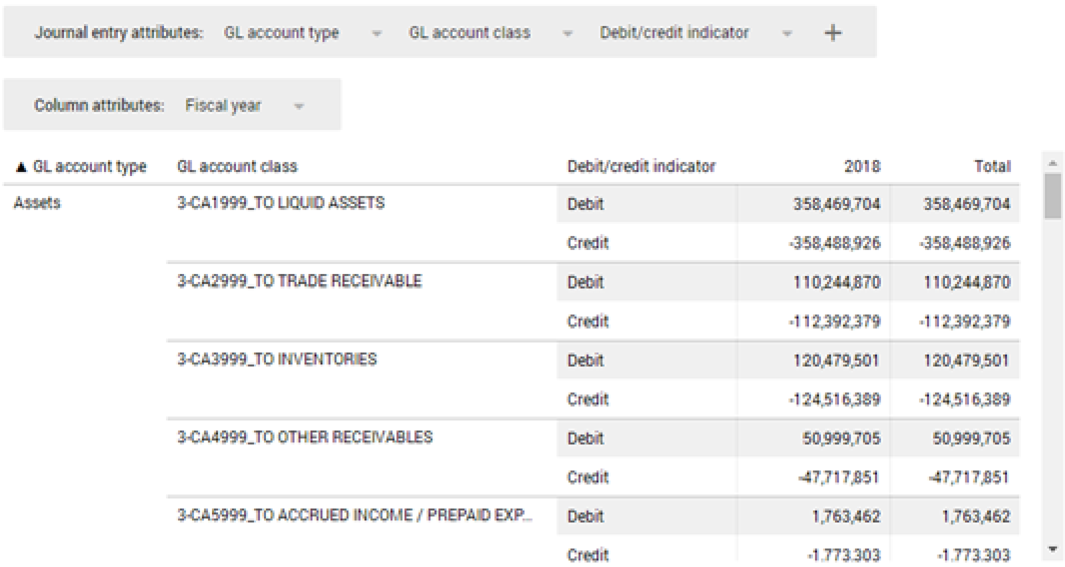
As process mining focuses on the process perspective and has not been designed for a joint investigation of process and financial data, several challenges have been identified when piloting the first prototype[364] supporting the reconciliation between both data sources.[365]

The purpose of the reconciliation between the process data and the financial data is to understand which financial accounts are involved in the process, the nature and volume of the transactions covered by the process data as well as the recording of these transactions throughout the process. Consequently, the analysis has been reorganized into a new menu page "Financial reconciliation" with different analyses supporting the auditor in meeting these objectives:

- The "General ledger activity" analysis enables reconciling the general ledger data used in process mining to the balance sheet and income statement accounts in the entity's trial balance.
- The "Account coverage" analysis facilitates understanding the nature and volume of (a) transactions covered by the process mining analyzer and (b) transactions that are expected to relate to the process but are not included in the process mining analyzer.
- The "Booking pattern" analysis supports identifying how journal entries are posted in the process by displaying the general ledger accounts that are debited and credited upon the execution of related activities in the process.

Figure 27 shows the "General ledger activity" dashboard that is based on the general ledger data exported from the ERP system.

[364] Cf. Figure 22.
[365] Cf. Chapter 3.1.2.



▲ GL account type	GL account class	Debit/credit indicator	2018	Total
Assets	3-CA1999_TO LIQUID ASSETS	Debit	358,469,704	358,469,704
		Credit	-358,488,926	-358,488,926
	3-CA2999_TO TRADE RECEIVABLE	Debit	110,244,870	110,244,870
		Credit	-112,392,379	-112,392,379
	3-CA3999_TO INVENTORIES	Debit	120,479,501	120,479,501
		Credit	-124,516,389	-124,516,389
	3-CA4999_TO OTHER RECEIVABLES	Debit	50,999,705	50,999,705
		Credit	-47,717,851	-47,717,851
	3-CA5999_TO ACCRUED INCOME / PREPAID EXP...	Debit	1,763,462	1,763,462
		Credit	-1,773,303	-1,773,303

Figure 27: General ledger activity for reconciliation

The dashboard shows the balance sheet and income statement accounts and the related debit and credit movements of the audit period that can be reconciled to the trial balance prepared by the auditee. The analysis may be adjusted to display the monthly debit and credit transaction movements throughout the period, respectively. To support identifying any non-reconciling items for investigation, additional attributes (for example, the individual accounts, transaction codes or posting keys) can be added to the analysis.

Besides the reconciliation of the data, the analysis supports the auditor in understanding the nature and volume of transactions as required by ISA 315. The auditor may:

- assess the total general ledger activity for each significant account class and identify and filter the account classes being part of the purchase to pay process (for example, the account classes “trade payables”, “inventory” and “cash”),
- identify and filter general ledger accounts within purchase to pay by adding the individual accounts to the analysis (for example, the GR/IR account),
- identify general ledger accounts that are new or unusual or have movements throughout the period that are unexpected and
- understand the transaction codes used for processing transactions within purchase to pay and verify the appropriate and consistent use throughout the period.

Typical posting activities in the purchase to pay process include the posting of goods receipts, invoices, payments and bank clearings. As these activities are triggering the recording of journal entries, a link between the process data and the general ledger accounts can be established using the document references in the journal entries as well as the posting key and transaction type information. To understand the debit and credit amounts of the financial accounts related to the process that are covered by the process mining data, the auditor may use the "Account coverage" analysis illustrated in Figure 28.[366]

GL Account Type	▲ GL Account	Coverage	Debit	Credit	Net Activity
Liabilities	4-CL1999_TO BORROWINGS CURRENT	Not Covered	352,597,229	349,919,191	2,678,038
		Covered		848,312	-848,312
	4-CL2999_TO TRADE PAYABLES	Covered	103,925,489	107,799,010	-3,873,521
		Not Covered	62,792,419	62,977,191	-184,772
	4-CL4999_TO OTHER PAYABLES	Not Covered	29,386,031	30,085,956	-699,926
		Covered	5,168,925	4,930,452	238,473
	4-CL5999_TO ACCRUED EXPENSES	Not Covered	13,328,091	13,505,788	-177,697
		Covered	16,259		16,259
	4-CL6999_TO PROVISIONS	Not Covered	847	736	111
Assets	3-CA1999_TO LIQUID ASSETS	Not Covered	258,316,444	217,018,668	41,297,776
		Covered	53,316,022	94,648,351	-41,332,330
	3-CA2999_TO RECEIVABLE	Not Covered	79,632,264	79,320,464	311,800
	3-CA3999_TO INVENTORIES	Not Covered	25,883,503	76,009,167	-50,125,664
		Covered	49,244,083	2,244,193	46,999,890

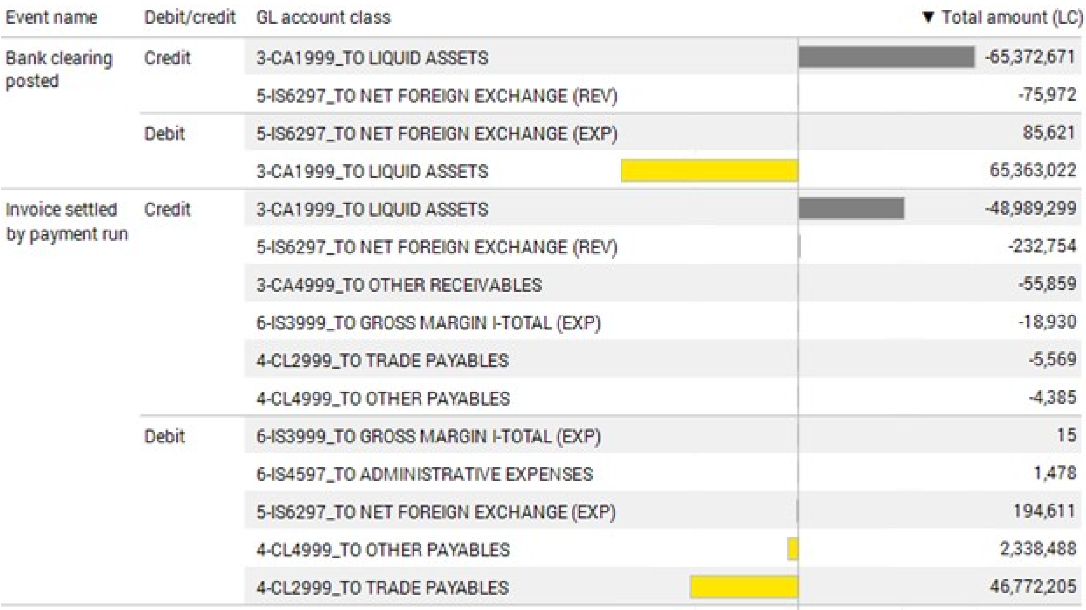
Figure 28: Account coverage by general ledger account class

The analysis shows the debit, credit and total movements of the balance sheet and income statement accounts throughout the period. For each account, the part of the debit and credit movements that is covered by the process data is distinguished from transaction volumes not covered. The analysis supports the auditor in designing appropriate audit procedures outside the process mining application to address risks of material misstatement for those amounts of a significant account not covered by the process mining analyzer. Common transactions related to purchase to pay that are not covered by the process data include accruals, prepayments, value added tax (VAT) entries, payments of other liabilities and manual adjustments posted to these accounts, for example, year-end valuation adjustments like the reclassifications of creditors with debit balances. As these transactions are not related to a purchase order, they are not considered as cases in the process mining application. The auditor may use the

[366] The initial prototype of the account coverage dashboard has been introduced in Chapter 3.1.1.

"General ledger activity" dashboard[367] to further analyze the transactions related to the process but not covered by the process data and drill down to the detailed journal entries as necessary for designing appropriate audit procedures for non-covered transactions.

The third analysis incorporated for analyzing financial data is the "Booking pattern" dashboard illustrated in Figure 29.



Event name	Debit/credit	GL account class	▼ Total amount (LC)
Bank clearing posted	Credit	3-CA1999_TO LIQUID ASSETS	-65,372,671
		5-IS6297_TO NET FOREIGN EXCHANGE (REV)	-75,972
	Debit	5-IS6297_TO NET FOREIGN EXCHANGE (EXP)	85,621
		3-CA1999_TO LIQUID ASSETS	65,363,022
Invoice settled by payment run	Credit	3-CA1999_TO LIQUID ASSETS	-48,989,299
		5-IS6297_TO NET FOREIGN EXCHANGE (REV)	-232,754
		3-CA4999_TO OTHER RECEIVABLES	-55,859
		6-IS3999_TO GROSS MARGIN I-TOTAL (EXP)	-18,930
		4-CL2999_TO TRADE PAYABLES	-5,569
		4-CL4999_TO OTHER PAYABLES	-4,385
	Debit	6-IS3999_TO GROSS MARGIN I-TOTAL (EXP)	15
		6-IS4597_TO ADMINISTRATIVE EXPENSES	1,478
		5-IS6297_TO NET FOREIGN EXCHANGE (EXP)	194,611
		4-CL4999_TO OTHER PAYABLES	2,338,488
		4-CL2999_TO TRADE PAYABLES	46,772,205

Figure 29: Booking pattern by account class

The analysis displays the journal entry activity on the general ledger accounts related to the activities in the process. For the transactions identified as covered by the process mining analyzer, the auditor may confirm if the debit and credit accounts used in recording journal entries correspond to the expected bookkeeping principles in the purchase to pay process of the entity. The analysis may be disaggregated to the individual accounts included in the account classes or the transaction codes, document types or posting keys involved to determine if the entity's booking behavior within the process complies with applicable generally accepted accounting principles (GAAP). In a typical purchase to pay process, the significant booking activity covered by the process results from the posting of goods receipts, invoices, payment runs and bank clearings. Figure 29 shows the general ledger movements triggered by the „invoices settled by payment run" and "bank clearing posted" activities. The analysis considers the transaction volume for all events that relate to the execution of these activities in the period investigated. In the example provided, the payment run is debiting third party trade

[367] Cf. Figure 27.

payables and crediting the bank clearing account. The posting of the bank clearing performed within the bank account class is debiting the bank clearing account and crediting the bank account.

At (2): Three-way-match between the purchase orders, goods receipts and invoices

A common control established by management to address the risks of material misstatement at the assertion level within the purchase to pay process is the three-way-match of the purchase order, the goods receipt and the invoice. The term "three-way-match" refers to an application control within the ERP system designed to prevent misstatements by validating the accuracy of the invoice before a payment is issued to the vendor delivering the goods.[368] Figure 30 illustrates the concept of the three-way-match.

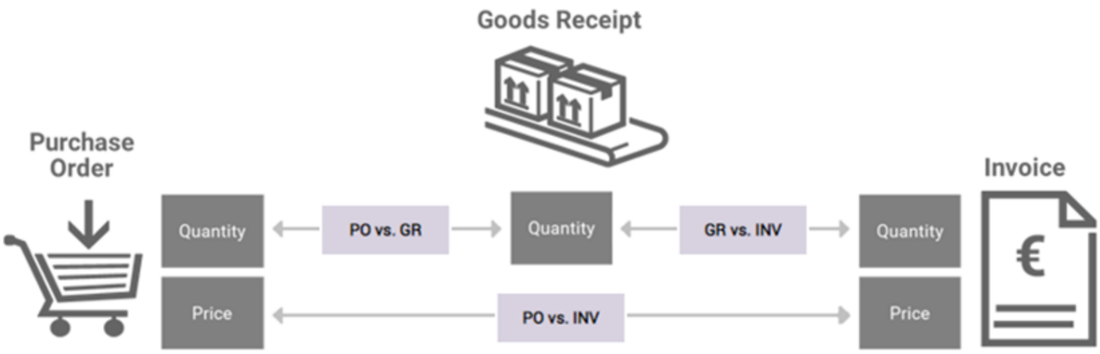


Figure 30: Concept of the three-way-match between purchase order, goods receipt and invoice

By performing a reconciliation between the purchase order, goods receipt and invoice related to a transaction, the three-way-match ensures that the entity both requested and received the goods the invoice is issued for. The quantity of goods stated on the goods receipt document is compared to the quantity as per the purchase order (first match) and the quantity on the invoice (second match). Further, the expected price based on the purchase order is matched to the actual price included on the invoice (third match).[369]

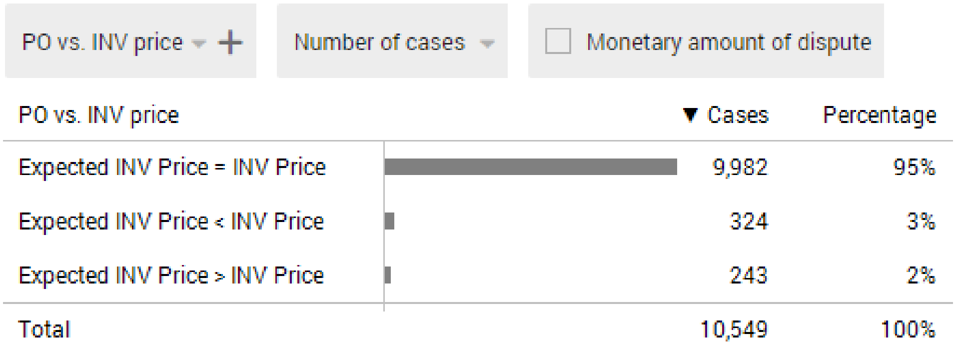
While the feedback indicates that the teams have been able to use process mining to identify key controls in the process, these controls did not include the three-way-match. The review of the pilot teams' prior period's audit plans, however, confirmed

[368] Cf. ELDER, RANDAL J. et al. (2020), p. 608.

[369] In SAP, the three-way-match is controlled using the GR/IR clearing account. When goods are received, they are recorded using the price on the purchase order and the quantity of the goods received. The goods receipt is recorded with a journal entry debiting inventories and crediting the GR/IR account. When the invoice is received, it is clearing the GR/IR account with a journal entry debiting the GR/IR account and crediting the related vendor's trade payables account.

that every audit team identified the three-way-match as a control relevant to financial reporting.

The three-way-match in SAP is performed on the level of a purchase order line item, which has been determined as the process instance for the audit firm's process mining analyzer. As the analyzer includes the price and quantity information of the documents related to a case, the three-way-match performed by the system may be reperformed with process mining by reconciling the document information for each individual case. Figure 31 illustrates how the match between the purchase order and the invoice price may be reperformed with process mining.



PO vs. INV price	▼ Cases	Percentage
Expected INV Price = INV Price	9,982	95%
Expected INV Price < INV Price	324	3%
Expected INV Price > INV Price	243	2%
Total	10,549	100%

Figure 31: Implementation of the two-way-match between purchase order and invoice price

In the example provided, the invoice price matches the expected price based on the purchase order for 95 percent of all cases. For a total of 567 cases, mismatches are identified. With process mining, the auditor may not only obtain information about the number of cases with mismatches but also the total monetary amount of disputes identified.

As blocking all invoices with minor disputes for payment would lead to significant delays in the daily business, organizations usually configure absolute and/or relative tolerances for the three-way-match.[370] Specific tolerance keys are used to automatically block an invoice for payment to the vendor, for example, if:

- there is no goods receipt related to a purchase order line item,
- the invoice quantity differs from the quantity stated on the purchase order or the goods receipt by more than the specified absolute or relative tolerance or
- the invoice price deviates from the purchase order price by more than the specified absolute or relative tolerance.[371]

[370] Cf. HARTKE, LARS/HOHNHORST, GEORG/SATTLER, GERNOT (2010), pp. 368 and 383.
[371] Cf. ibid., pp. 377 and 382f.

The payment block of an invoice may only be removed if appropriate actions are taken to resolve the dispute, for example, recording the related goods receipt, correcting the invoice posting or adjusting the quantity or price information in the purchase order, respectively. As such, besides the three-way-match application control, the auditor frequently identifies an IT dependent manual control related to payment blocks as relevant for the audit, requiring that exceptions above tolerances are followed-up upon on a timely basis.

With process mining, the three-way-match tolerances applied by the auditee may be exported from the ERP system. The auditor may review the history of tolerances, identify any changes throughout the audit period or when compared to the prior period and evaluate the reasonableness of the tolerances applied. Figure 32 illustrates how the tolerances may then be applied to the exceptions identified in reperforming the three-way-match.

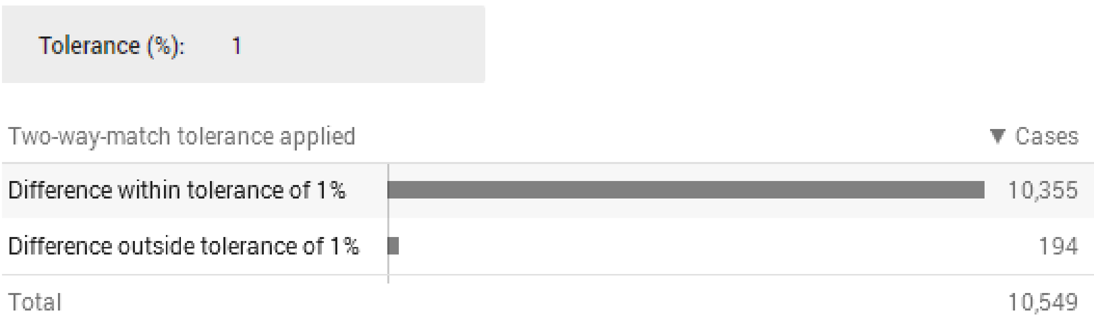


Figure 32: Application of a price tolerance of one percent to the three-way-match

The total number of 10.549 cases corresponds to the cases subject to the reperformance of the three-way-match in Figure 31. After applying the price tolerance of one percent to the exceptions identified, the number of exceptions is reduced from 567 in Figure 31 to 194 cases in Figure 32. For these cases outside tolerances, the auditor may expect that the invoice has been blocked for payment until appropriate actions have been taken to address the disputes identified in the three-way-match.

Consequently, activities related to the setting and removal of a payment block as well as activities related to changes of documents in the process have been integrated into the process mining analyzer. In the example provided in Figure 32, the auditor may filter the cases outside the tolerance and use the process graph to investigate whether the related invoices have been blocked for payment and appropriate activities have been performed before the payment has been made to the vendor.

At (3): Critical path dashboard facilitating the analysis of process variations

UiPath's process mining solution provides a functionality to compare different process variations. Figure 33 shows the comparison of the two most frequent process variations in an invoice approval process, including 498 and 110 cases, respectively.

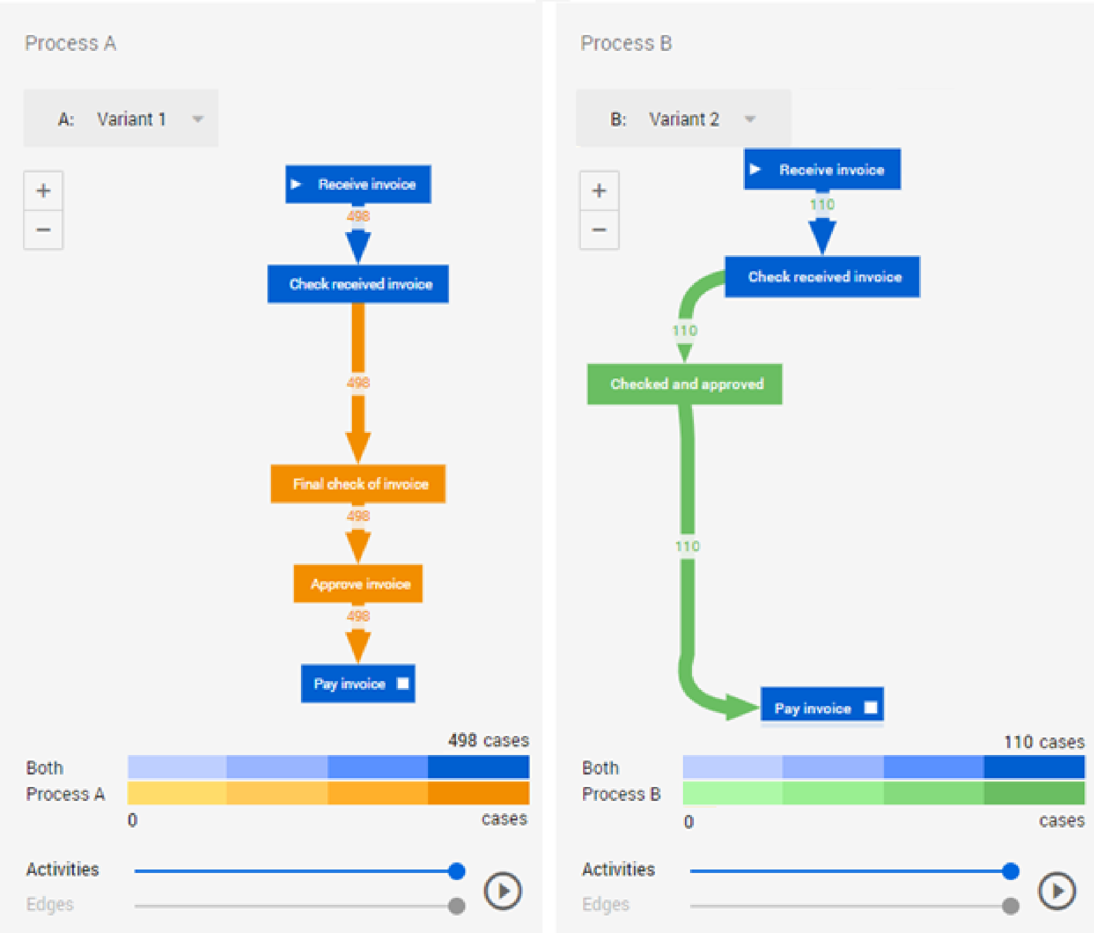


Figure 33: UiPath's dashboard for comparing process variations[372]

For variation 1 displayed at the left side, the sequence of activities is as follows: "receive invoice, check received invoice, final check of invoice, approve invoice, pay invoice". Activities that have only been performed for the variation on the left are colored orange. Any additional activities performed in the variation on the right are colored green. For variation 2 displayed at the right side, the final check of the invoice and the approval of the invoice have been performed jointly, represented by the activity "checked and approved". In comparing the different process variations, the auditor may determine if the sequence of activities is in line with his or her expectation of the invoice approval workflow of the entity and whether the variations are reflective of

[372] The figure has been derived from the demo application in the training environment of the vendor.

the entity's business and organization. In the example provided, the auditor might identify the activity "checked and approved" performed for variation 2 as a mitigating control activity for the activities "final check of invoice" and "approve invoice" performed for variation 1 and conclude that the two process variations represent routine process paths that do not indicate a risk of material misstatement.

However, provided the multitude of resulting process variations in practice, the manual comparison of two different process traces was not perceived as an appropriate instrument to conclude on the appropriateness of the individual process executions. In exploring their auditee's data with process mining, all audit teams further stated that additional guidance is needed regarding the analysis of process variations.

When compared to the most frequent process variation, changes in the ordering of activities, repetitions of individual activities or missing or additional activities lead to a new process variation. As such, among other factors, the number of process variations is influenced already by the number of activities in the process graph. A large number of variations does not necessarily lead to an increased risk of material misstatement. However, if process mining techniques to not consider the financial perspective, that is, the impact of an individual process execution on the financial accounts, the concept of materiality traditionally applied by the auditor in assessing risks and identifying misstatements is not applicable to the analysis of variations. Consequently, the auditor needs to apply professional judgment in evaluating whether the variations are reflective of the entity's business and in determining which variations to investigate in detail. To support this judgment beyond the provision of additional technical and methodical guidance, a new dashboard for evaluating different executions of the process has been integrated (Figure 34).

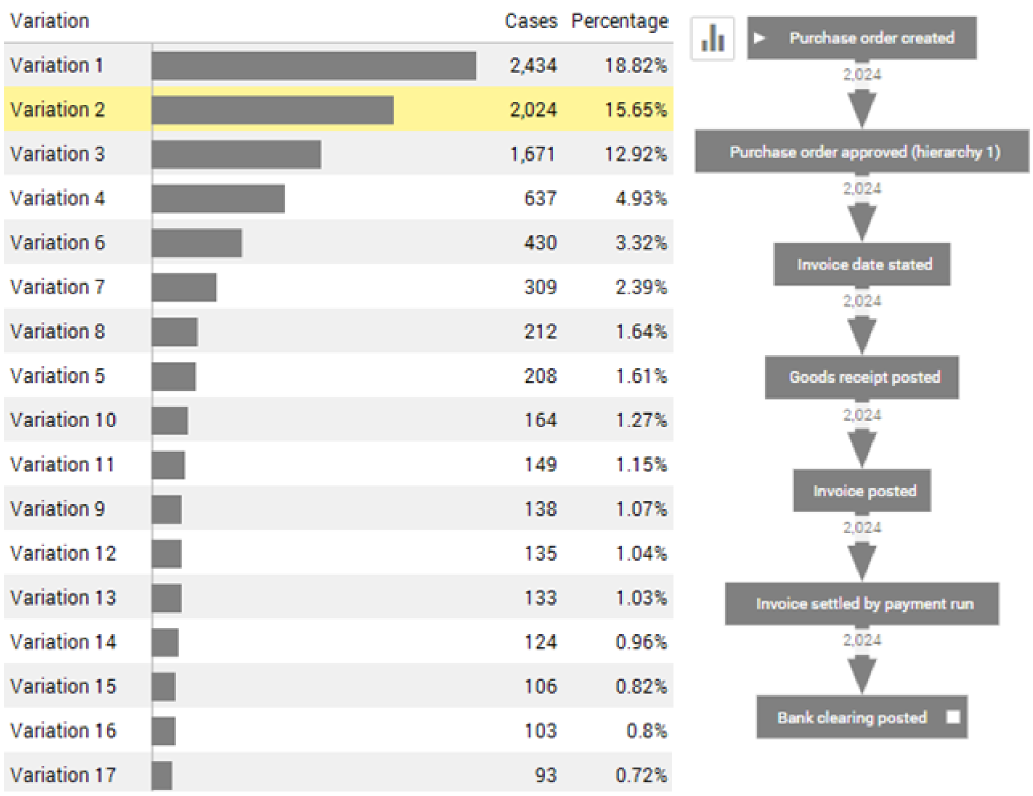


Figure 34: Critical path dashboard supporting the evaluation of process variations

The dashboard provides an overview of all process variations, the number of cases processed within a variation and the percentage of the total number of cases covered by an individual variation. Similar to the “Compare” dashboard presented in Figure 33, upon selecting an individual variation, the related process graph is visualized. In the example provided, the distribution of cases over the process variations shows that around 52 percent of all cases are already covered by the four most frequent process variations. A high frequency of cases indicates a routine process path that may be subject to different risks of material misstatement than non-routine process variations.

In preparation of the feasibility assessment, quantitative case information has been added to the process mining analyzer by integrating information on the amount of the purchase order line item processed within a specific case.[373] By this, the auditor may not only investigate the number or percentage of cases that have been processed, for example, by case type, material group or purchasing department, but also obtain information about the monetary value related to the case. Based on the audit teams’ feedback, the invoice amount related to a case has been added as additional case attribute

[373] Cf. Chapter 3.1.1.

to the case table[374], enabling to enhance the analysis of process variations by information about the total invoice volume processed (Figure 35).

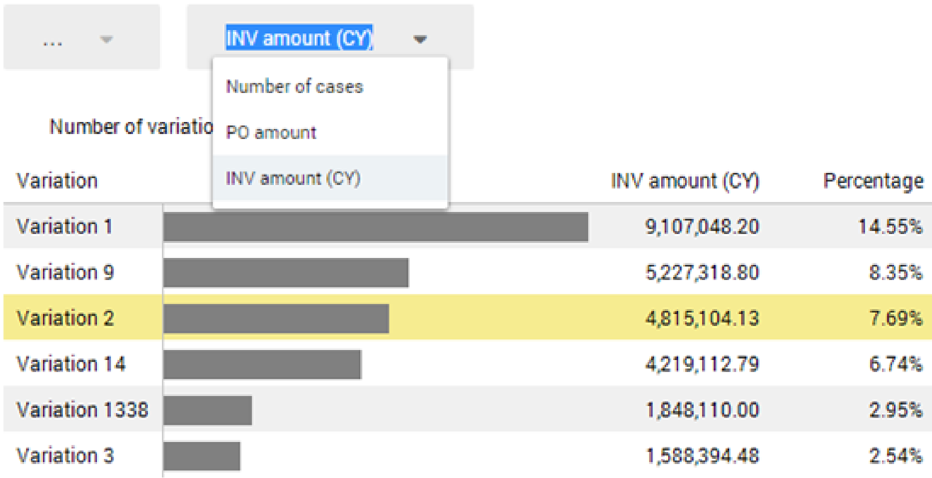


Figure 35: Analysis of process variations based on the total invoice amount

Depending on the scope and design of the data extraction, a process variation may include events from more than one period, for example, if the purchase order has been created in the prior period, but the related goods and the invoice have been received and recorded in the audit period under review.[375] To display the invoice amount relevant to the audit period under review, in the analysis presented in Figure 35, the calculation of the invoice amount related to a variation has been limited to invoices recorded in the audit period. Information on the number and invoice amount related to a variation may support the auditor in identifying routine and non-routine paths of the process and assess related risks of material misstatement. For example, from the excerpt of process variations shown in Figure 35, the auditor may identify variation 1338 as the process variation containing the fifth highest total invoice amount. Variation IDs are assigned according to the number of cases processed, starting from variation 1 with the highest number of cases.[376] In comparing the invoice amount processed in a specific variation to the number of cases, the auditor may conclude that variation 1338 is used to process a small number of high-volume transactions separate to the routine transactions and thus may involve a higher risk of material misstatement.

At (4): Integration of vendor-related analyses

In understanding the entity and the environment in which it operates, the auditor obtains an understanding of those the entity conducts business with, including third-party

[374] Cf. Table 2.
[375] The data extraction strategy and its impact on the analyses is discussed in Chapter 4.1.6.
[376] Cf. Figure 34.

vendors and related parties.[377] If the resource information in the process mining input data includes information on the vendor related to a case, the technology supports analyzing the types and structure of suppliers and changes therein compared to the previous audit period. To support the auditor in identifying risk factors and determining risks of material misstatement, a new menu page for analyzing supplier relationships has been integrated. Figure 36 shows the number of cases by vendor.

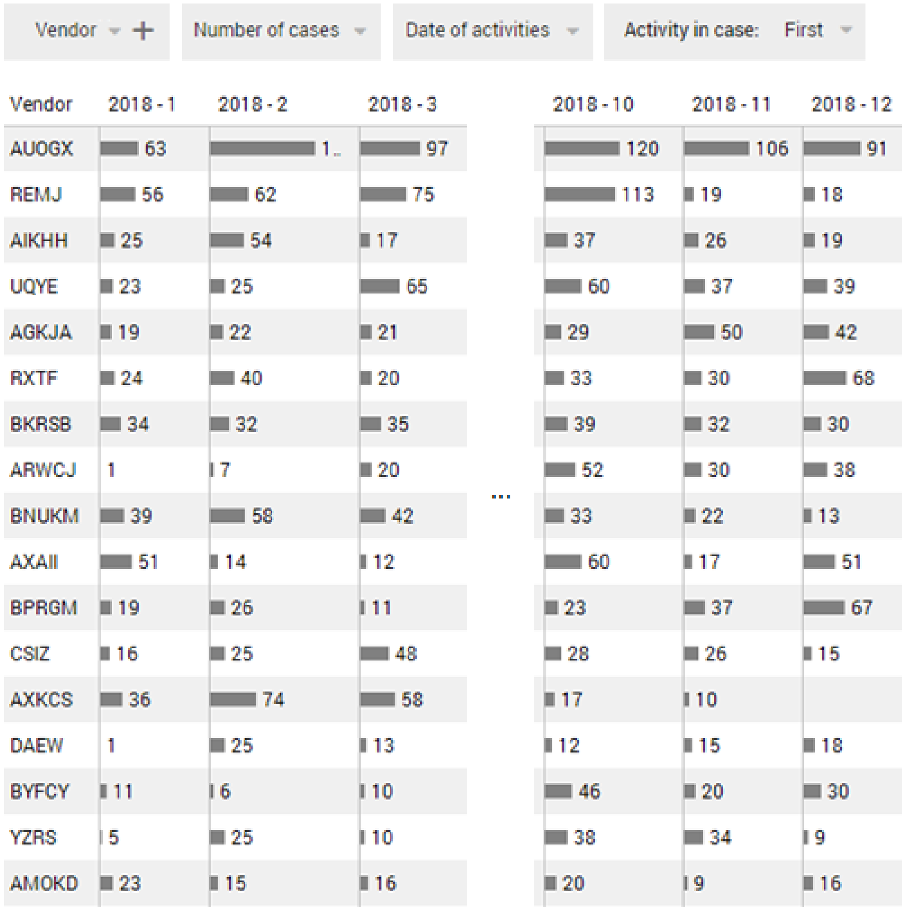
Vendor		▼ Cases	Percentage
AUOGX	1,183	1,183	9.4 %
REMJ	706	706	5.6 %
AIKHH	484	484	3.8 %
UQYE	472	472	3.7 %
AGKJA	401	401	3.2 %

Figure 36: Distribution of the number of cases by vendor

The metric of the analysis may be adjusted to investigate the purchase order or invoice amount related to a specific vendor instead of the number of cases, respectively. The analysis supports the auditor in understanding the vendor base and identifying key vendors based on the number of cases or, alternatively, the related invoice amount processed throughout the period.[378] The auditor may consider additional attributes in the analysis to support identifying risks of material misstatement. For example, adding information on the purchased material supports the auditor in understanding the nature of transactions conducted with key vendors. Information on significant purchasing volumes related to materials only purchased from a single vendor may indicate vendors the entity is highly dependent on. Information on the process variations related to specific vendors supports the auditor in identifying vendors related to transactions outside the normal course of the business that may be subject to different risks of material misstatement. Combining the vendor information with the "Booking pattern" analysis[379] introduced earlier in this chapter further supports the auditor in identifying related parties whose transactions are processed using trade payables intercompany accounts.

[377] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A57, A68 and Appendix 1; IFAC (2021), ISA 550, para. 9; Chapter 2.3.3.
[378] Cf. Chapter 2.3.3.
[379] Cf. Figure 29.

When performing vendor-related analytical procedures, the auditor usually disaggregates the data over time to identify significant changes or fluctuations in the business activity throughout the period or when compared to the prior period. Process mining may support understanding the vendor-related business activity throughout the period by showing the monthly development of cases by vendor (Figure 37).



Vendor	2018 - 1	2018 - 2	2018 - 3	…	2018 - 10	2018 - 11	2018 - 12
AUOGX	63	1..	97		120	106	91
REMJ	56	62	75		113	19	18
AIKHH	25	54	17		37	26	19
UQYE	23	25	65		60	37	39
AGKJA	19	22	21		29	50	42
RXTF	24	40	20		33	30	68
BKRSB	34	32	35		39	32	30
ARWCJ	1	7	20		52	30	38
BNUKM	39	58	42		33	22	13
AXAII	51	14	12		60	17	51
BPRGM	19	26	11		23	37	67
CSIZ	16	25	48		28	26	15
AXKCS	36	74	58		17	10	
DAEW	1	25	13		12	15	18
BYFCY	11	6	10		46	20	30
YZRS	5	25	10		38	34	9
AMOKD	23	15	16		20	9	16

Figure 37: Distribution of the number of cases by vendor and month

As the activities performed within a case may fall into different months, a case may not be allocated exclusively to a specific month based on the timestamps of its events. As such, by default, the analysis enables to determine for each vendor how many cases have been initiated in a specific month by allocating the cases to the month where the first activity has been performed. Alternatively, the cases may be counted in the month where they ended (because no additional event happened in the period investigated). Based on the understanding obtained on the cycle times of the process and throughput times between individual activities, the auditor may also determine to allocate the cases based on an individual activity, for example, to the month the (first or last) invoice has been posted for a case. Significant changes (or the absence of changes) identified in the processing of cases may be corroborated with the auditor's expectation based on the understanding of the entity and its environment, including the entity's business model.

The last modification made to the process mining analyzer after the initial feasibility assessment is the integration of the vendor master data and the changes made to this data. The continuous and consistent maintenance of master data is an integral part of the effective and efficient processing of transactions. The vendor master data contain descriptive and controlling-relevant information about the vendors the entity is conducting business with, including:

- general data on vendor level (for example, the address, VAT number and bank details),
- accounting data on the level of the company code (for example, control accounts, dunning procedures and payment methods) and
- purchasing data (for example, the minimum ordering volume and payment terms).[380]

In SAP, the purchasing department may only create a purchase order if all vendor-related purchasing data is available in the master data. As such, a large number of master data entries without maintained purchasing data may indicate a risk related to the existence of creditors.[381] Similarly, an invoice received may only be recorded by the accounting department if the accounting perspective is maintained.[382] Figure 38 shows the vendor master data that has been exported from the ERP system and integrated into the process mining analyzer.

[380] Cf. HARTKE, LARS/HOHNHORST, GEORG/SATTLER, GERNOT (2010), p. 345.
[381] Cf. ibid., p. 340.
[382] Cf. ibid., pp. 344f.

Vendor | Vendor ID | Country | Terms of payment | +

Number of cases

Vendor	Vendor ID	Country	Terms of payment	▼ Cases
AUOGX	0120001770	A	Z060 : Within 60 Days Due Net	1,510
REMJ	0320000082	A	Z030 : Within 30 Days Due Net	723
	0320000966	HP	Z014 : Within 14 Days Due Net	0
	0320002205	A	Z045 : Within 45 Days Due Net	0
	0320000362	HP	Z014 : Within 14 Days Due Net	0
ARWCJ	0120006355	F	Z000 : Pay Immediately W/o Deduction	515
AIKHH	0120002931	A	Z030 : Within 30 Days Due Net	500
UQYE	0120000094	J	Z000 : Pay Immediately W/o Deduction	492
BNUKM	0120000065	J	Z025 : Within 25 days Due Net	479
AXAII	0120006432	CN	Z000 : Pay Immediately W/o Deduction	475
AXKCS	0120000216	BM	ZZ22 : Cash against Documents	436
AGKJA	0120000082	A	Z000 : Pay Immediately W/o Deduction	429

Figure 38: Vendor master data dashboard of the process mining analyzer

By reviewing the master data table, the auditor may confirm the completeness and consistency of the general vendor data, the accounting data and the purchasing data for each vendor the entity conducts business with. To facilitate identifying vendors for which transactions have been recorded in the period under review, the master data table is enhanced by information on the total number of cases related to each vendor.

Appropriately maintained vendor master data reduce the effort in recording transactions and represent an important control related to the completeness and accuracy of transaction processing in purchase to pay.[383] Risks of material misstatement may arise, for example, from a missing segregation of duties between the master data maintenance and the processing of purchase orders, from frequent changes made to payment-related master data or from the concurrency of master data changes and transaction processing. Consequently, as part of the risk assessment procedures, the auditor obtains an understanding of the entity’s policies and procedures designed and implemented with regard to the creation, modification and deletion of master data. Figure 39 shows the analysis of master data changes designed to support the auditor in understanding the master data change process, identifying risks of material misstatement and identifying and evaluating related control activities.

[383] Cf. HARTKE, LARS/HOHNHORST, GEORG/SATTLER, GERNOT (2010), p. 340.

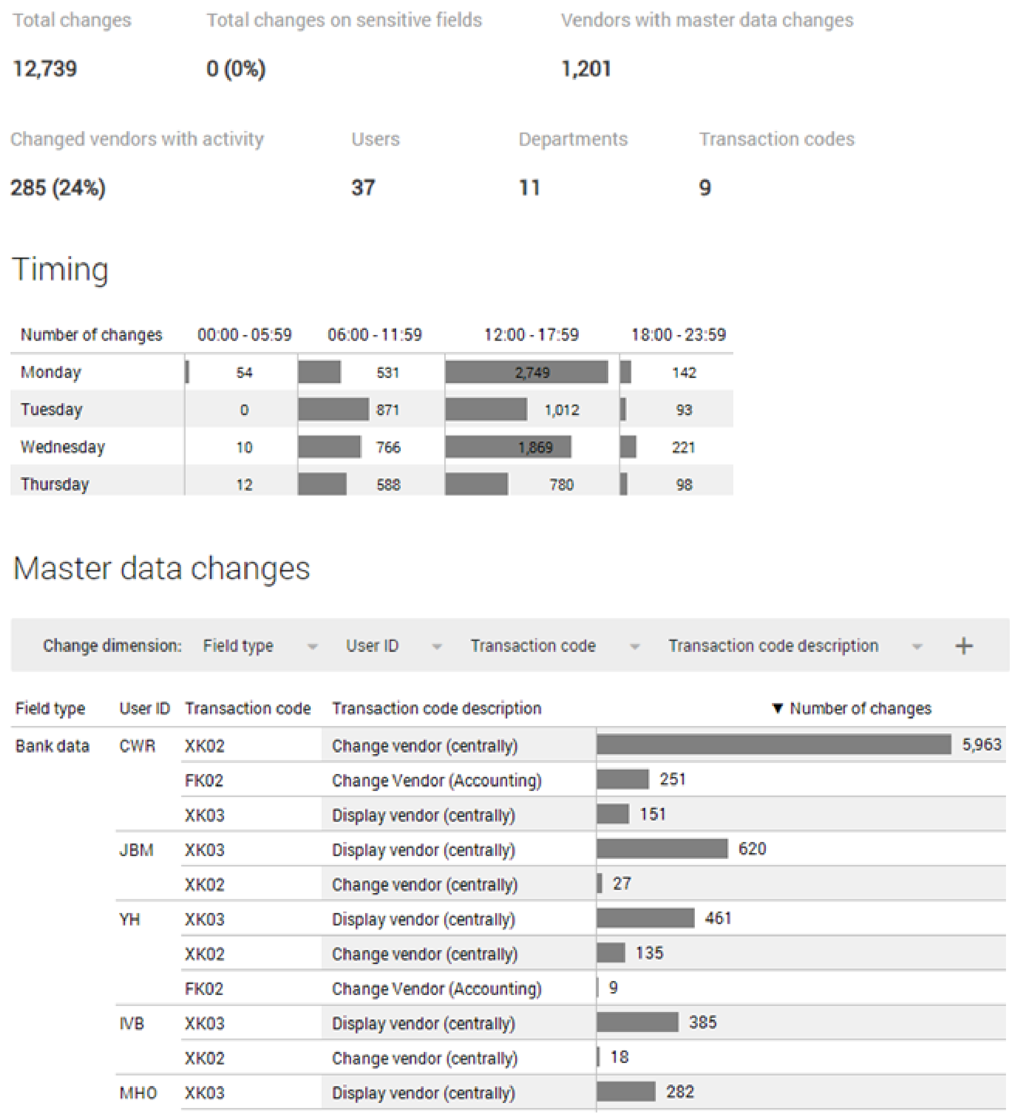


Figure 39: Analyzing master data changes using process mining

The KPIs at the top of the dashboard provide a summary of changes made to the vendor master data, including the total number of changes, the number of individual vendors changed, the number of users and departments involved and the transaction codes used. If the entity customized the master data by marking specific data fields as sensitive, the KPIs separately summarize the number of changes made to sensitive fields. When a sensitive field is changed in SAP, the related vendor is automatically blocked for payment-related activities until a second authorized individual reviews and approves or rejects the modification of the data.[384] As such, sensitive fields are frequently configured for bank data information, enforcing the adherence to the four-eyes principle when changes are made. If the entity makes use of sensitive fields, the auditor may filter the respective KPI and identify the related fields from the table at the bottom of

[384] Cf. HARTKE, LARS/HOHNHORST, GEORG/SATTLER, GERNOT (2010), p. 353.

the dashboard that includes the field types that have been changed throughout the period.

As changes made at unusual times may indicate a risk factor according to ISA 240, the dashboard further includes an analysis summarizing the changes made by the day of the week and the timestamp information. In analyzing the time perspective, the auditor may filter master data changes made outside the normal business hours of the entity for further investigation.

The table at the bottom of the dashboard supports obtaining an understanding of the master data changes by summarizing the changes by the type of the field that has been changed, the user initiating the change and the transaction code used. The auditor may add the vendor information to the analysis to investigate master data changes made to key vendors identified.

The analyzer further provides the detailed list of vendor master data changes, including the timestamp, user and transaction code information, the vendor, the field that has been changed as well as the old and the new value of the field. If the entity does not make use of the dual control for sensitive fields or the auditor does not consider the sensitive fields specified as being sufficient for preventing unauthorized modifications of master data, the auditor may manually analyze the employee’s involvement in the master data change process by:

- analyzing the roles and responsibilities of users involved in master data changes by investigating the transactions codes used by individual users (for example, the use of the transaction code for removing a payment block should be restricted to authorized individuals only),
- investigating the organizational segregation between the master data maintenance and transaction processing (by comparing the users modifying master data throughout the period with the users processing transactions by executing activities in the process) and
- determining whether appropriate segregation of duties between changes related to the purchasing perspective and the accounting perspective of the vendor master data is in place.

Appendix III summarizes the structure and key analyses of the process mining application used for piloting in 2018.

3.2.2 Audit team feedback and results from quality assurance procedures

Feedback obtained from the piloting project

Feedback interviews have been conducted with all pilot teams in the period between April 2018 and June 2019. The feedback received relates to:

(1) the process of extracting, transforming and loading the data,
(2) the integration of the procedures performed with process mining into the audit approach,
(3) the tool handling and tool performance
(4) the analyses provided in the analyzer and
(5) any additional comments from the audit teams.

The feedback obtained within each category summarizes as follows:

At (1): Feedback on the ETL process

- Similar to the feasibility assessment, most of the audit teams highlighted the extensive involvement of IT specialists and the detailed understanding of the entity's IT landscape required for setting up the process mining analyzer.
- All audit teams stated that the process of extracting and transforming the data took considerably more time when compared with other automated tools and techniques used in the audit.
- In some instances, the data model has been customized to include further activities with regard to (1) the scanning and release of invoices using preprocessing software outside the SAP system, (2) purchase requisitions in the SAP Supplier Relationship Management (SRM) workflow and (3) the invoice workflow for SAP FI invoices.
- Many audit teams requested additional guidance on the standard data model and customization options available. One audit team challenged why there is no comprehensive list with all activities for the default configuration of common ERP systems as well as existing customization options available, enabling the auditor to select the activities applicable to the entity's IT landscape.
- Some of the audit teams fed back that when analyzing the initiating activities in the process, cases have been identified that do not start with the creation of a purchase order. However, when exporting these cases and inquiring the entity to identify the root-cause, a purchase order document was provided. One team

determined that these cases relate to purchase orders for framework agreements that are used over a long period of time and identified that some of these orders are not included in the data.

At (2): Feedback on the integration into the audit approach

- All audit teams highlighted that process mining contributed to a more detailed process understanding and provided insights in the process execution on a level that has not been achieved using the traditional approach of auditing processes and controls. Most of the teams mentioned the high complexity of the overall process graph that was not reflected in the entities' process narratives and not evident in the audit teams' documentation of their traditional walkthrough procedures. Figure 40 shows an excerpt of the purchase to pay process graph from a retail company where the number of activities and edges that are displayed is set to the maximum.

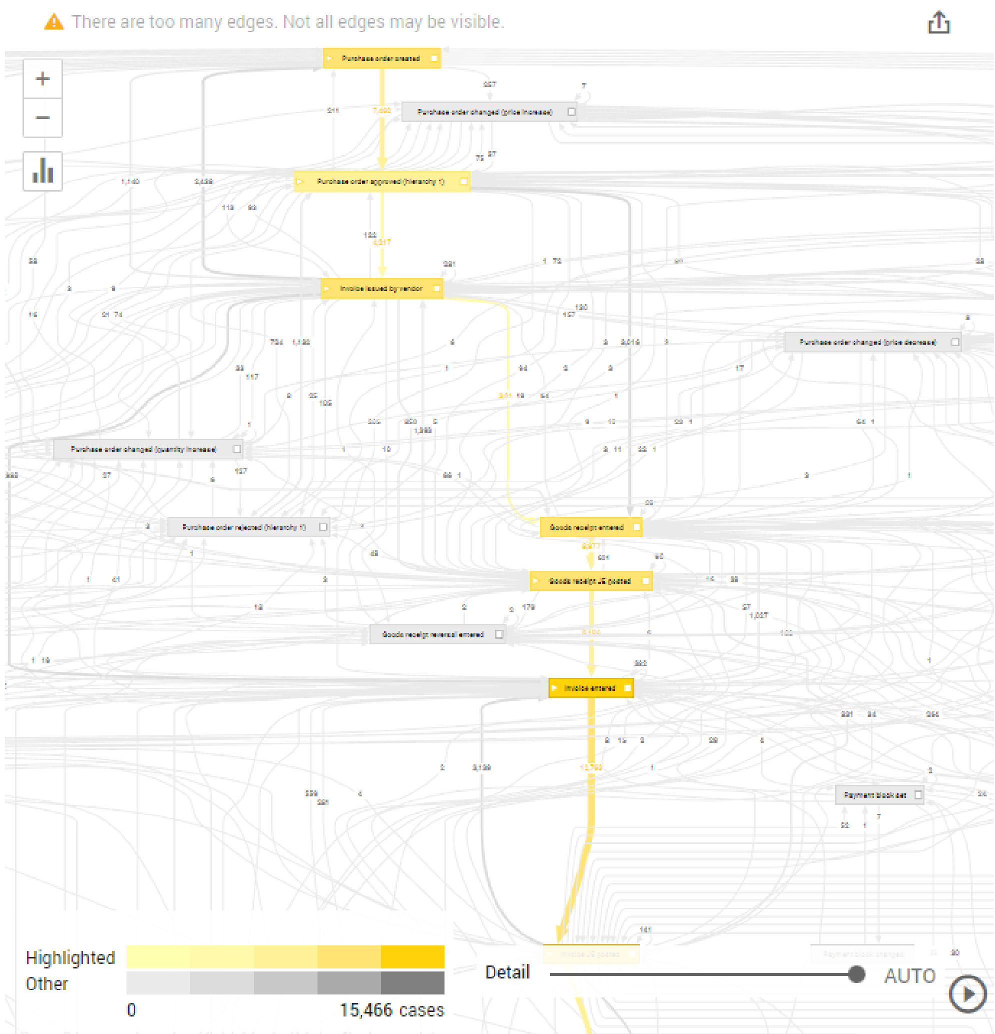


Figure 40: Excerpt of the overall process graph of a retail company

Although the retailer's process does not include an extensive number of cases and activities, the warning message at the top of the process graph indicates that the process includes too many edges for the process mining application to visualize them in their entirety.

- However, even provided the complexity of the overall process graph, most of the teams fed back that process mining supports the development of an audit plan for the analyzed process, including the identification of inherent risks at the assertion level and related controls addressing these risks.
- One of the audit teams stated that besides supporting the understanding of the process, process mining improved the approach to select transactions along the critical path of the process that do not match the audit team's expectation with regard to an appropriate initiation, recording, processing and reporting of transactions in the business process. Using the three-way-match analysis performed for the entire population of transactions instead of a sample only as an example, the team pointed out that using process mining to identify and test internal controls would significantly increase audit quality.
- Another audit team confirmed that auditing processes and controls based on the entire population of transactions rather than a sample only is perceived as a natural extension of the data-driven audit approach that is already established in other areas of the audit. However, in exploring how to perform tests of controls with process mining, the team highlighted that control testing with process mining may contradict the evidence obtained based on a sample only. The team remarked that they did not identify any exceptions in a sample of purchase orders selected to test the purchase order approval control, but the test based on the entire population showed that the control has not been performed for some transactions.
- One of the teams experienced that the overall time effort when using process mining was shorter than the time spend with traditional procedures, especially with regard to risk assessment procedures. Audit teams that have already been exploring the use of process mining in the previous audit period confirmed that with the previous experience, the information on the data model and the technical and methodical guidance available, the efficiency of the process analysis increased significantly.

- The audit teams agreed that the redesigned reconciliation[385] between the process data and the general ledger data helps to identify flows of transactions related to purchase to pay. This includes any significant parts of the process that are not covered by the process mining application as the transaction volume is not related to a case. However, for some of the audit teams, the portion of the transaction movement on the financial accounts that was covered by the process data was quite low. Non-covered transactions related to the purchase to pay process included, in particular, transactions that do not require a purchase order. As direct purchases through the SAP FI module are not related to a purchase order, they are not included in the process. The teams noted that in case of low coverage of the transaction volumes on relevant financial accounts by the process data, process mining may not replace the traditional procedures performed to understand the process and identify and address risks of material misstatement but is performed in addition to these procedures, decreasing audit efficiency. The same applies to manual parts of the process that are relevant to the audit and thus need to be audited alternatively as they are not evident in the process data.
- The audit teams again stated that more guidance is required on how to handle the extensive number of different process variations, including examples of factors that may support the auditor's professional judgment. The new dashboard supporting the analysis of variations,[386] including the integration of the invoice amount processed over a specific process path, was perceived positively. However, teams have been unsure on the absolute and relative thresholds to apply when evaluating the number of cases or invoice amount processed over a specific process path.
- One team specifically challenged whether there must be a "process materiality" supporting the quantitative evaluation of process variations. Another team stated that for the qualitative assessment of specific paths through the process, the variations should be summarized by different criteria. For example, the team was not interested in the sequence of activities that have been performed in a variation but rather in whether an activity has been performed at all. According to the team, grouping the variations by their activities while

[385] Cf. Chapter 3.2.1.
[386] Cf. Figure 34.

disregarding the ordering of the activities within the process trace would reduce the total number of variations to evaluate significantly.

- More detailed guidance on the analyses to perform in order to gain sufficient audit evidence was requested. One team fed back that the guidance should include the assertion and risk covered by a specific analysis. The team further requested a description of what is specifically not covered by an analysis, i.e., the audit procedures that need to be performed in addition to the process mining analyses.
- It was not always intuitive whether an audit procedure performed with process mining is an analytical procedure used as risk assessment procedure, a test of control or a substantive procedure. One of the teams indicated that the procedures to identify and address risks of material misstatement sometimes merge when using process mining. While the team confirmed that reasonable assurance with regard to the identified risks has been obtained with the procedures performed, the concern was raised that the distinction of the nature of the procedures may be important for applying materiality and for documentation purposes.

At (3): Feedback on the tool handling and tool performance

- The audit teams did not face any challenges regarding the general application handling such as navigating through the analyzer, using the filter functionalities provided and customizing the analyses to the objective of the audit procedure.
- Most of the audit teams did not notice any performance issues. One team mentioned that the initial loading of the process mining application was quite long for large datasets, however, after the application was loaded, no further issues regarding the performance were identified.
- Two audit teams mentioned that sometimes when switching dashboards in the analyzer, it is not intuitive which filters are carried over and which filters are removed from the data. For example, filters got lost when switching from the "Financial reconciliation" menu page to the process graph. Vice versa, the concern was raised that it is not possible to filter a specific variation in the process and investigate the general ledger accounts involved in the transactions related to this process trace.

At (4): Feedback on the analyses provided in the analyzer

- Positive feedback was provided on the analyses specifically dedicated to the understanding of the vendor structure[387] and changes therein according to ISA 315. However, the vendor analysis[388] was difficult to interpret when compared with an analysis of the vendor subledger data. The audit teams specifically struggled in analyzing vendor activity based on the total invoice amount by vendor, as the amount is not reconcilable to the transactions recorded in the trade payables subledger.
- The booking pattern analysis[389] was perceived as useful instrument to understand both the activities in the process and the related recording of transactions on the general ledger accounts, facilitating to understand the critical path of the process and identify, for example, risks of material misstatement that may be related to transactions flows involving seldom used financial accounts.

At (5): Additional feedback received from the audit teams

- Every audit team faced a number of cases that relate to long-standing purchase orders that are used for many transactions and frequently include a large number of (partial) deliveries and invoices. These purchase orders significantly increase the complexity of the process and, due to the large number of events used in processing the related cases, each long-standing purchase order frequently has a unique process trace and thus its own process variation. As a consequence, some teams requested a separate analysis to isolate, deconstruct and analyze long-standing purchase orders to reduce the overall process complexity.
- The interpretation of activities in the analyzer was not always intuitive. One team challenged that, for example, without studying the extensive technical guidance, it is not clear if the activity “invoice settled by payment run” refers to the initiation of the payment run in the IT system or the posting of the related journal entry on the general ledger accounts.
- Two teams fed back that a functionality to compare the process data with a comparative period and identify significant changes to the process (or the ab-

[387] Cf. Chapter 3.2.1.
[388] Cf. Figure 36.
[389] Cf. Figure 29.

sence of these changes) would significantly increase audit efficiency and support the use of process mining for audit teams with tight reporting deadlines at period end.

Feedback evaluation and results of quality assurance procedures

The second empirical application of process mining in the audit practice confirmed the potential of the solution to increase the quality of the process understanding and support both the auditor's risk identification and risk response procedures at the assertion level. While some of the initial concerns of the audit teams could be addressed by additional functionalities, guidance and appropriate training and coaching on using process mining, however, other methodical and technical challenges initially identified have been confirmed and additional obstacles have been identified.

Key challenges identified as part of the piloting and the quality assurance procedures performed relate to the ETL and delivery process, including the data model customization, as well as handling the actual process complexity that is not evident in the traditional process documentation.

ETL and delivery process and data model customization

Despite the audit firm's ongoing standardization of the data model for SAP and the continuing automation of the ETL process, all audit teams highlighted the extensive time consumed for data preparation and the detailed technical understanding of the process and the information system required, going far beyond the prerequisites for other data analytics solutions applied in the field of auditing. Expert support was needed especially with regard to understanding and customizing the data model to account for additional activities used in custom workflows in the SAP system. The feedback on expected activities that are missing in the application confirms that additional guidance is required, including the data extraction period and its implications on the data and events (not) available in the analyzer. The feedback obtained regarding the lack of guidance and the efforts in understanding and customizing the process mining data model are partly related to the fact that today's auditor is no IT auditor. Although the auditor is required to obtain an understanding of the information system relevant to the audit,[390] auditors are primarily educated in business economics and not in auditing IT systems. In an audit of financial statements, the entity's IT systems and

[390] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 25.

integrated application controls are frequently audited by a team of domain experts with specialized technical knowledge.

Both a missing audit background of the IT specialist preparing the data as well as lacking technical knowledge of the audit team may lead to quality issues when applying process mining in the field of auditing. For example, some audit teams decided to include multiple business units in the same process mining analyzer due to their similar process design and the same users involved in the process. However, in many cases, the business units were located in different countries with different local currencies. Thus, for an appropriate representation of the metrics in the analyzer, the default analyses need to be customized to distinguish between different currencies when summarizing, for example, the total purchase order or invoice amount within a case or a process variation. Another example involves the number of activities included in the analyzer. While an activity may contribute to the understanding of the actual process execution, each additional activity increases the complexity of the overall process graph and the number of different process executions. In consequence, there is a trade-off between the comprehensiveness and the complexity of the process. Quality assurance procedures conducted by the author of this thesis showed that some audit teams included a lot of activities referring to the same process step, for example, the approval of a purchase order (the activity "purchase order approved") was separated into individual activities for the approvals on different levels of the authorization hierarchy (for example, "purchase order approved - hierarchy level 1"). Instead of multiplying the "purchase order approved" activity with the number of hierarchy levels, the information on the hierarchy level of the individual approving the document may also be stored as event attribute of the "purchase order approved" activity. By carefully evaluating which information – and which activities – are required to evaluate the effectiveness of the purchase order approval control, the complexity of the overall process may be reduced while the information relevant to the audit is kept.

Another example relates to the behavior of process mining when working with analyses that are based on different data sources. Some audit teams tried to analyze the processing of cases that are recorded on a specific general ledger account by filtering the account in the trial balance and then switching to the process graph for further investigation of the related transactions. However, the trial balance used for the financial reconciliation is based on the journal entry information, i.e., the financial data, while the process graph is using the process data, i.e., the purchase order information that defines an individual case. As both data sources are not linked by default within the SAP system, the links between the data fields need to be manually created and any filters used on data fields not available in one of these sources are lost when switching between both analyses. As such, depending on the integration of the underlying data sources, the behavior of process mining analyses may not always be intuitive. If the auditor is not aware of the characteristics of the data model, unexpected behavior of the tool, such as the loss of filters, may go undetected when performing the analyses but impact achieving the audit procedure's objective.

The examples demonstrate that preparing the data and using process mining in an audit of financial statements requires both in depth technical domain knowledge and an understanding of the data model's implications on the audit procedures that may or may not be performed. The findings with regard to data quality, including the mixture of different currencies, indicate that a detailed review of each process mining application is required before it is used in an audit of financial statements. Further, besides appropriate technical and methodical training of the IT team preparing the data and the audit team using the process mining application, challenges identified regarding the behavior of different analyses depending on their data sources (and thus the design of appropriate audit procedures with process mining) necessitate a dedicated quality assurance process for the audit procedures performed using the application.

Integration into the audit approach

Most of the feedback on the integration of process mining into the audit approach relates to the complexity of the overall process and its implications on different audit procedures. The level of insights in the process execution obtained with process mining was significantly higher than the understanding obtained from traditional walkthrough procedures. Figure 41 illustrates a comparison of the flowchart usually drawn by the auditor as part of understanding a business process with the process graph reconstructed with process mining.

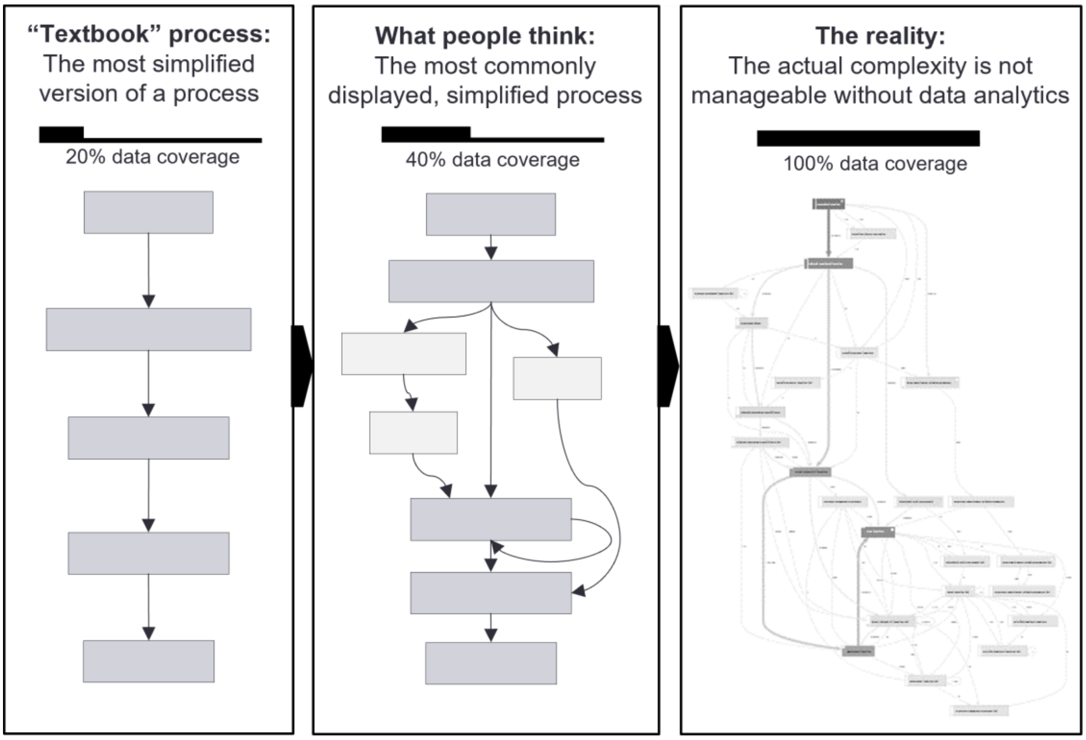


Figure 41: The complexity of "textbook" processes compared to real processes

The process graph on the left only shows the key activities within the process. It is limited to the most frequent process variation that covers around 20 percent of the total number of cases included in the analyzer. This process usually complies with the initially designed target process that is frequently evident in process handbooks or narratives available at many companies. The graph in the middle shows a simplified process that covers around 40 percent of the population of cases. The level of detail may be compared to the auditor's flowchart supporting the understanding of the critical path of the process and the identification of risks of material misstatement. The process on the right also includes seldom-used process paths. It covers the total population of cases and shows the actual process complexity resulting when applying process mining on real-life business processes. The feedback and quality assurance procedures performed during the piloting, including comparing the process picture obtained with process mining with the documentation of the critical path of the process from previous audit periods, confirm that business processes are usually not executed as described in companies' process narratives or as documented as part of the auditor's walkthrough procedures.

As already identified as part of the feasibility assessment,[391] one of the key challenges in handling the process complexity in an audit of financial statement is evaluating the resulting process variations, i.e., obtaining reasonable assurance that the different ways transactions are processed throughout the period do not result in a material misstatement of the financial statements. When asked about the evaluation of process variations, one audit team stated that they analyzed so many variations that the rest was below the performance materiality determined for the general ledger accounts involved in processing the transactions. They applied professional judgment by not investigating every sequence of activities but only those that had larger invoice amounts. The total invoice amount processed in a variation may provide an indication of the magnitude of a misstatement that may result from a risk of material misstatement related to the process path. However, it is no indicator for the completeness of transactions, as weaknesses in the process execution may result in invoices not being recorded. As such, the analysis of variations is subject to both quantitative and qualitative criteria, which may include, for example, seldom-used process paths that may be used for processing transactions outside the normal course of the business[392] or variations with missing control activities such as the purchase order or invoice approval. All audit teams agreed that due to the large amount of resulting process variations they may not be audited in their entirety in practice. However, the extent of variations that have been investigated and evaluated in detail by the audit teams varied significantly across the pilot engagements. While one of the teams focused on process variations above tolerable error, two of the audit teams investigated both the ten variations with the highest invoice amount and the highest ten variations according to the number of cases processed. Other teams analyzed the variations up to a certain percentage of their coverage of either the total number of cases or the total invoice amount. These findings suggest that the auditor needs to be supported with appropriate techniques for identifying and addressing any risks of material misstatement that may result from the different ways the process is executed.

The feedback obtained from the second wave of piloting has been evaluated and addressed in the next version of the analyzer that has been approved for global use and widely implemented across different regions and engagements of the audit firm.

---

[391] Cf. Chapter 3.1.2.

[392] Cf. IFAC (2021), ISA 240, para. 33(c).

3.3 Implementation of process mining in the audit practice in Europe, Middle East, Africa and Japan in 2020

3.3.1 Global certification and scope of the implementation

Global certification of the audit firm's process mining solution

Based on the results of the feasibility assessment and piloting of the process mining technology in an audit of financial statements, the audit firm decided to conduct a global certification of the process mining analyzer built for the purchase to pay process. The certification process ensures that all globally certified solutions (1) have been tested and independently reviewed to withstand regulatory inspection, (2) provide audit teams with appropriate technical and methodical guidance and (3) have an appropriate operating model in place for the software or infrastructure required to use the solution. The firm's global certification enables audit teams to use the insights and analysis results obtained with process mining as audit evidence instead of applying process mining in addition to the traditional audit procedures.

The certification considered the results from the different waves of piloting and included a review of both the back- and front-end of the process mining application and the technical and methodical guidance supporting the audit teams in obtaining sufficient and appropriate audit evidence. Over the course of the certification process, both enhancements of the methodical guidance and adjustments of the process mining application have been identified and implemented. Clarifications of the methodical guidance on how to perform audit procedures with process mining were made especially regarding the reconciliation of process and financial data, the execution of professional judgment when evaluating process variations and the required level of detail when documenting the procedures performed using process mining.

Further, a far-reaching decision has been made about the integration of the procedures into the audit approach by limiting the implementation of process mining to engagements where the audit team follows a substantive audit approach. The predominant use case for process mining in the field of auditing identified in related research is evaluating internal controls.[393] However, besides a few theoretical examples,[394] no study has been identified that specifically included a holistic theoretical and empirical evaluation of the technology's appropriateness to support testing the design and

[393] Cf. Chapter 2.2.
[394] Cf. CHIU, TIFFANY/BROWN-LIBURD, HELEN/VASARHELYI, MIKLOS A. (2019), pp. 56f.

operating effectiveness of internal controls over financial reporting. Although the audit firm's feasibility study and the piloting conducted in advance to the certification confirmed the technology adds value to auditing and may help to identify and evaluate specific control activities, the firm determined that a broader empirical exploration of the suitability to test the design and operating effectiveness of controls is required. As part of the certification, it was decided to limit the subsequent implementation of process mining to a substantive audit approach, i.e., to using the technology for risk assessment procedures and to support substantive audit procedures addressing identified risks. Information obtained with process mining about the effectiveness of the entity's system of internal control may not be used as audit evidence.[395] Consequently, the implementation of the certified solution further excluded those audits performed in accordance with PCOAB auditing standards where the auditor expresses an opinion on the effectiveness of the auditee's internal controls over financial reporting.

The changes of the process mining application identified and implemented as part of the certification are discussed in more detail in the next section.

Certified process mining application used for the subsequent implementation

Based on the findings from the global certification and the piloting results several adjustments have been made to the process mining application over the course of the certification. Figure 42 shows the overview dashboard of the audit firm's certified process mining application implemented in 2020.

[395] The firm decided to further explore the use of process mining to test internal controls detached from the general implementation of the solution.

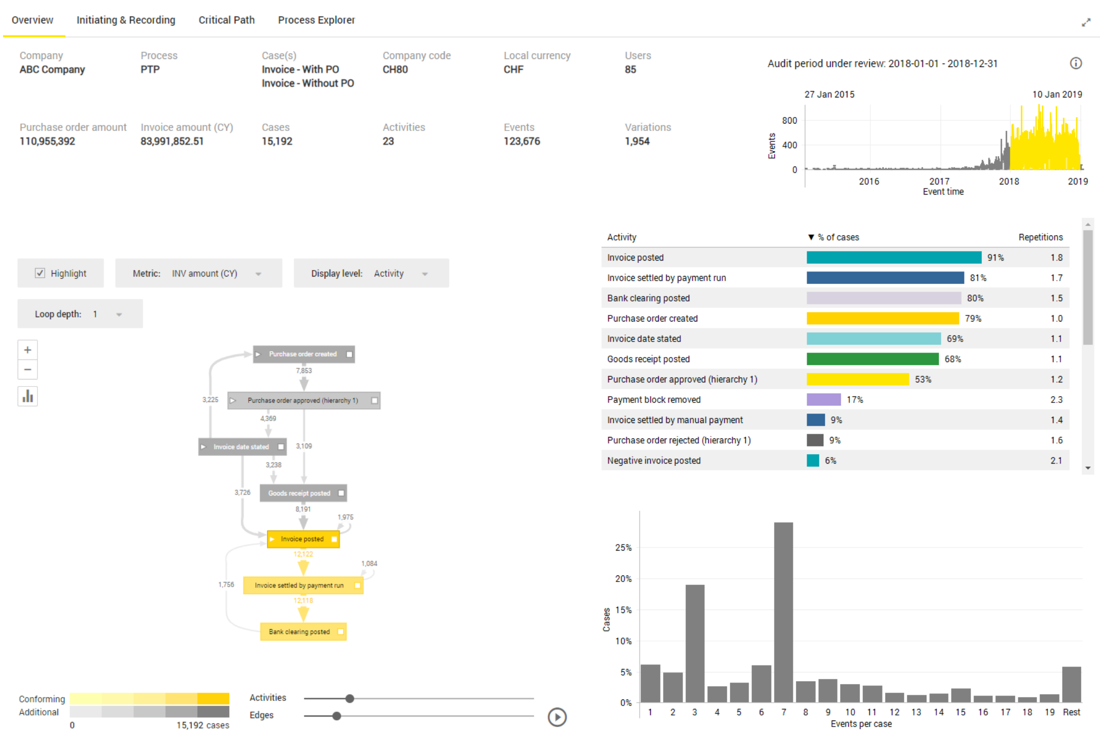


Figure 42: Overview dashboard of the process mining application in 2020

Major adjustments of the application made as part of the global certification include:

(1) the integration of SAP FI invoices as additional case type in the application,

(2) the implementation of an analysis supporting the identification of initiating and recording activities in the process and

(3) the flagging of CpD (conto pro diverse) vendor accounts.

At (1): Integration of SAP FI invoices

The initial version of the audit firm's process mining application only considered SAP MM transactions that are related to a purchase order. However, in practice, there are invoices in SAP that do not require a purchase order. These invoices are created directly in the financial accounting module of SAP and are typically used for the processing of transactions that do not expect a purchase order, goods receipt or service entry sheet, respectively (for example, tax or rental payments).[396] Limiting the process instance to the purchase order systematically excludes those invoices from the analysis that do not relate to a purchase order.

To increase coverage of trade payables and related accounts by the process data, a second case type has been integrated into the process mining analyzer that accounts

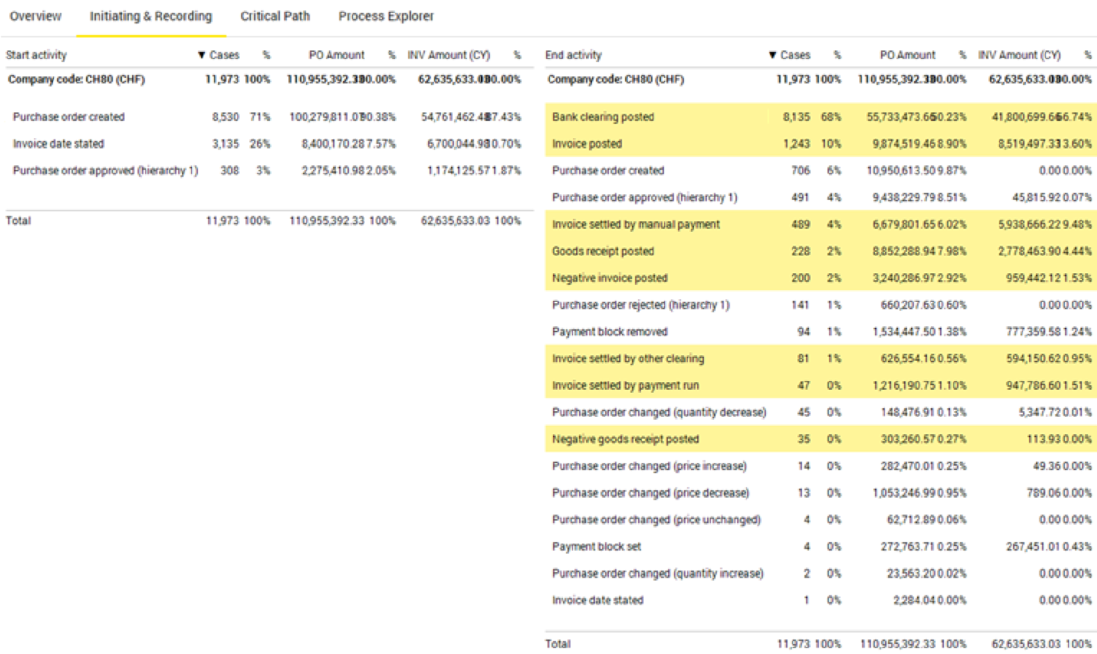
[396] Cf. Chapter 3.2.1.

for direct SAP FI transactions not involving a goods receipt or purchase order document. For cases that are related to a purchase order, a purchase order line item is the process instance. These transactions are processed through the SAP MM module. For direct purchases where no purchase order is available, the process instance is an invoice line item. These transactions are processed through the SAP FI module. As both types of transactions are usually related to different risks of material misstatement, the auditor may choose to deconstruct the significant class of transactions and perform the audit procedures separately for each case type.

At (2): Identifying initiating and recording activities

Understanding the flow of information through a business process involves understanding the initiation, recording, processing and reporting of transactions in the general ledger and the entity's financial statements.[397] To help evaluating the key elements of the flow of information without investigating the complex overall process graph[398] or each individual process variation, a new dashboard has been implemented.

The "Initiating & Recording" dashboard presented in Figure 43 supports analyzing the activities in the process in greater detail. It is used to understand the initiation of the process and to identify the activities leading to the recording of journal entries.



Start activity	▼ Cases	%	PO Amount	%	INV Amount (CY)	%
Company code: CH80 (CHF)	**11,973**	**100%**	**110,955,392.3**	**100.00%**	**62,635,633.0**	**100.00%**
Purchase order created	8,530	71%	100,279,811.0	90.38%	54,761,462.4	87.43%
Invoice date stated	3,135	26%	8,400,170.28	7.57%	6,700,044.9	10.70%
Purchase order approved (hierarchy 1)	308	3%	2,275,410.98	2.05%	1,174,125.57	1.87%
Total	11,973	100%	110,955,392.33	100%	62,635,633.03	100%

End activity	▼ Cases	%	PO Amount	%	INV Amount (CY)	%
Company code: CH80 (CHF)	**11,973**	**100%**	**110,955,392.3**	**100.00%**	**62,635,633.0**	**100.00%**
Bank clearing posted	8,135	68%	55,733,473.6	50.23%	41,800,699.6	66.74%
Invoice posted	1,243	10%	9,874,519.46	8.90%	8,519,497.3	13.60%
Purchase order created	706	6%	10,950,613.50	9.87%	0.00	0.00%
Purchase order approved (hierarchy 1)	491	4%	9,438,229.79	8.51%	45,815.92	0.07%
Invoice settled by manual payment	489	4%	6,679,801.65	6.02%	5,938,666.22	9.48%
Goods receipt posted	228	2%	8,852,288.94	7.98%	2,778,463.90	4.44%
Negative invoice posted	200	2%	3,240,286.97	2.92%	959,442.12	1.53%
Purchase order rejected (hierarchy 1)	141	1%	660,207.63	0.60%	0.00	0.00%
Payment block removed	94	1%	1,534,447.50	1.38%	777,359.58	1.24%
Invoice settled by other clearing	81	1%	626,554.16	0.56%	594,150.62	0.95%
Invoice settled by payment run	47	0%	1,216,190.75	1.10%	947,786.60	1.51%
Purchase order changed (quantity decrease)	45	0%	148,476.91	0.13%	5,347.72	0.01%
Negative goods receipt posted	35	0%	303,260.57	0.27%	113.93	0.00%
Purchase order changed (price increase)	14	0%	282,470.01	0.25%	49.36	0.00%
Purchase order changed (price decrease)	13	0%	1,053,246.99	0.95%	789.06	0.00%
Purchase order changed (price unchanged)	4	0%	62,712.89	0.06%	0.00	0.00%
Payment block set	4	0%	272,763.71	0.25%	267,451.01	0.43%
Purchase order changed (quantity increase)	2	0%	23,563.20	0.02%	0.00	0.00%
Invoice date stated	1	0%	2,284.04	0.00%	0.00	0.00%
Total	11,973	100%	110,955,392.33	100%	62,635,633.03	100%

Figure 43: Initiating and recording activities for cases related to a purchase order

397 Cf. IFAC (2021), ISA 315 (Revised 2019), para. 25(a)(i)a.
398 Cf. Figure 40.

As the processing of transactions throughout purchase to pay differs for SAP MM and SAP FI transactions, usually, the analysis of activities is performed separately for each case type. In the example provided, the analysis is performed for purchases relating to a purchase order. The table at the left side of Figure 43 lists the initiating activities, i.e., all activities that started a case. Common initiating activities within the purchase to pay process are the creation of a purchase order for SAP MM transactions and the posting of invoices for SAP FI transactions. If cases are identified that start with activities different to typical initiating activities in purchasing, the auditor may specifically consider the total number of cases and the total purchase order or invoice amount related to these cases. If these cases are numerous or material individually or when aggregated, they may indicate a separate significant class of transactions. On the other hand, unexpected initiating activities may also be caused by long-standing purchase orders.[399] As only events from up to four years before the balance sheet date are included in the analyzer,[400] long-standing cases might appear to start with another activity than the purchase order creation. In the dataset presented in Figure 43, three percent of all cases are starting with a purchase order approval. The auditor might filter these cases for further investigation and review the related case details, including the timing of events and the related source documents.

The table at the right side of Figure 43 includes the ending activities of the cases, i.e., the last activities within the cases as per the end date of the data extraction period. The highlighted activities represent the recording activities in the process. Typical recording activities include the posting of goods receipts and the posting and settlement of invoices and payments. In the example provided, most cases ended with the posting of the bank clearing or the invoice. As at period end not all cases have run through the complete process, there may be other activities included. For example, for some of the cases, a purchase order has just been created.

Based on the understanding obtained, the auditor may determine if the initiating and recording activities align with the understanding of the purchase to pay process or, in case they do not align, if they lead to a risk of material misstatement.

[399] Cf. Chapter 3.2.2.

[400] The data extraction period is discussed in more detail in Chapter 4.1.6.

At (3): Flagging of CpD vendor accounts

As a result of the feasibility assessment, an analysis of the entity's vendor structure has been added to the analyzer.[401] As part of the certification process, the vendor information has been enhanced with information on vendor accounts used as CpD accounts. These accounts are usually created to process transactions with multiple vendors that are infrequently used. For CpD accounts, the individual vendor's details are not included and maintained in the vendor master data. Instead, information such as the vendor name, bank details and payment terms are required to be captured when the invoice is entered into the information system.[402] As for CpD accounts the segregation of duties related to maintaining the vendor master data and entering invoices does not apply, these accounts are subject to increased risks of unauthorized or potentially fraudulent payments and should not be used for processing multiple transactions related to an individual vendor.

Scope and design of the subsequent implementation in 2020

After the global certification process was completed, in the period from June 2020 until March 2021, the audit firm conducted a broader implementation of the certified process mining solution for purchase to pay in Europe, Middle East, Africa and Japan. In total, 104 applications have been deployed to 47 audit teams of entities with varying sectors and sizes in 16 countries. The objective of the implementation was to validate the changes made to both the guidance and the application based on the feedback received during piloting and the adjustments made as part of the certification. The ETL process has been decentralized to the individual countries' digital and implementation functions. Each audit team further received technical and methodical execution support from local representatives that have been nominated to accelerate the adoption of process mining in their region. Over the course of the implementation project, the local representatives received central support from the author of this thesis and additional experience-sharing sessions with the audit teams have been set up on a monthly basis.

The certified release used for the broader implementation in the audit practice consists of a new version of the application, a functional guide providing technical guidance on each dashboard and analysis provided and a methodical audit guide summarizing the audit approach using process mining.

[401] Cf. Chapter 3.2.1.
[402] Cf. HARTKE, LARS/HOHNHORST, GEORG/SATTLER, GERNOT (2010), p. 379.

3.3.2 Audit team feedback and results from quality assurance procedures

Feedback obtained from the implementation

Feedback interviews have been conducted with all audit teams in the period between November 2020 and April 2021. The feedback has been structured similar to the previous piloting projects[403] and relates to:

(1) the process of extracting, transforming and loading the data,
(2) the integration of the procedures performed with process mining into the audit approach,
(3) the analyses provided in the analyzer and
(4) any additional comments from the audit teams.

The feedback obtained within each category summarizes as follows:

At (1): Feedback on the ETL process

- Teams perceived the ordering process of the analyzer as very technical. Most of the teams needed support from a technical expert to understand and obtain the required information.
- The audit teams again remarked that the data extraction and transformation process for process mining is much more complex than it is for other automated tools and techniques in the audit. Many teams fed back that understanding this process and the prerequisites for applying process mining is time-consuming but essential to appropriately use the technology to support audit procedures.
- Several issues in the ETL process reported by the local delivery teams led back to country specific characteristics of SAP. For example, most of the tables in SAP store amount fields with two decimals. However, currencies from some countries (for example, Hungary, Japan and Korea) do not use decimals. As a result, 100 JPY are stored as 1.00 JPY in SAP. When extracting the data from the system, the extraction routines need to account for these specifics in order to report the correct values. Similar issues may arise when exporting Japanese text from the system. For example, as the number of characters in Japanese is more than 256, they cannot be encoded using a single byte. Thus, Japanese is encoded using two or more bytes, referred to as “double-byte” or “multi-byte”

403 A summary of the feedback obtained about tool performance is omitted as no significant performance issues have been reported by the audit teams.

encoding. This may lead to shifts in columns and manual steps in data preparation using extraction routines that have been developed based on a language encoded using a single byte.

At (2): Feedback on the integration into the audit approach

- All audit teams confirmed that with process mining, they obtained a detailed understanding of the significant class of transactions and reached an appropriate level of understanding of the entity's process. Also, they have been able to identify and assess risks using process mining.
- Four audit teams fed back that they usually did not differentiate between different types of transactions (SAP MM and SAP FI transactions) when auditing the purchase to pay process. However, process mining confirmed that the flow of transactions within the significant class of transactions is very different and related to different risks of material misstatement. The teams noted that this deconstruction of the significant class of transactions led to additional work but increased audit quality, as different risks of material misstatement are related to the different processing.
- Many audit teams noted that comparing data between different process mining applications was very inefficient. The teams suggested to integrate a comparison functionality directly within the application to compare the data as of period end with the data already analyzed in the previous audit period or at an interim date.
- One team emphasized the potential of process mining to provide an automated roll-forward functionality, highlighting changes in the process or new process variations.
- Many teams indicated they have been able to identify controls relevant to the audit and using process mining to test the design and operating effectiveness of controls felt more intuitive that using the solution to perform substantive procedures. As a result, many audit teams suggested to reconsider the limitation of applying process mining to support substantive procedures only.
- Most of the teams pointed out that the methodical guidance has been used extensively and helped to guide the procedures supported by process mining. One team reported that more extensive guidance and real-life examples are needed for the analyses provided in the process mining application. The team would

further appreciate "best practices" to demonstrate how to support audit procedures with analyses performed using process mining.

At (3): Feedback on the analyses provided in the analyzer

- Teams stated that cases flagged as being open in the process mining analyzer do not reconcile to the vendors' open items as of period end.
- Several teams fed back that the total balances on the "Vendor structure" dashboard do not reconcile to the liabilities recorded in the subledger.
- Further, teams indicated that the invoice amount of a vendor does not seem to be displayed correctly when viewed by month. They have not been able to reconcile the invoice amount of a vendor throughout the period to the financial data at a specific point in time.
- The technical guidance on the functionalities of the process mining analyzer was perceived as complex but necessary.

At (4): Additional feedback received from the audit teams

- While many teams confirmed they will continue to apply process mining in their audits, some teams highlighted the high costs to set up the application, understand the data model and appropriately perform the procedures as a potential blocker for next year's application.
- Teams asked for process mining solutions applicable to other ERP systems and business processes.

Feedback evaluation and results of quality assurance procedures

Despite the positive feedback obtained on the level of detail provided by process mining that increases the quality of the auditor's process understanding and related audit procedures, some of the feedback obtained throughout previous waves of piloting persists. Even with a comparably mature process mining solution that has been subject to extensive piloting and certification procedures, audit teams keep struggling with evaluating process variations, comparing data at different points in time and the extensive process and IT understanding required already in advance of the analysis.

Main challenges identified by either the feedback provided or the quality assurance procedures performed throughout the implementation of the process mining analyzer relate to:

(1) the audit teams' understanding of the entity's process and information system,

(2) challenges regarding the analysis of the vendor structure and related information, in particular with regard to:

   (a) cases marked as "open cases" in the process mining analyzer that do not reconcile to the individual open items by vendor at period end,

   (b) a total invoice volume by vendor that does not reconcile to the vendor's account payables balances at period end,

   (c) identifying transaction flows related to a vendor that are not covered by process mining (and thus need to be audited alternatively using other automated tools or techniques that do not include process data but are based on subledger data),

   (d) performing analytical procedures on the development of the invoice amount by vendor throughout the period and

   (e) multiple vendors involved in a specific case.

At (1): Understanding of the entity's process and information systems required for using process mining

The broader implementation in the audit practice across multiple regions confirmed that audit teams are frequently not entirely familiar with the different SAP modules and configurations used by the entity and additional third-party applications of the information system relevant to the analyzed business process. However, this understanding is essential for appropriately setting up the process mining analyzer, including required custom activities that are not part of the standard data model. For some audit teams, the coverage of the accounts related to purchase to pay by the process data was comparatively low, as some components of the auditees' information systems were not considered in the process mining data model by configuring appropriate custom activities. For the (parts of the) financial accounts not covered by the transactions included in the process data, additional audit procedures need to be performed outside of process mining to meet relevant audit objectives. In addition, if aspects of the process are performed outside the entity's information system, the auditor needs to perform traditional audit procedures to identify and assess any risks of material misstatement related to these manual activities. Consequently, provided the extensive prerequisites of process mining, including the detailed process understanding and technical knowledge required to set-up and interpret the analyses appropriately, some audit teams perceived a mismatch between cost and benefits of using process mining.

At (2a): Case status differentiating between "open" and "closed" cases

UiPath's process mining solution assigns each case a "case status", differentiating between cases that are still open and cases that are closed.[404] For the purchase to pay process, this classification of the case status is made based on whether the case already has a payment activity or not. In this scenario, the payment activity is considered as the ending (and thus, closing) activity of the case. The observations and quality assurance procedures conducted as part of the implementation confirm that such a predefined case status is problematic in auditing:

- Audit teams sometimes confuse the case status "open" with an open liability. If a purchase order has just been created, the cases related to the purchase order line items contain the event "purchase order created" only. As no goods and no invoice have been received and no payment has been made yet, according to the definition of the case status, the case is "open" but no liability is existing yet.
- Irrespective of the choice of the process instance, a case may have multiple deliveries, invoices and payments. Consequently, the case status cannot be determined based on whether a payment event is existing. Otherwise, a case might be marked as being "closed" although the related invoice amount has not been fully paid. Determining whether a case is fully paid and consequently can be considered as "closed" at the end of the audit period needs to be based on the amount that is expected to be paid according to the invoice. This requires information not only about the invoice amount of the case as a whole but the invoice amount of the case at a specific point in time.
- Cases categorized as "closed" may still have business activity. This may happen, for example, if there is a frame agreement with a supplier for which a single purchase order reference is used.

Consequently, the classification of the case status should be removed from the process mining analyzer or replaced with an appropriate indicator for a case's inclusion in the liabilities at period end.

---

[404] Cf. UiPath (2021).

At (2b): Reconciliation of vendor balances between the process and the financial data

In the audit firm's process mining solution, the quantitative information of a case is limited to the metrics stored as case attributes.[405] This includes, for example, the total invoice amount of the cases processed in a specific process variation[406] or related to a specific vendor. However, a total invoice amount related to the activity of entering the invoice into the system will not reconcile to the vendor balances in the financial statements as of period end, as it does not inform about actual postings, reversals, payments, or other transactions affecting trade payables.

When performing substantive procedures, the auditor is not particularly interested in the total invoice amount entered for a specific vendor but in the occurrence of purchases throughout the year and the completeness of the vendor-related liabilities as of period end.

At (2c): Performing procedures for "non-covered" parts of the vendor balances

If there are significant transaction volumes or account balances relating to trade payables that are not covered by the process data, further audit procedures to identify and address any related risks of material misstatement need to be designed. However, without considering the subledger data and reconciling it to the vendor movements that are covered by the process data, the auditor cannot identify the individual transactions not covered by the process analyses (for example, credit memos posted directly to the subledger without referencing the related purchase order line item).

At (2d): Analytical procedures on the monthly development of the vendor-related activity

The audit firm's initial prototype built based on UiPath's process mining application did not include detailed information on the vendors in the process beyond the number of cases related to a specific vendor.[407] Consequently, quantitative information on how the individual cases affect the related vendors has been incorporated in the analyzer, including the purchase order and the invoice amount. However, if the information on these metrics is only available as case attribute and not related to the individual events

[405] Cf. Table 2.
[406] Cf. Figure 35.
[407] Cf. Chapter 3.1.1.

in the event log, the vendor related information may only be displayed for the entire period.[408]

To provide at least a broad indication of the distribution of individual cases by vendor throughout the period, the analysis presented in Figure 37[409] enables the auditor to determine for each vendor how many cases have been initiated or ended (because no additional event happened in the period investigated) in a specific month. Some teams adjusted the analysis to investigate the distribution of the invoice amount related to specific vendors throughout the period.

Vendor	2018 - 1	2018 - 2	2018 - 3		2018 - 10	2018 - 11	2018 - 12	▼ INV amount (CY)
UQYE		268,269.65	46,600.29		2,222,340.70	73,032.95	1,223,964.39	6,152,899.23
BYGHI	778,151.46	248,706.97	400,598.99		566,961.86	694,964.14	501,719.94	5,942,739.57
BFGJO	3,614.00	14,332.00	53,897.00		1,346,655.00	14,963.00	41,541.00	4,358,440.22
AGKJA	2,111,718.05	32,721.55	4,911.13		77,113.41	256,224.22	215,621.57	3,172,928.69
LEN	218,242.91	15,075.33	10,127.12		190,343.00	25,778.35	9,292.04	2,984,957.46
AOPXU	894,812.51	189,291.91	112,716.96		156,841.70	85,162.46	27,775.87	2,592,635.29
BVKBS	54,719.93	454,532.65	210,349.72		289,857.39	137,116.41	414,301.46	2,543,455.55
AXKCS	0.21	82,118.15	204,340.37		86,670.18	136,821.61	167,247.09	2,350,757.47
ARWCJ	101,636.92		47,285.64	...	266,018.69	453,359.48	323,865.93	2,322,647.66
ALSIF	9,873.88	9,692.01	17,119.18		62.85	9,090.32	2,982.11	2,137,632.43
YICM	321,920.68 -...	236,612.09	310,995.09		141,938.51	70,570.11	256,696.40	2,057,716.76
AXAII	236,811.77	137,359.18			188,877.32	498,072.71	371.84	2,010,493.94
BOBYL	100,461.52	163,589.40	69,514.70		198,991.62	208,799.10	83,688.84	1,930,658.51
AMOKD	182,458.18		587,828.24			94,301.97		1,840,742.26
AUOGX	631,225.29	38,216.58	89,660.37		493,273.89	57,016.49	109,029.72	1,698,796.74
RXTF	94,309.55	223,488.65	93,096.88		77,527.85	168,696.56	68,164.29	1,571,092.47
BPRGM	22,106.98	16,420.33	26,949.56		348,084.96	196,237.86	25,329.79	1,536,163.27
BJAZJ	1,994.80	461.62	5,892.61		372,584.45	198,291.24	3,100.96	1,510,493.00

Figure 44: Invoice amount of the cases initiated in a specific month by vendor

However, as the processing of individual cases overlaps, the invoice amount related to a vendor at a specific point in time throughout the period cannot be determined. The implementation presented in Figure 44 requires the auditor to decide if the invoice amount related to a case is displayed in the month of the first or the last activity of the case. By this, even if the invoice amount posted (and not entered) is considered, time-related analyses are misleading.

[408] Cf. Figure 36. The metric of the analysis may be changed to display the total invoice amount by vendor instead of the number of related cases.

[409] Cf. Chapter 3.2.1.

Consequently, the invoice amount cannot be incorporated appropriately in this dashboard and procedures involving the movement of transaction volumes between individual months are not supported by the process mining analyzer.

At (2e): Consideration of multiple vendors involved in a case

The case table illustrated in Table 2 includes information about the vendor related to a specific case. However, the vendor receiving the purchase order, the vendor providing the goods, the vendor sending the invoice and the vendor to whom the payment is made can differ. Vendors may vary, for example, due to the set of conditions defined for a specific purchase order item. These conditions include, for example, the gross price, the planned delivery cost and taxes. In SAP, multiple vendors may be related to a business transaction by assigning different partner functions to a document.[410] A "partner function" is a common term used to define the roles, rights and responsibilities of different business partners involved in a particular transaction. Standard vendor partner functions related to a purchase order include the ordering address, the goods supplier, the invoicing party and the payee. Figure 45 summarizes the concept of multiple vendors related to an individual process instance.

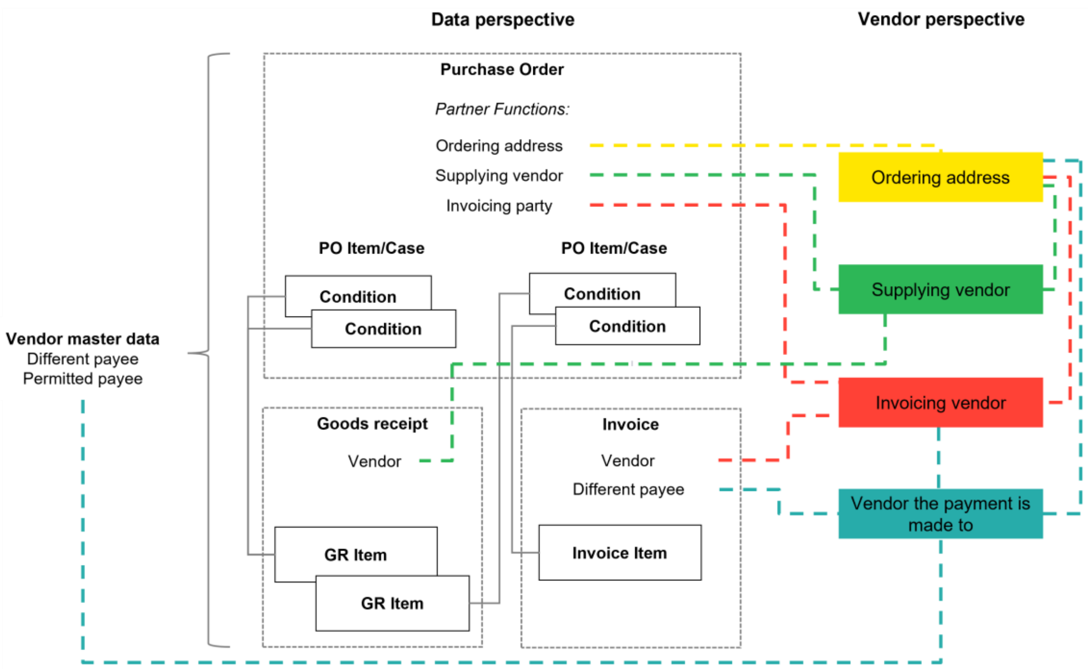


Figure 45: The concept of multiple vendors related to a process instance

In the example provided, three different partner functions are defined in the purchase order document. The vendor receiving the purchase order is defined by the ordering

[410] Cf. HARTKE, LARS/HOHNHORST, GEORG/SATTLER, GERNOT (2010), p. 346.

address. The vendor supplying the goods has a reference to both the purchase order and the goods receipt. Similarly, the invoicing party is specified by its partner function in both the purchase order and the invoice document. Finally, the vendor the payment is made to may differ from the invoicing party, for example, in the scenario of factoring. The information on permitted payees is maintained in the vendor master data.

The example demonstrates that the varying roles of a vendor cannot be considered if the vendor information is stored as case attribute. For example, if the vendor included in the case table is the vendor receiving the purchase order, all related case metrics (for example, the purchase order amount, the quantities of the related goods receipt and the invoice amount) will be related to this vendor. However, if the case has a different invoicing or supplying party or the payment is made to another vendor, analyses solely based on the case table will be misleading.

Consequently, when analyzing vendor related information, specific consideration needs to be devoted to the role of the vendor in the process, i.e., to displaying information appropriately for cases that that include events related to different vendors.

The feedback obtained with regard to the analysis of the vendor structure and related information demonstrates that the auditor is not used to analyze process information using automated tools or techniques, or more specifically, to analyze process data without considering the financial aspect. Substantive audit procedures performed on trade payables and related vendor accounts usually include analytical procedures on subledger data, for example, to investigate unusual balances or monthly credit and debit movements on the one hand and selecting samples and key items of balances and transactions for substantive tests of details on the other hand.[411] As this financial information is not included in today's event logs, challenges result when using process mining to support substantive procedures.

3.4 Interim conclusion

Scientific literature on process mining includes predominantly positive statements on the technology's applicability to an audit of financial statements.[412] Chapter 2.3 of this

[411] Cf. ELDER, RANDAL J. et al. (2020), pp. 579f.
[412] Cf. Chapter 2.2.

thesis discusses how process mining may be integrated into the risk based audit approach and concludes that from a theoretical point of view, the statements made in the process mining related research can be supported. Chapter 3 deals with the empirical validation of the suitability of the technology to support designing and executing audit procedures. Over the period from 2017 to 2021, one of the Big Four audit firms extensively tested the practical application of process mining in the audit. In total, 157 process mining applications have been deployed to engagement teams across 20 different countries. The empirical evaluation strongly confirms the potential of the solution to increase the quality of the auditor's process understanding and to support risk assessment procedures, tests of controls and substantive audit procedures.

However, despite the theoretical suitability of process mining to support a financial statement audit, several challenges related to the practical application are identified. Many findings relate to education and training needs inherent to the adoption of a new technology in the audit practice. These findings include acceptance challenges similar to those observable in the audit profession when first introducing automated tools and techniques to electronically analyze the entire population of general ledger data. For example, some auditors decided to not adopt process mining on their engagement and adjust or challenge an audit approach for processes and internal controls that is performed almost unchanged from a methodical perspective since many years. Other findings have been addressed throughout the piloting process by customizing the process mining application to the requirements of the audit and integrating additional functionalities. However, numerous methodical and regulatory questions and technical challenges related to the practical application of process mining are identified that need to be answered and resolved, respectively, to enable a broader application of the technology in the field of auditing. The findings and challenges related to the second research question of this thesis are identified based on the audit teams' feedback obtained over the course of the piloting process, the quality assurance procedures performed by the author of this thesis and discussions with various domain experts conducted as part of the application development and the global certification process.

The following Chapter 4 discusses the major findings, questions and requirements resulting from the practical application of process mining in auditing. At the same time, propositions on how to resolve or handle the challenges in future implementations of process mining are developed. From the challenges identified, the following aspects are of significant importance from an audit methodology and application perspective:

- In practice, process mining does not reconstruct the process from the ERP data automatically. Many auditors perceive the extensive process understanding and technical knowledge required in advance to using process mining as a blocker for adopting the technology on their audit engagement. Limitations regarding the completeness and accuracy of the input data and the performance of the process mining application on large datasets further reduce the relevance of process mining for the audit (Chapter 4.1).
- To fully benefit from a data-driven process analysis in the audit, a connection between the process data and the accounting data needs to be established. The auditor is not particularly interested in the total invoice amount processed for a specific vendor but in the occurrence of purchases throughout the year and the completeness of the vendor-related liabilities as of period end. To avoid over-auditing of classes of transactions with a limited risk of material misstatement and under-auditing of significant classes of transactions, an integration of the process data with the financial data is required (Chapter 4.2).
- There is a trade-off between the comprehensiveness of the event log and the complexity of the resulting process graph. In concurrent research, the datasets used to evaluate process variations are usually limited to a short period of time or a specific vendor only. Resulting process variations classified as unusual are handed over to the auditor for further investigation.[413] However, in practice, the transparency achieved with process mining comes along with the limitation that the numerous process variations cannot be audited in their entirety. The evaluation of piloting results confirms the audit teams' uncertainty about (a) the degree to which variations need to be analyzed and (b) the qualitative and quantitative criteria applicable to this analysis (Chapter 4.3).
- Current process mining applications do not support comparing the period end data of the audit period to the data that has already been audited in the prior period or at an interim date of the audit period. The empirical evaluation indicates that the efficiency of using process mining in the audit could be increased significantly if the analyzer supported a comparison of the current period's data with an interim period or the prior period (Chapter 4.4).

[413] Cf. CHIU, TIFFANY (2018), p. 22.

- The audit firm's implementation of process mining was limited to engagements where the auditor follows a substantive audit approach without reliance on internal controls over financial reporting. To some extent, this decision was based on an incomplete methodical integration and the circumstance that the suitability of process mining for testing the design and operating effectiveness of controls has not been empirically evaluated so far. However, team feedback confirms that process mining may support both understanding the process, including identifying controls relevant to the audit, and confirming the understanding. The findings give reason to further explore the use of process mining as an instrument to test the design and operating effectiveness of internal controls over financial reporting (Chapter 4.5).

4 Key challenges for the application of process mining in an audit of financials statements

4.1 Descriptive character of process "mining"

4.1.1 Problem definition

Audit firms, software vendors and researchers promote process mining as an instrument to achieve full transparency on internal processes and integrated controls.[414] There is general agreement in theory and practice that process data includes audit relevant information that is independent from the auditee[415] and may enhance all audit procedures that deal with internal processes. The term „mining" suggests that the solution reconstructs the process based on the data in an explorative and largely automated manner. However, in practice, extensive knowledge of the process, the relevant IT and ERP systems, their interfaces as well as existing configuration and customization options is required already in advance to the actual analysis. As with any other automated tool or technique, there is trade-off between costs (i.e., resources and time to setup and perform the analyses) and benefits that needs to be considered when assembling, transforming and presenting the data to the auditor.[416] This poses challenges to both the development of a process mining solution that is scalable to a large number of different entities and the application of process mining on a specific auditee's data. Decisions made in extracting, transforming and loading the data, in developing the analyses and in customizing the solution to a specific entity directly impact the analyses that can be conducted[417] and the audit evidence that can be obtained. The empirical evaluation reveals that many audit teams perceive these prerequisites as a blocker for adopting process mining on their audit engagements. Provided the detailed process understanding and technical knowledge required to set-up and interpret the analyses appropriately, they do not see the additional value process mining may provide to their audit.

---

[414] Cf. Deloitte (2020); EY (2019); PwC (2019); KPMG (2018); Chapter 2.2.
[415] Cf. Chapter 2.1.
[416] Cf. Chapter 3.1.2; Jans, Mieke/Alles, Michael G./Vasarhelyi, Miklos A. (2013), p. 3.
[417] Also see Jans, Mieke (2019), p. 59.

Decisions and challenges in preparing the data, designing the analyses and customizing the application include:

- understanding the ERP system, including the concept and limitations of relational databases, the interfaces to other IT applications and entity specific configurations and customizations,
- determining the activities to consider in the event log,
- determining the process instance and its granularity,
- considering the timing and duration of events,
- determining the data extraction period and its implications on the analyses and
- assessing the completeness and accuracy of the data.

4.1.2 Understanding the IT landscape and scoping the event log

Identifying the relevant data sources and scoping the event log

Analyzing a business process with process mining requires input data that is extracted from an entity's ERP system.[418] Today's ERP systems like SAP, Microsoft Navision and Oracle use relational databases that store data in tables with rows and columns. ERP systems have integrated workflows supporting organizations in executing business processes. However, unlike workflow-engine-based systems frequently used for case studies in process mining research,[419] ERP systems are not process-oriented. As such, they do not provide integrated functionalities to support data extraction for process mining in the form of a process-specific log file that may be exported and loaded into the process mining application.[420] The data that is relevant to a specific process is stored in various tables at numerous locations and the data structure varies across different ERP providers. The tables are linked based on data they have in common, enabling queries of specific information using the database language „Structured Query Language" (SQL). SQL is the standard language[421] for interacting with management systems and supports, for example, creating a new table of data by extracting information from one or more tables of the relational database.

In contrast to workflow management systems, ERP systems require to assemble the relevant input data for process mining from multiple tables and data fields across the

[418] Cf. Chapter 2.1.
[419] Cf. Chapter 2.2.
[420] Cf. IEEE TASK FORCE ON PROCESS MINING (2012).
[421] Cf. KRANAKIS, EVANGELOS (2013), p. 385.

information system. The actual "mining" part of process mining starts when the relevant input data is loaded into the process mining analyzer. It includes the discovery of relationships in the data and the reconstruction and visualization of the process graph.[422] However, before the data can be loaded into the analyzer, the information required to create the event log (or more precisely, the tables with the input data that are loaded into the process mining analyzer) need to be located, extracted and transformed into a data model suitable for the process mining analyzer. Consequently, creating an event log involves various decisions that directly impact the analyses that are supported by process mining.[423] This includes decisions with regard to the process instance, the relevant activities and additional information on these activities (for example, attributes or meta-data), as these determine the tables and data fields that need to be considered when extracting the data from the ERP system.[424] As a prerequisite, detailed knowledge about the ERP system, other information systems relevant to the process as well as the interfaces between these IT applications is required.

Weaknesses of an ETL process based on relational databases

Figure 46 outlines the ETL process[425] for the analyzer introduced in Chapter 3.

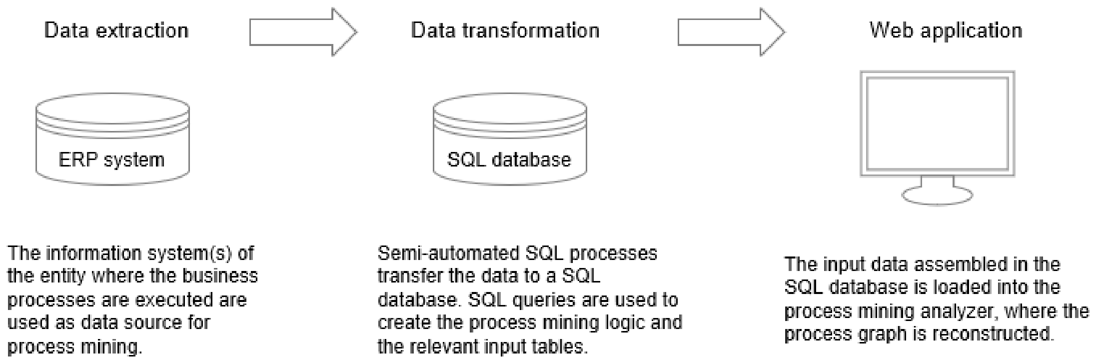


Figure 46: Current process mining approach based on a relational database

The data that is extracted from the entity's ERP system is transferred to a SQL database, where it is restructured and transformed into a format suitable for process mining.

The SQL database is structured in data cubes, i.e., individual data sources that contain the relevant input tables. The decisions made in assembling the event log directly im-

422 Cf. GEHRKE, NICK/WERNER, MICHAEL (2013), pp. 4f.
423 Cf. VAN DER AALST, WIL M. P. (2016), pp. 144ff.
424 Cf. JANS, MIEKE (2019), p. 60.
425 Cf. Chapter 3.1.1.

pact the number and types of data sources included and the perspectives of the process[426] that may be analyzed. For the case audit firm's process mining purchase to pay analyzer, the most important cubes in the SQL database are cases, events, user information, vendor information, vendor changes, journal entries, invoices, goods receipts, bank data and tolerances. The case cube contains the data on the purchase order line items, like case ID, amount or case owner. The event cube contains the individual events which occurred in the case and represents the event log. All user-related data, including user IDs and departments, are stored in the user cube. General information on the vendor, like name, address, or bank data, is saved in the vendor cube. Information which relates to changes in this information is stored in the vendor changes cube. The journal entry cube contains all journal entry-related data like the posting date or amount. Data relating to invoice documents like invoice amount, invoice date or invoice quantity is stored in the invoice cube and, similarly, information on goods receipts, like quantity or date, is stored in the goods receipt cube. The bank data cube contains all relevant bank data like account numbers, account owner, or bank names. The tolerances used in two- and three-way-matches are stored in the tolerance cube. Figure 47 provides a simplified overview of the data cubes and their relationships.

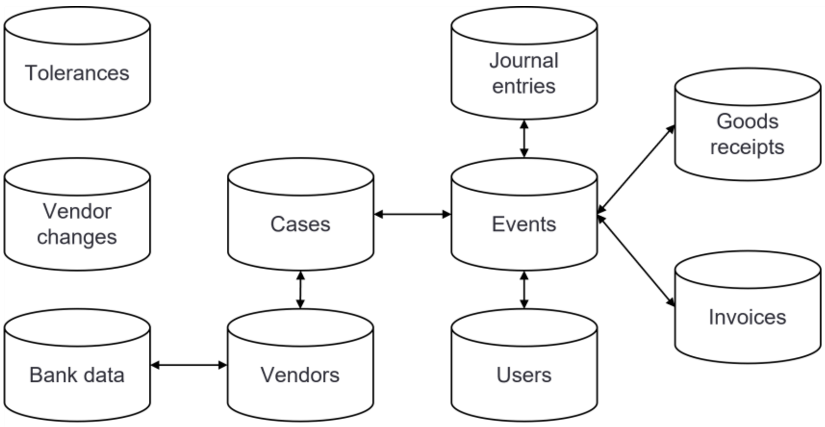


Figure 47: Process mining data sources (simplified presentation)

The links between the individual data sources need to be established manually before the data can be loaded into the process mining application. They are used to access data from one cube and combine it with information of cubes that are directly or indirectly linked.

[426] Cf. Chapter 2.1.

However, several challenges audit teams were facing relate back to the underlying data model and the relational database structure.[427]

For example, some audit teams reported they were not able to filter for the significant accounts in the purchase to pay process on the "Account coverage" dashboard[428] to then investigate the processing of the related cases using the process graph on the "Overview" dashboard[429]. In analyzing process variations, another audit team tried to identify the general ledger accounts used for processing the most frequent process variations. The team raised the issue that when filtering variations on the "Critical path" dashboard[430] and then switching to the trial balance on the "General ledger activity" dashboard[431], the specified filters get removed. At the same time, the dashboards on the "Financial reconciliation" menu page[432] do not support filtering for specific process variations. These issues result as the analyses on the "Financial reconciliation" menu page are based on the journal entry data cube, i.e., they use the general ledger data, whereas the analyses on the "Understand process" menu page[433] are based on process information, i.e., they use the case data cube as input. Switching between dashboards that use different data cubes as data source will result in all filters being removed unless the links between the data sources have been manually established.[434]

Further, fitting the complex relationships of a business process into a relational database can be inefficient due to redundant information storage. The data needs to be stored and modeled for each analysis perspective that is provided, immediately impacting the performance of the analyzer especially for large datasets.

Consequently, during the implementation of process mining in the audit practice, the integration of the data model was a continuous task to provide the auditor with the required flexibility in designing audit procedures while ensuring performance of the application. At the same time, a significant portion of the training and coaching effort

---

427 Cf. Chapter 3.2.2.
428 Cf. Figure 28.
429 Cf. Appendix II.
430 Cf. Figure 34.
431 Cf. Figure 27.
432 Cf. Chapter 3.2.1.
433 Cf. Appendix III.
434 In the implementation described in Chapter 3.2, the link between both data sources is established using the "Booking pattern" analysis, enabling the auditor to identify the general ledger accounts used for specific recording events within the process, cf. Figure 29.

needed to be devoted to the data model underlying the process mining analyzer and its impact on the analyses that may or may not be performed.

Implications on the application of process mining in the audit practice

When companies started to introduce the first IT solutions to support their daily business, the number of applications and data flows were relatively limited. Over the last decades, IT became a corporate asset and technological innovation and changing business demands led to a rapid growth of IT solutions. As a result, the IT environment of today's organizations includes a vast number of IT applications and interfaces managing the data flows for the daily business. Understanding the role of IT and its use by the entity is an important element in understanding the process and the entity's system of internal control according to ISA 315.[435] As inputting information, processing transactions and recording the transactions into the general ledger involves the use of IT, the auditor usually identifies the accounting relevant IT applications concurrently with obtaining an understanding of the process.[436] Challenges arise if on the one hand, process mining is promoted to facilitate this process understanding by automatically reconstructing the actual process execution from the relevant data sources, while on the other hand, the information on the IT applications and activities relevant to the process is a prerequisite of setting up the process mining analyzer in the first place.

To reduce the complexity of the ETL process and provide both meaningful and interpretable analyses, current process mining solutions are limited to the analysis of a single process at a time, with the prerequisite that this process has as few dependencies from other processes as possible.[437] Still, in expert surveys conducted by CLAES and POELS and – eight years later – by MARTIN et al., practical challenges identified for process mining include in particular the data availability and acquisition, cost of data transformation and insufficient data quality.[438] In addition, extensive manual data preparation and data cleaning is required as part of building the event log.[439] Even if the analysis is limited to a single business process only, the configuration of the process may vary depending on the ERP system, preprocessing IT systems, interfaces between the IT applications and entity specific customizations. If the data relevant for

[435] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 25.
[436] Cf. ibid., para. A143.
[437] Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 359.
[438] Cf. MARTIN et al. (2021), p. 521; CLAES JAN/POELS, GEERT (2012), p. 190.
[439] Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 359.

process mining is scattered over multiple tables or even different information systems, the events and their corresponding cases are difficult to identify.[440] This is especially challenging when dealing with processes that are historically grown and include many interrelated IT systems.[441] Besides understanding the IT landscape itself, the auditor needs to have knowledge about how the entity stores and handles its data.[442] For example, IT systems usually do not store as much data as they theoretically could, because such recording slows down the system.[443]

Due to the heterogeneity of IT landscapes, creating a repository of ERP systems and other relevant IT applications with all configuration and customization options available is impossible in practice. Locating and scoping the relevant data sources to assemble an event log for process mining requires a rigorous and defensible method of structuring the data[444] and by this, domain knowledge and the involvement of IT specialists.[445]

At the current state of process mining and with the extensive involvement of domain specialists required, it is not feasible in practice to develop a custom process mining application for each individual auditee.[446] Consequently, decisions had to be made by the case audit firm in order to continue exploring the use of process mining beyond the first feasibility assessment. After the initial introduction of the technology to the audit practice, the process mining solution, the underlying data model and the related ETL process have been standardized to enable later scalability of the solution and facilitate both quality assurance and appropriate technical and methodical support structures. This was achieved by limiting the first process mining application to the analysis of the purchase to pay process only and developing and maintaining a common process mining data model for the ERP system SAP. The common data model was then customized to the specifics of the individual audit engagement.

The limitation to a single process (purchase to pay) and ERP system (SAP) significantly reduces the potential application scope of process mining in the audit. Still, even

---

440 Cf. IEEE Task Force on Process Mining (2012), p. 179.
441 Cf. ibid.
442 Cf. Jans, Mieke/Alles, Michael G./Vasarhelyi, Miklos A. (2014), p. 1756.
443 Cf. ibid.
444 Cf. Jans, Mieke/Alles, Michael G./Vasarhelyi, Miklos A. (2010), p. 10.
445 Cf. Aguirre, Santiago/Parra, Carlos/Sepulveda, Marcos (2017), p. 111; van der Aalst, Wil M. P. (2016), p. 144.
446 Cf. Chapter 3.1.2.

with a standardized data model, several decisions are required in assembling and customizing the input data to the specific entity.

4.1.3 Determining the activities to consider in the event log

Due to the information that is automatically recorded, researchers emphasize the independent and objective verification of an entity's process execution when using process mining.[447] Process discovery techniques[448] are promoted for providing an unbiased view on the process solely using a given event log, i.e., without using additional a priori information and without the need to manually design a process model.[449] However, the log files extracted from the systems are typically not limited to a single process only. Data transformation and filtering is required in order to scope the input data to the process of interest. The scoping is supported by determining the activities that are known to belong to the same process.[450] Hence, an important step in creating the event log is identifying and understanding the activities that are performed in the process being audited.[451] For example, when analyzing the purchase to pay process, common activities include "purchase order created", "purchase order approved", "goods receipt posted", "invoice posted" and "invoice paid". Additional activities may include, for example, quantity or price changes made to the purchase order, the rejection of a purchase order or the blocking of an invoice for payment. As the number of activities directly impacts the complexity of the event log and the resulting process graph,[452] the relevance of each activity for the process understanding and the audit procedures planned needs to be carefully evaluated. For example, while the activity "invoice workflow started" may be of interest for process improvement initiatives of the auditee, it might not represent an activity relevant to the audit considering the results of the auditor's risk assessment procedures.

---

[447] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2014), p. 1755; ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), pp. 355f.
[448] Cf. Chapter 2.1.
[449] Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), pp. 355f.
[450] Cf. GEHRKE, NICK/WERNER, MICHAEL (2013), p. 4. A prerequisite when selecting the relevant activities is that the execution of the activity is captured by the system, cf. Chapter 4.1.7.
[451] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), p. 11.
[452] Cf. Chapter 3.2.2.

Consequently, in practice, process mining does not replace the need to manually understand the process. The understanding of the designed process and the involved IT systems is required to identify the activities that need to be included in the event log.[453]

Several findings of the empirical evaluation relate back to decisions made regarding the activities to include (or not to include) in the event log. Challenges include:

(1) identifying activities and assessing their relevance to the audit and

(2) determining the granularity level of activities.

At (1): Identifying activities in the process and assessing their relevance to the audit

The information on the relevant activities and their source systems is a prerequisite of setting up the process mining analyzer. Each activity needs to be defined from a technical perspective to extract the relevant raw data from the entity's information systems. As the activities in the process mining analyzer directly impact the possible analyses, in theory, the auditor should identify the activities in the process and determine if they are relevant to the audit and thus need to be included in the event log.[454] However, even with the limitation of the ERP system to SAP, there is a magnitude of different modules, configurations and customization options available. The practical implementation demonstrates that determining the activities to consider in the event log requires both extensive technical knowledge and a detailed understanding of the entity's process execution. As such, in determining the activities to include in the event log, the auditor needs to be supported by an IT specialist.

However, starting from a blank piece of paper for each individual auditee is not feasible in practice and leads to varying definitions, descriptions and labeling for the same activities performed in the ERP system.[455] To overcome this limitation, a default set of activities related to the purchase to pay process in SAP has been determined centrally as part of the implementation described in Chapter 3. In a second step, the set of activities for the standard application may be adjusted and extended by "custom" activities based on the ERP configuration and customization of the specific entity. The final decision on the activities to include for an individual audit engagement is then

[453] Consequently, from a practical perspective, the target process needs to be added as input data for process mining discovery techniques in Figure 6.

[454] Cf. Jans, Mieke (2019), p. 62.

[455] Cf. Chapter 3.1.2.

made by the auditor based on the understanding of the process, previous experiences with the entity and the auditor's professional judgment.

Although the entities that qualify for the implementation of process mining have been selected based on their use of the SAP MM and FI modules, several audit teams identified activities that are part of the auditee's purchase to pay process but are not included in the standard repository of activities.

Frequently, the feedback was related to entities using shopping carts in the procurement process that are created in the SRM module in SAP. In this scenario, the purchase to pay process is not initiated by the activity "purchase order created" in the SAP MM module but by the creation of the respective shopping cart in the SAP SRM module. The process mining analyzer introduced in Chapter 3 only covers purchases processed using the SAP MM or SAP FI module. If activities executed in the SAP SRM module are not considered in the event log, the purchase to pay process model reconstructed with process mining may be incomplete or misleading, depending on the specific configuration of the auditee's SAP system. Figure 48 illustrates the four different implementation strategies available for the SAP SRM module.[456]

[456] As the implementation described in Chapter 3 is limited to SAP, in the following, the back-end ERP system is assumed to be SAP. Technically, the SAP SRM module may also be integrated with a material management function outside of SAP.

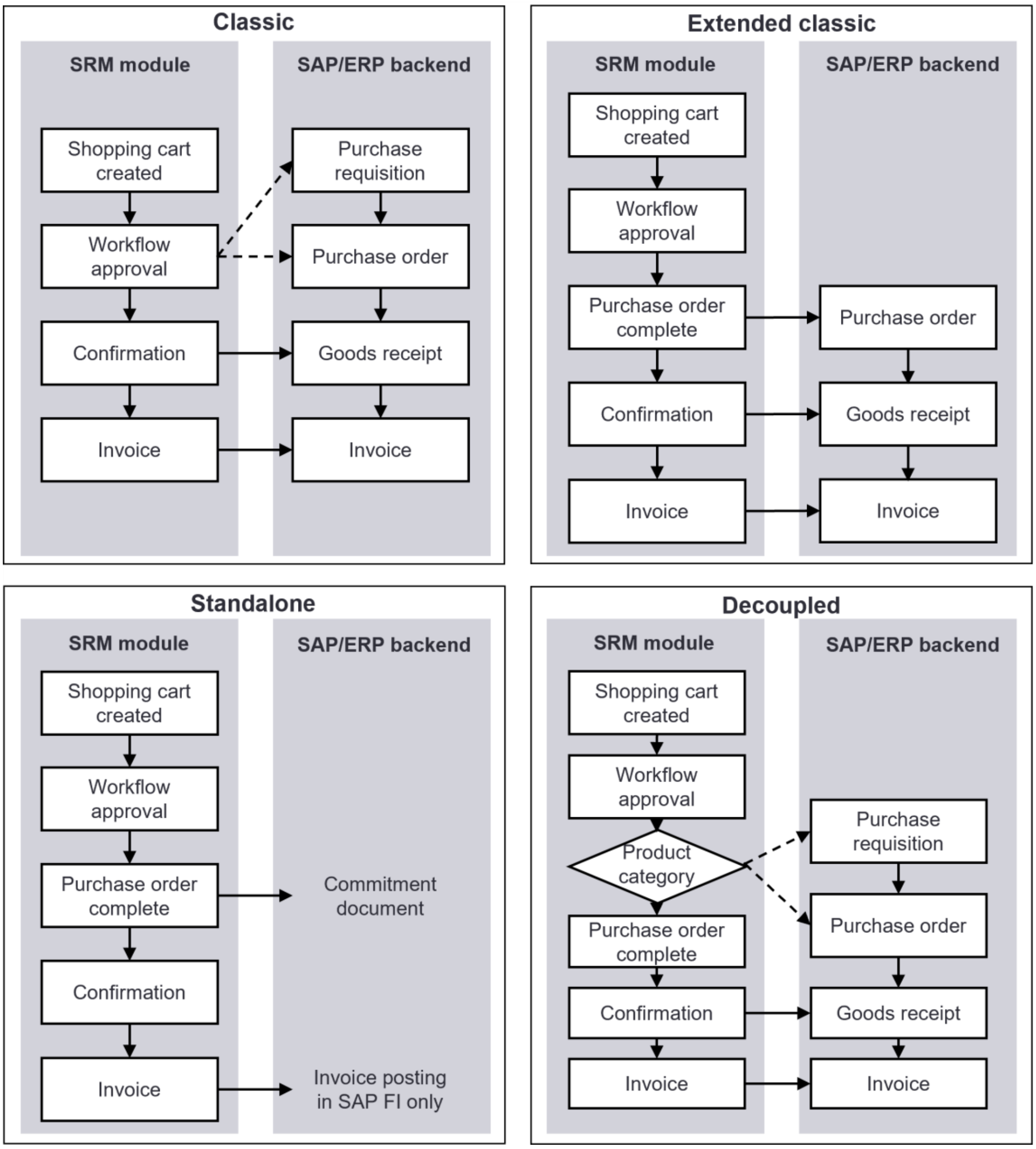


Figure 48: Implementation scenarios for the SAP SRM module[457]

In the classic scenario, the shopping cart is created and approved in the SAP SRM module, but the remainder of the purchasing process is located in the SAP back-end system. In this scenario, the SAP SRM module is linked to the SAP MM module and all related documents are located in the SAP MM module. This includes the purchase requisition that is automatically created with the workflow approval of the shopping cart, the purchase order, the goods receipt and the invoice.

In the extended classic scenario, the purchase requisition is not transferred to the SAP back-end system. All documents following the purchase requisition, for example, pur-

[457] Cf. SAP (2022).

chase orders, goods receipts and invoices, are created in the SRM module and replicated to the SAP MM module. It is important to note that these documents can only be changed in the SAP SRM module and not in the back-end system.

In the standalone scenario, the SAP SRM module is not integrated with the SAP MM module but used for the entire purchasing process. The shopping cart and the related documents are processed directly in the SAP SRM module. The related invoice is replicated to the accounting system.[458]

In the decoupled scenario, both modules are used in parallel. The product category determines if the purchasing process is handled using the SAP SRM module or the SAP MM module.

The example demonstrates both the importance and the complexity of assembling the appropriate data sources for a repository of activities in audit practice. Even if the analysis is limited to a single process and a single ERP system only, there is a multitude of configuration options available with different degrees of integration between the individual modules. If the event log is limited to data extracted from the SAP MM and FI modules, events related to the creation and approval of the shopping cart in the SAP SRM module are not extracted from the system by default. If there is no activity related to the purchase order requisition,[459] the process graph starts with the activity "purchase order created" but the actual initiation of the process in the SAP SRM module and the resulting purchase order requisition are missing in the data model. In the classic and extended scenario, the auditor needs to determine beforehand if the information on the creation and rejection or approval of the internal purchase requisition document is relevant to the audit, or if the accounting relevant significant class of transactions is initiated by the creation of the corresponding purchase order that is sent to the vendor. Further, in the extended classic scenario, changes to documents following the purchase order requisition can only be made in the SAP SRM module. As a result, any corresponding events related to changes of documents will not be included in an event log that is based on data extracted from the SAP MM module only. Similarly, if the entity uses the decoupled scenario of the SAP SRM implementation, event data for

[458] As a prerequisite for piloting process mining was that the entity makes use of the SAP MM and FI modules, the standalone scenario has not been identified for any of the auditees.

[459] This applies, for example, to the extended classic scenario where the purchase requisition is not transferred to the SAP MM module.

product categories handled in the SAP SRM module is missing if no SAP SRM-related activities are included in the data model.

Even if the purchase to pay process is predominantly executed using the SAP MM and FI modules, many auditees use some preprocessing systems (for example, invoice scanning and release applications like Readsoft) or integrated custom workflows (for example, invoice approval workflows for SAP FI transactions). If activities executed in these IT applications are considered relevant to the audit and thus covered by traditional process understanding and walkthrough procedures, they need to be incorporated in the event log to achieve an accurate representation of the actual process execution. However, as there is a trade-off between the information content of the event log and the complexity of the resulting process graph, the auditor needs to critically evaluate the necessity for each individual activity for the process understanding and the planned audit procedures.

To illustrate the practical challenges, the example of a large consumer goods company is used. In determining the activities to include in the analyzer, the audit team selected 23 activities from the standard repository and created 32 additional custom activities, resulting in a total of 55 activities. Table 7 lists the custom activities used by the audit team.

Custom activity	**Number of events**	**Number of cases**	**Percentage of total cases**
Workflow ended MM	63.616	51.552	36%
Workflow started MM	63.195	51.251	36%
Approval level 1 MM	27.986	19.962	14%
Shopping cart approval workflow ended (document reviewed)	16.763	16.763	12%
Shopping cart approval workflow started	16.709	16.708	12%
Shopping cart created	16.622	16.622	12%
Shopping cart sub-workflow automatic approval	11.677	11.677	8%
Workflow price checked	5.944	4.704	3%
Shopping cart approval workflow ended (approved)	5.233	4.757	3%
Workflow ended FI	4.776	4.776	3%
Workflow started FI	4.712	4.657	3%
Approval level 1 FI	4.406	4.404	3%
Approval level 2 MM	3.281	2.171	2%
Shopping cart sub-workflow approved	1.872	1.759	1%
Refused MM	966	612	0%
Workflow quantity checked	929	741	1%
Approval level 2 FI	669	669	0%
Approval refused 1 MM	496	338	0%
Shopping cart sub-workflow cancelled	335	272	0%
Approval refused 1 FI	254	236	0%
Invoice approval workflow master data change request	220	209	0%
Invoice approval workflow master data finish	219	208	0%
Invoice approval workflow master data change	193	177	0%
Invoice approval workflow master data confirm	161	157	0%
Shopping cart approval workflow ended (cancelled)	137	127	0%
Approval refused AP FI	56	46	0%
Refused FI	52	52	0%
Approval level AP MM	27	27	0%
Workflow price SSC accepted	22	17	0%
Approval refused 2 FI	4	4	0%
Approval refused 2 MM	4	4	0%
Workflow quantity SSC accepted	2	2	0%

Table 7: Custom activities included for a large consumer goods company

The activities relate to the SAP SRM module and the entity’s customized approval workflow for SAP MM and SAP FI invoices, respectively. Without investigating the technical definition of the activities, the interpretation of many activities is not straight-forward to a third person. If the custom activities have been created and labeled by an IT expert, interpretation difficulties may even arise for those members of the audit team that have not been involved in the setup of the process mining analyzer. However, some findings may already be derived from the aggregated overview of the custom activities presented in Table 7.

All the activities are purely workflow related, that is, no custom activity involves the posting of a document and the recording of a journal entry.[460] Provided a total number of 141.524 cases in the audit period, only six of the 32 custom activities have been

[460] These activities are already included in the audit firm’s default set of standard activities described earlier in this chapter.

performed for more than ten percent of all cases. The most frequent activities include the start and end of the SAP MM workflow and the SAP SRM approval workflow. Similar activities are available for the start and end of the SAP FI workflow. Provided that the initiating, recording, processing and reporting activities for these workflows are included in the event log, it is questionable whether the separate activities indicating the start and end of the workflows add additional value to the audit of the critical path of the process.

18 activities have been performed for less than one percent of all cases, giving rise to critically assess their relevance to the process. If an activity is only performed for a very small number of cases, it might be less important for the overall process while increasing the complexity of the process graph. For example, these activities include four activities related to master data changes (change request, change, confirmation of the change and completion of the workflow). Although the purchase to pay process relies on appropriate master data maintenance for the related vendors, the two processes are separated in the ERP system. While the event data is following the processing of a specific case (i.e., a document within the purchase to pay process) through the system, the master data is maintained on the level of the vendor. Incorporating data sources from both processes in a joint process graph may lead to interpretation difficulties, as a single change of an individual vendor (i.e., one vendor related change event at a specific point in time) will be duplicated into an individual event for all cases affecting the respective vendor. The challenges in interpretation increase if a case is related to multiple vendors,[461] for example, an invoicing vendor that differs from the vendor the purchase order has been sent to. In addition, the activities only incorporate master data related activities that are performed as part of the invoice approval workflow, i.e., master data changes outside this workflow are not considered in the process graph. As the concurrency of master data changes with transaction processing may give rise to risks of material misstatement, the analyzer provides a dedicated dashboard[462] to analyze the entire population of vendor master data changes throughout the data extraction period. Consequently, the auditor may decide to refrain from incorporating parts of the master data process as activities into the process graph in order to reduce the process complexity. Further, in the example provided, the activity “workflow quantity SSC (Shared Service Center) accepted” is only performed for two cases

[461] Cf. Chapter 4.2.4.
[462] Cf. Chapter 3.2.1.

throughout the period and, in addition, might be redundant to the three-way-match[463] of the quantity information between purchase order, goods receipt and invoice.

There is a total of 16 different custom activities explicitly limited to either SAP MM or SAP FI transactions. Distinguishing between SAP MM and SAP FI transactions is important, as they may relate to different risks of material misstatement. However, whether a specific case is related to the SAP MM or SAP FI workflow is already incorporated in the case type definition.[464] Hence, the auditor may identify if an approval is related to a SAP MM invoice or a SAP FI invoice simply by filtering the data to the respective case type. Considering the information already stored as case attribute, the number of the respective custom activities may be reduced from 16 to eight.

The 55 activities included in the event log of the consumer goods company result in 1.458.961 events that are executed for 141.524 cases. The corresponding process graph constitutes 10.205 distinct process variations. Setting up the analyzer for the same audit period but considering only the 23 activities from the standard repository results in 5.608 process variations, reducing the number of variations by 54,9 percent.

The example demonstrates the practical challenges that arise when process mining is planned to be scaled to many different entities. Many audit teams tend to include as much information as possible in the process graph. If the entity makes use of an invoice approval workflow and the SAP SRM module, incorporating the activities relevant to the accounting process into the process graph supports addressing identified risks of material misstatement. However, as each activity needs to be validated and tested on the one hand and increases the complexity of the process model on the other hand, cost and benefit of each activity need to be carefully evaluated. As such, the audit teams need to be supported in understanding the set of standard activities, identifying any additional activities available in the process and making informed decisions about their relevance to the audit.

At (2): Determining the granularity level of activities

After identifying the activities to include in the event log, the auditor needs to determine their level of disaggregation. Piloting shows that process mining applications may include activities at a granularity not relevant or not required for end-users. For example, instead of using the standard activity "purchase order approved", some audit

[463] Cf. Figure 30.
[464] Cf. Chapter 3.3.1.

teams decided to refine the activity to directly incorporate the hierarchy level of the individual approving the purchase order.[465] As a result, there are multiple activities referring to the same procedure of approving a purchase order ("purchase order approved - hierarchy 1", "purchase order approved - hierarchy 2", …, "purchase order approved - hierarchy N").

As an excerpt from Table 7, Table 8 includes ten activities related to the approval or rejection of an invoice.

Custom activity	Number of events	Number of cases	Percentage of total cases
Approval level 1 MM	27.986	19.962	14%
Approval level 1 FI	4.406	4.404	3%
Approval level 2 MM	3.281	2.171	2%
Approval level 2 FI	669	669	0%
Approval refused 1 MM	496	338	0%
Approval refused 1 FI	254	236	0%
Approval refused AP FI	56	46	0%
Approval level AP MM	27	27	0%
Approval refused 2 FI	4	4	0%
Approval refused 2 MM	4	4	0%

Table 8: Invoice approval related custom activities for a large consumer goods company

The event log already stores information on the user performing the respective event as event attribute.[466] By comparing the user-related information of the event log with the entity's hierarchy of authorization, the auditor may easily identify cases that have not been approved by an authorized individual without the need of disaggregated approval activities for each individual level of authorization.[467] Considering the information already stored as case or event attributes, the number of approval-related custom activities in Table 8 may be reduced from ten to two ("invoice approved" and "invoice rejected").

Another example relates to the eight shopping cart related activities included in Table 7 and summarized in Table 9.

[465] Cf. Chapter 3.2.2.
[466] Cf. Table 1.
[467] Cf. Figure 16.

Custom activity	Number of events	Number of cases	Percentage of total cases
Shopping cart approval workflow ended (document reviewed)	16.763	16.763	12%
Shopping cart approval workflow started	16.709	16.708	12%
Shopping cart created	16.622	16.622	12%
Shopping cart sub-workflow automatic approval	11.677	11.677	8%
Shopping cart approval workflow ended (approved)	5.233	4.757	3%
Shopping cart sub-workflow approved	1.872	1.759	1%
Shopping cart sub-workflow cancelled	335	272	0%
Shopping cart approval workflow ended (cancelled)	137	127	0%

Table 9: Shopping cart related custom activities for a large consumer goods company

As none of the custom activities defined for the entity[468] relates to the recording of a journal entry or the change of a document within the process, the entity likely makes use of the classic scenario[469] of the SAP SRM implementation. As such, the audit team may challenge whether:

(a) the information about the internal shopping cart is required from an audit perspective in addition to the standard activities “purchase order created”, “purchase order approved” and “purchase order rejected” and, if it is required, whether

(b) eight different activities are needed to describe the creation and approval (or rejection) of a shopping cart.

A lower level of the activities’ granularity than required significantly increases the complexity of the process graph, the resulting number of process variations and impedes all analyses involving the sequence of activities.

Overall, the empirical evaluation highlights the importance of reviewing the activities in the process, determining their appropriate level of disaggregation and critically evaluating their importance for the end-user.

4.1.4 Determining the process instance and its granularity

Determining the process instance

A fundamental prerequisite of assembling the input data for process mining is determining the process instance and its granularity. In current process mining solutions, choosing the process instance is equivalent to determining the document[470] whose processing throughout the activities in the process is analyzed. By this, the choice of a

[468] Cf. Table 7.
[469] Cf. Figure 48.
[470] Cf. JANS, MIEKE (2019), p. 61.

process instance directly impacts the analyses that can be performed and thus, the audit evidence that can be obtained.

Common process instances that are chosen in the purchase to pay process are the purchase order or the invoice. It is important to consider that a single purchase order can result in multiple invoices and a single invoice can cover multiple purchase orders. For example, if the process instance is determined to be the purchase order document, the purchase orders 1 and 2 may reference to the same invoice INV-001. As a result, the event „invoice INV-001 entered“ would appear in the process traces of both cases.[471] This needs to be considered when designing and performing the analyses, for example, when summarizing and analyzing the total invoice amount related to specific variations of the process. Similarly, if the process instance is the invoice, invoices 1 and 2 may reference to the same purchase order PO-001 and any events related to the purchase order PO-001 would appear in both cases.[472] This will impact, for example, analyses related to the purchase order approvals, changes and rejections but on the other hand facilitate analyzing invoice approval activities.[473]

Determining the granularity of the process instance

In ERP systems like SAP, the header and item information are stored in different tables at various locations.[474] This requires to not only select a process instance but also determine its granularity, i.e., to decide if the case that is followed through the process is the entire document (for example, the purchase order) or a single line of the document (for example, a purchase order line item).[475]

Like the process instance also its granularity has direct implications on the interpretation of process flows and the results of the analyses. In SAP, approval and rejection activities typically refer to the full document, while payments, goods receipts and invoices refer to an individual line of the document.[476] For example, to address the risk that purchases received are not recorded, CHIU/BROWN-LIBURD/VASARHELYI propose to inspect whether for all process instances the activities "goods receipt" and "payment" are performed.[477] As the authors chose the entire purchase order as a process

[471] Cf. JANS, MIEKE (2019), p. 61.
[472] Cf. ibid.
[473] Cf. ibid., pp. 59f.
[474] Cf. Chapter 4.2.1.
[475] Cf. JANS, MIEKE (2019), pp. 62f.
[476] Cf. ibid., p. 63.
[477] Cf. CHIU, TIFFANY/BROWN-LIBURD, HELEN/VASARHELYI, MIKLOS A. (2019), p. 56.

instance,[478] there may be partial deliveries and partial payments for one or more line items of the purchase order that need to be considered. If the entire purchase order document is selected as process instance, the number of goods receipts that are recorded for a process instance depend on the number of line items of the purchase order and the number of either partial or complete deliveries for these line items. In this scenario, the evidence of a goods receipt and a payment may still not evidence that the purchase has been delivered and recorded completely. Similar interpretation difficulties arise when an individual purchase order line item is chosen as the process instance and some activities, like approving the purchase order, are performed on the level of the full document.[479]

Implementation at the case audit firm and challenges identified

In order to develop a standardized process mining application that can then be scaled and customized to the specific entities, the decision on the process instance was made centrally during the implementation. The extensive implications of a predefined process instance have already been identified as part of the initial feasibility assessment. In the first prototype that only included limited adjustments made to the vendor software, the document used as a process instance was the purchase order. However, limiting the process instance to the purchase order systematically excludes those invoices from the analysis that do not relate to a purchase order. The analyzer that was finally implemented is based on a dynamic process instance with the aim of maximizing the parts of the process that can be analyzed.[480]

Still, the necessity to predefine the process instance may lead to acceptance challenges in practice, for example, when audit teams do not distinguish between direct purchases and purchases related to a purchase order when auditing the purchase to pay process using traditional procedures. If process mining demonstrates that the entity makes use of both types of transactions and the critical path of these transactions, including the control activities performed, differs significantly, the quality of the process understanding increases. However, in this scenario, additional audit procedures to identify and address related risks of material misstatement need to be designed. Acceptance challenges may also arise even when the entity exclusively uses the SAP MM and FI modules to process purchase-to-pay-related transactions. As many transactions posted

[478] Cf. CHIU, TIFFANY/BROWN-LIBURD, HELEN/VASARHELYI, MIKLOS A. (2019), p. 56.
[479] Cf. JANS, MIEKE (2019), p. 63.
[480] Cf. Chapter 3.3.1.

to trade payables are not referring to a purchase order or an invoice (for example, credit memos posted to the vendor account), even in this scenario, the trade payables will not be fully covered by the process data.[481]

As documents in an ERP system cannot be connected in a one-to-one relationship, even an individual line item of a purchase order may refer to multiple goods receipts and invoices and the invoices may be paid using multiple payment runs. Figure 49 illustrates the event log and the process graph of a purchase order line item with multiple deliveries.

Event ID	Activity	Quantity	Amount	User ID
1	Purchase order created	5	3.050	LJE
2	Purchase order approved	5	3.050	KXW
3	Goods receipt posted	3	1.830	KXW
4	Invoice posted	3	1.830	DIC
5	Invoice settled by payment run	3	1.830	MTF
6	Bank clearing posted	3	1.830	JHK
7	Goods receipt posted	2	1.220	LJE
8	Invoice posted	2	1.220	DIC
9	Invoice settled by payment run	2	1.220	MTF
10	Bank clearing posted	2	1.220	JHK

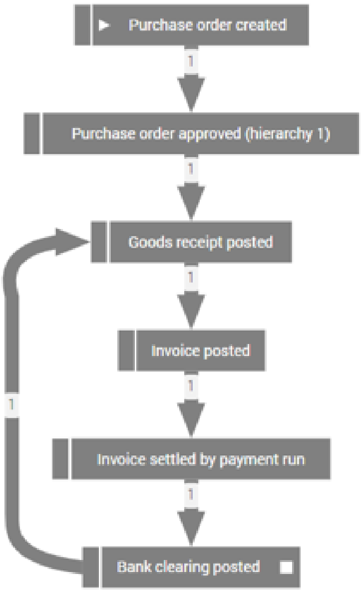


Figure 49: Event log and process graph for a case with partial deliveries

The purchase order line item includes five positions for which two separate deliveries and invoices have been received. The invoices have been settled by two different payment runs. For similar cases where the invoice posting, settlement and clearing (event IDs 8 to 10) of the second partial delivery (event ID 7) are missing, the process graph is identical. As such, the existence of a goods receipt event, an invoice posting event and a payment event within a specific case does not evidence that the purchase has been delivered, invoiced and paid completely. Similarly, if an activity is performed multiple times within a case, this does not necessarily represent a repetition of the same activity, for example, a duplicate payment. To conclude on the completeness of the goods receipts and invoices posted, the auditor needs to consider the quantity and price

[481] Cf. Chapter 4.2.1.

information related to the documents. That is, the quantity and price between the purchase order line item and the related goods receipt and invoice line items need to be matched (usually referred to as the three-way-match between purchase order, goods receipt and invoice) and finally compared to the payment.

Still, choosing an individual line item of a purchase order instead of the full document as process instance reduces the complexity of matching subsequent documents. However, interpretation difficulties exist for activities related to the purchase order, as the approval or rejection of a purchase order is performed at the level of the full document and not at the level of the individual line item.[482] Consequently, the event log includes events (for example, the approval of the purchase order document PO-100) that are, in fact, referring to multiple cases (i.e., each line item of the purchase order document PO-100). As a result, the activity "purchase order approved" is duplicated by the number of line items of purchase order PO-100.

It is important for the auditor to understand these relationships between cases and events in the event log. For example, as part of understanding the roles and responsibilities within the process, the auditor may investigate the distribution of purchase order approval events by user, as presented in Figure 50.

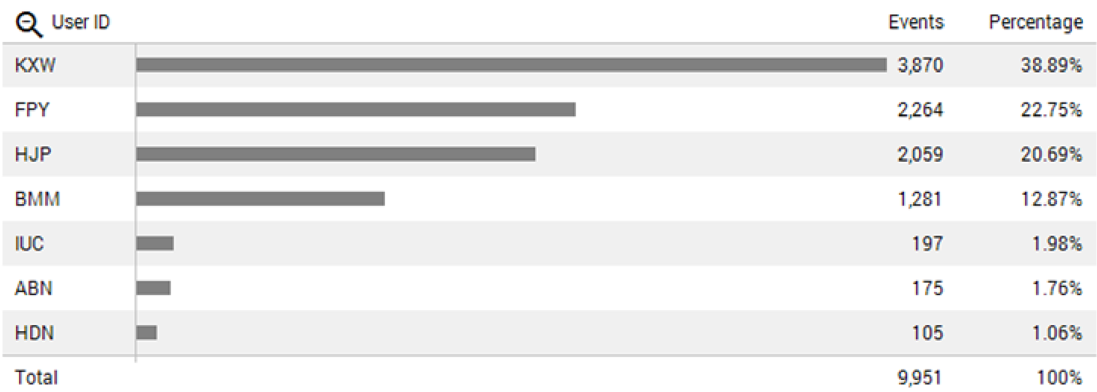


Figure 50: Number of purchase order approval events by user

In total, the dataset presented in Figure 50 includes 11.973 cases that relate to a purchase order (i.e., 11.973 purchase order line items). For 9.951 cases, a purchase order approval event is performed. However, filtering the cases passing through the "purchase order approved" activity results in the number of line items that have been approved, whereas the actual approval is performed on the level of the full purchase order document, i.e., for multiple cases together. Analyzing the underlying raw data reveals

[482] Cf. JANS, MIEKE (2019), p. 63.

the actual number of 4.272 purchase orders in the process, of which 2.390 purchase orders have been approved during the period for which the data has been extracted. Table 10 summarizes the actual number of purchase order approvals by user.

User	Purchase order approvals	Percentage
KXW	1.133	47,41%
FPY	683	28,58%
HJP	278	11,63%
BMM	156	6,53%
IUC	67	2,80%
ABN	65	2,72%
HDN	8	0,33%
Total	**2.390**	**100,00%**

Table 10: Number of purchase orders approved by user

While the users' ranking according to their number of purchase order approvals did not change for the dataset presented, audit teams need to be aware that the presentation of approval events is inherently biased by the number of line items existing for the purchase orders approved by a specific user.

A significant portion of training and coaching of the audit teams needed to be devoted to the implications of the selected process instance and its granularity on the audit procedures.

4.1.5 Timing and duration of events

Timestamp information and the impact on the process flow

Visualizing and analyzing the execution of a business process requires that the events within a case can be ordered, i.e., each event has a unique position within a process trace.[483] Usually, this ordering is achieved by including the date and timestamp information of events in the event log.[484] However, depending on the activity type and data source, the availability of date and timestamp information may vary for different activities. Challenges include the co-occurrence of events, missing timestamp information and differences between the actual time the event is performed and the timestamp in the event log.[485]

[483] Cf. Chapter 2.1
[484] Cf. ibid.
[485] Cf. VAN DER AALST, WIL M. P. (2016), p. 143.

Depending on the type and data source of the activity, there may be instances where events related to the same case occur at exactly the same time.[486] For example, a journal entry related to the posting of an invoice may be entered into the ERP system and at the same time set effective. Provided the entry and posting of an invoice is separated into two activities, the events “invoice entered” and “invoice journal entry posted” will have the same date and timestamp information. For these scenarios, a default logical ordering sequence of events needs to be determined in order to visualize the process flow. To deal with the co-occurrence of events, in the case audit firm’s process mining analyzer for purchase to pay, first any purchase order-related activities are displayed. Any goods receipt-related activities are assumed to come second. Thirdly, the invoice-related activities are displayed. In the fourth category are the settlement-related activities and, fifth and finally, the clearing activities are displayed. As a general principle to order activities on the same document (that have the same date and timestamp), the creation or entering event is always put first. Potential changes are assumed to come second and, lastly, the posting activities are set.

For some activities, only the date and no timestamp information may be available.[487] For example, in SAP, this will be the case when the journal entry is set effective at a different date than its entry date. While the system captures the entry date and time of the journal entry, only the date but no timestamp information is available for the effective date. Missing timestamp information needs to be added or the process mining algorithm needs to be adjusted in order to visualize the process flow. A practical solution might be to assign a “0:00” timestamp to these back- or forward-dated posting events. However, altering the timestamp information needs to be considered when performing time-related analyses. For example, analyzing the overall throughput time in the process between the activities “goods receipt entered” and “invoice journal entry posted” will be biased and analyses related to the sequence of activities will be less precise.

[486] Cf. VAN DER AALST, WIL M. P. (2016), p. 143.
[487] Cf. ibid.

Sometimes there may be significant differences between the actual time an event takes place and its timestamp in the event log, for example, because certain events are querying before being recorded.[488] As a result, the sequence of events for a given trace in the event log may be unreliable.

The challenges in capturing the date and timestamp information for certain activities need to be considered when analyzing the time perspective[489] of the process, including the throughput times between activities and the activity sequence (i.e., the ordering of activities). Especially for cases with short processing times, where many activities are performed in a short period of time, the course of business can appear unusual and it might include many concurrent events.

Challenges in determining the duration of activities

Events in an event log can have multiple timestamps, for example, a timestamp for the start of the event, a timestamp for the end of the event and a third timestamp calculating the duration of the event. Event logs may further distinguish between start activities and completion activities. For example, the activities presented in Table 6 include activities indicating the start and the end of a specific workflow in the ERP system.

The event log presented in Table 1 only includes a single timestamp attribute. Frequently, the duration of activities in an ERP system cannot be determined accurately. For example, when approving a purchase order in SAP, only the click on the approval button is recorded with a timestamp. The duration of the approval activity, i.e., the time spent by the approver to review the information of the purchase order, cannot be precisely measured. To facilitate interpretation, the approval activity is not labeled "approve purchase order" but "purchase order approved", as the timestamp indicates the date and time the activity has been completed.

In the example provided, the auditor cannot obtain evidence on the quality or diligence of the approval activity, for example, by identifying purchase orders that have been approved within seconds. Further, key resources involved in the process may only be identified by the number of activities performed but not by their actual time contribu-

[488] This scenario may occur, for example, in an X-ray machine where different components have local clocks and actions are querying before being recorded, cf. VAN DER AALST, WIL M. P. (2016), p. 143.

[489] Cf. Chapter 2.1.

tion to the process execution. Consequently, when information on the start and completion of events are missing, throughput times of individual activities and resource utilization cannot be measured accurately.[490]

4.1.6 Data extraction period and its implications on the analyses

From its initiation to the final settlement, a process instance might span over more than one reporting period, for example, when a case was processed in the audit period but was started in the previous period and completed in the period after the balance sheet date. VAN DER AALST points out that event logs usually provide just a snapshot of a longer running process and suggests removing incomplete cases from the input data.[491] However, on the one hand, this approach presumes that the initiating and the final activity of a process instance are known.[492] The purchase to pay process usually starts with the creation of a purchase order and ends with the payment of the invoice. However, there may be, for example, partial deliveries or payments being recorded for a purchase order, preventing a precise determination of a case's status as "open" or "closed" without analyzing the history of the case in detail.[493] On the other hand, removing incomplete cases might lead to the removal of cases that are relevant to the audit, for example, because they have (or should have had) a recording event in the audit period affecting the entity's account balances. Ensuring that all cases that are relevant to the audit are included in the analysis is critical to avoid systematical errors related to the completeness of the population. However, even the best data extraction strategy has its limits, for example, regarding server storage capacity, storage cost and performance of the ETL process. Determining the data extraction strategy and understanding the implications on the data (not) included in the process mining analyzer are two fundamental tasks when applying process mining in the audit.

The process mining analyzer introduced in Chapter 3 uses process data from up to five years. In general, all cases that have an event in the audit period under review are included. For these cases, the analyzer considers the events in the audit period and up to three years before the audit period. Depending on the data extraction date, the analyzer may also include events that have been performed in the period after the

[490] Cf. VAN DER AALST, WIL M. P (2016), p. 143.
[491] Cf. ibid.
[492] Cf. ibid.
[493] Cf. Chapter 3.2.2.

balance sheet date. The definition of the data extraction period is illustrated by the examples provided in Figure 51.

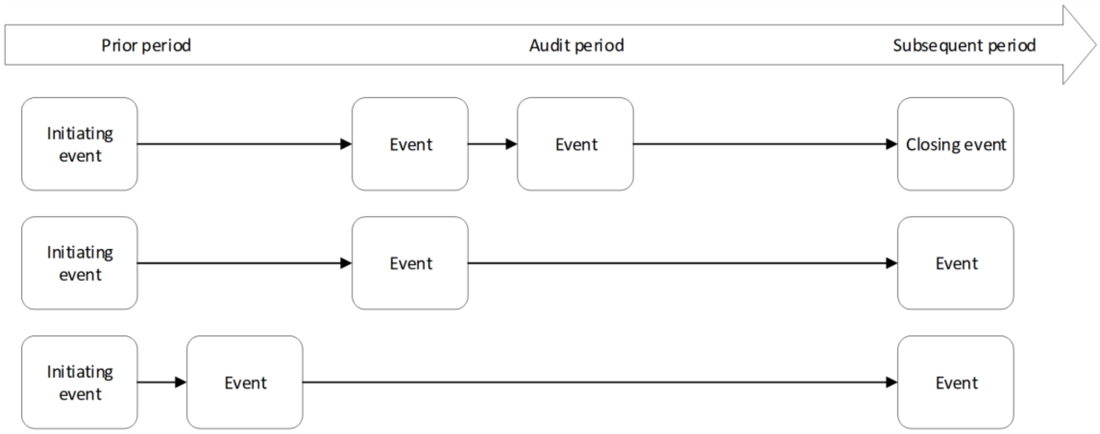


Figure 51: Data extraction period and inclusion criteria for cases in the process mining analyzer

The first two cases have events in the audit period under review and thus are included in the analyzer. It is not relevant whether the initiating event falls within the audit period or not and whether the case is already completed during the audit period or not. As there is activity in the audit period, both cases are included. The case in the third example does not have any events in the audit period. That is, no journal entries have been recorded and the case has not been processed in the audit period under review. This case is not included in the analyzer.

Even an optimized data extraction strategy is limited by the period for which data is (or can be) recorded and extracted. Understanding the data range included in the analyzer is especially important when it comes to long-standing purchase orders. A purchase order may have been created five years ago and there may still be activities performed related to that purchase order in the audit period under review.[494] For this process instance, only events within the past four years before the balance sheet date are included. The creation event of this long-standing purchase order that happened five years ago will not be included in the audit firm's analyzer. This will impact, for example, analyses related to the initiation of transactions as required by ISA 315.[495] As a consequence, the process mining solution needs to provide means to identify these transactions where the displayed initiating activity might differ from the actual initiation activity due to limitations in data extraction. In addition, if the purchase order

494 In practice, these long-standing purchase orders are frequently used for framework agreements with specific suppliers.
495 Cf. IFAC (2021), ISA 315 (Revised 2019), para. 25(a); Chapter 3.3.1.

document that is linked to the respective "purchase order created" event is not available in the data, further analyses may be impacted. For example, the invoice documents related to the process instance cannot be matched to the missing purchase order, resulting in errors or mismatches in the two-may-match of the quantity and price information between the purchase order and the invoice. Long-standing purchase orders are not uncommon in practice and exist at all entities that have been analyzed in the empirical evaluation. The limitations that are resulting for the analyses could be overcome by extracting way more data than actually needed, for example, to ensure that the initiating activities are complete and all invoices have a corresponding purchase order line item. However, due to technical challenges in practice, including storage cost, server capacities and possible impacts of the data extraction on the performance of the entity's IT system, such an extended extraction strategy has not been empirically evaluated so far. As a consequence, it is important for the auditor to understand the data extraction strategy, its limitations and resulting inclusion criteria for cases in the analyzer in order to avoid systematic errors when addressing completeness[496] risks.

4.1.7 Reliability of the data

Despite the progress made in the research discipline on process mining over the last decade, there is no established "best practice" for assembling the relevant input data and creating the event log yet.[497] However, as with any other automated tool or technique, the audit evidence that can be obtained with process mining depends on the quality and reliability of the input data. The event log is created from information produced by the entity. When using such information to perform audit procedures, ISA 500 requires the auditor to evaluate "*whether the information is sufficiently reliable for the auditor's purposes*"[498], including the completeness and accuracy of the information.[499]

---

[496] "Completeness" refers to both the completeness assertion for account balances and transactions derived from management's assertions as well as the completeness of information produced by the entity, cf. IFAC (2021), ISA 500, para. 9(a), also see Chapter 4.1.7.
[497] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 6.
[498] IFAC (2021), ISA 500, para. 9.
[499] Cf. ibid., para. 9(a).

Completeness of the input data

Completeness refers to the risk that relevant process data is missing in the process mining application, resulting in an incomplete and misleading process model. For example, if the auditor tests a population of transactions for appropriate authorization, the audit evidence obtained is less reliable if the population of transactions is not complete.[500]

The completeness of the process model depends on:

- the degree of digitalization and automation of the analyzed process,
- the information that is actually recorded and stored in the relevant IT systems and
- the procedures performed in extracting and transforming the raw data and loading the input data and into the process mining application.

An inherent limitation of process mining (and every other automated tool or technique) is that manual procedures performed without the use of IT are not recorded in the data.[501] For example, in a purchase to pay process, the ordered materials are usually checked for quantity and quality upon receipt.[502] However, if the physical receipt of purchased goods and related checks performed by the warehouse staff leave no digital trace in an IT system, these procedures cannot be captured in the event log. In this scenario, although the auditor might expect that goods are received and inspected before the goods receipt is entered into the system, the activity "goods receipt entered" does not support inferences to related manual activities that may have been performed. As the parts of business processes that are not embedded in an information system decrease as digitization continues, the application scope of process mining techniques grows continuously.[503] However, despite using process mining, the auditor still needs to perform traditional procedures, including inquiries and observations, to identify procedures performed outside the IT system and determine their relevance to the audit.

Another inherent limitation of process mining is that the information that is recorded and stored by the source system may depend on specific decisions and business practices of the entity. ERP systems can automatically record an extensive set of variables

[500] Cf. IFAC (2021), ISA 500, para. A60.
[501] Cf. WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021), p. 13; JANS, MIEKE (2012), p. 396.
[502] Cf. ELDER, RANDAL J. et al. (2020), p. 571.
[503] Cf. JANS, MIEKE (2012), p. 396.

and meta-data about the accounting relevant user entries.[504] However, as the computational demand increases with the extent of the information that is logged, the logging is frequently limited by having some of the logging capability disabled.[505] For example, instead of storing the entire history of change entries made to a document, an entity may decide to only record the latest entry made.[506] In this regard, completeness refers to the risk that not all transactions or other events that should be recorded have been recorded. As the adoption of process mining in practice increases, though, the auditor and the entity's internal audit function (or those charged with governance) might have a joint interest in increasing the scope and extent of the data that is tracked and recorded by the IT system.[507]

The completeness of the process model may further be affected by the procedures performed in extracting and transforming the raw data and loading the input data into the process mining application. Decisions made in determining the scope and extent of the data extraction directly impact the audit evidence that can be obtained from analyzing the population of process traces in the process mining application. For example, the third case presented in Figure 51 does not have any events in the audit period, i.e., no accounting documents have been recorded and the case has not been processed in the audit period under review. As defined in the data extraction strategy, this case is not included in the process mining analyzer. However, the auditor may have identified the risk that acquisition transactions occurred in the audit period but have not been recorded in the system. Hence, it is important for the auditor to understand the data extraction strategy and resulting inclusion criteria for cases in the analyzer to avoid systematic errors when addressing completeness risks. Similarly, decisions made in transforming the raw data into an event log and loading the input data into the process mining application may affect the completeness of the process model and impact the possible audit evidence. This not only includes the risk that data may be lost in the transfer, but also decisions made in assembling the input data, such as determining the process instance, the relevant activities, the attributes and meta-data about the cases and events and the references that are established between the individual data sources.

[504] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 16.
[505] Cf. ibid.; JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), p. 29.
[506] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 7.
[507] Cf. ibid., p. 15.

Accuracy of the input data

The accuracy of the process model depends on:

- the appropriateness of the recording of transactions and other events at the entity,
- the prevention of altering the log files of the information system,
- procedures performed in manipulating the data in creating the event log and
- the process mining algorithm used.

The accuracy of the event log may be impacted by information that is entered manually into the system. For example, user-entered parameters may be inaccurate due to typing errors when entering documents or missing document references when posting journal entries. Errors affecting the accuracy of data can also result from hardware failures or malfunctioning interfaces between different IT applications.[508] Further, even if recorded appropriately in the system, the log files may not accurately reflect the circumstances under which a transaction has been entered into the system. For example, an employee is usually required to login with an individual user ID and password before performing actions in the ERP system. The entity may have policies in place that manage, for example, regular password changes and locking the personal computer when leaving the workplace. However, these policies do not prevent that theoretically, the place of an employee could have been taken over by another person.[509] While this scenario is not unique to process mining, the user information of the related events would be inaccurate and bias the analysis of roles, responsibilities and segregation of incompatible duties in the process.

The potential of process mining to add value for auditing is especially driven by the reconstruction of the process based on data, including meta information about the data entered by the auditee that is automatically and independently logged by the system.[510] As the audit evidence obtained using process mining (or any other automated tool or technique in auditing) relies on the immutability of the input data,[511] mechanisms need to be in place that prevent users from modifying or overriding the automated logging by the IT applications. This includes, for example, that users involved in processing transactions do not have the privileges of a system administrator that would enable

[508] Cf. GEHRKE, NICK/WERNER, MICHAEL (2013), p. 4.
[509] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 14.
[510] Cf. Chapter 2.1.
[511] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A169.

them to override controls over the recording in the system as well as the prevention of manipulation.[512] As part of understanding the components of the entity's system of internal control, ISA 315 requires the auditor to identify the risks arising from the use of IT and the entity's general IT controls that address these risks.[513] If the general IT controls are not operating effectively or the auditor determines not to test the operating effectiveness, substantive procedures need to be performed that address related IT risks.[514]

With the increasing use of data analytics in the audit, audit firms start to develop competencies in data enablement, including the extraction, transformation and loading of data for use within an automated tool or technique. While audit firms are familiar with data enablement for financial data, data requirements for process mining are significantly more extensive. The process mining input data needs to be assembled from multiple data sources stored at various locations in the entity's IT systems, merged, manipulated and re-structured in order to create the event log.[515] This process requires the involvement of data scientists and IT experts and is prone to errors. Errors may result from both the computations or categorizations performed in the creation of the event log and the information created, added or changed within the process mining analyzer. For example, differences in the system time of multiple data sources may lead to an inaccurate ordering of events in a log file assembled from these data sources.[516] During the audit firm's feasibility assessment, many audit teams reported that the dashboards and analyses have been customized by an IT expert in parallel to the audit and several iterations of adjusting the analyses provided in the first prototype were required. The feedback includes occasions where issues resulting from a missing audit and accounting background of the data scientist preparing the analyzer have been identified.[517] For example, in a calculation comparing the purchase order price with the invoice price of the cases, differences resulted due to the use of the initial purchase order price at the time of the creation of the purchase order rather than the adjusted and approved purchase order price. Multiple company codes with different local currencies may result in the aggregation of different currencies if not carefully considered

---

[512] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2013), p. 9.
[513] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 26(c).
[514] Cf. ibid., para. A166.
[515] Cf. Chapter 4.1.1; VAN DER AALST, WIL M. P. (2016), p. 155.
[516] Cf. VAN DER AALST, WIL M. P. (2016), p. 143.
[517] Cf. Chapter 3.1.2.

in the design of the analyses in the process mining application.[518] Errors or assumptions made in preparing the data directly impact the interpretation of the analyses and conclusions drawn from procedures performed in the process mining analyzer. The case audit firm reacted to the risk of inappropriate data manipulation with a standardized process mining analyzer and central quality assurance procedures for any customizations made to ensure data integrity.

To reconstruct the process graph from an event log, process mining makes use of specific mining algorithms. A large body of process mining literature deals with the development and evaluation of process mining algorithms for process discovery.[519] However, the quality of the process models discovered with these algorithms varies significantly.[520] Algorithms used in an audit of financial statements may not provide false negative results and should provide as few false positive results as possible.[521] If the process model does not include all process executions evident in the event log, the algorithm systematically leads to under-auditing, as material misstatements or control exceptions may remain undetected. On the other hand, any process paths that are created by the mining algorithm but did not occur in practice are not relevant to the audit. Consequently, before applying any process mining solution in the audit, the quality and suitability of the mining algorithm needs to be validated.[522] This validation requires expert knowledge in both the domain of process mining and the domain of auditing.

Implications of the challenges regarding data reliability

Dealing with information produced by the entity is not new for the auditor. Related risks need to be addressed by the auditor as with any other automated tool or technique used to analyze information produced by the entity.

Evidence about the completeness and accuracy of information may be obtained by testing controls over the preparation and maintenance of the information, by designing and performing additional audit procedures or concurrently with performing the actual audit procedure that is using the information.[523] These procedures may involve tests of

518 Cf. Chapter 3.2.2.
519 Cf. Chapter 2.2.
520 Cf. ROZINAT, ANNE et al. (2008), p. 84.
521 Cf. WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021), p. 13.
522 Common evaluation criteria for the quality of a discovered process model are the fitness, precision, generalization and structure of the model. An overview is provided by ROZINAT, ANNE et al. (2008), p. 88.
523 Cf. IFAC (2021), ISA 500, para. A61.

details or any other type of procedure, including analytical procedures and reperformance.[524] For example, when considering the general ledger data in the process mining application, any financial data used may be reconciled to the financial statements. The coverage of the transaction volumes on the financial accounts relevant to the process by the process data, i.e., by the activities initiating journal entries, helps to evaluate the completeness and accuracy of the process model. Similarly, the audit procedures performed with process mining help to evaluate the reliability of non-financial data. ISA 500 explicitly names the reperformance of 100 percent of the items for testing information produced by the entity.[525] With process mining, this reperformance is done on the level of an individual case (i.e., a document). For example, the three-way-match[526] is reperformed for 100 percent of the transactions by matching quantities and prices for each set of purchase order, goods receipt and invoice. The digital traces for event data typically only exist in a digital format. In a traditional audit procedure of selecting specific items or in audit sampling, the procedure would be the same – but it would only be performed for a fraction of the total population of purchases. In addition, even in a traditional testing setup, the match between an invoice and a purchase order is performed by taking screenshots from the system with the same information that is used in process mining.

Although the examples demonstrate how procedures with regard to the reliability of information may be performed with process mining, the complexity of data preparation and the challenges regarding the reliability of the data reduce the relevance of process mining for the broader audit practice significantly. Extensive quality assurance is not only required with regard to the design of the analysis in the process mining application but also on the level of each individual dataset. If the risks related to the completeness and accuracy of the input data are not appropriately identified and addressed, the auditor may not obtain reliable audit evidence.[527] For example, without having validated data integrity, the auditor may not determine if a control deviation is the result of (a) an inappropriate process execution, (b) an incorrect event log or (c) an incorrect discovery of the event log by the process model. If the auditor has not obtained a detailed understanding of the entity's information system and the procedures

[524] Cf. IFAC (2021), ISA 500, para. A20 and A21.
[525] Cf. ibid., para. A53.
[526] Cf. Chapter 3.2.1.
[527] Cf. IFAC (2021), ISA 500, para. A60.

performed in preparing the event log, the evaluation of deviations may further be falsified by (d) an incorrect interpretation of the process mining analysis.[528]

4.1.8 Development of a solution approach for the challenges in data preparation

Summary of decisions and challenges in preparing the input data

Table 11 summarizes the decisions and challenges in setting up the process mining application and in customizing the analyzer to the specific auditee.

Decision or challenge	Implications for the setup of the process mining analyzer	Implications for the audit team
Understanding the IT landscape	To assemble the input data, extensive knowledge about the ERP system and interfaces to other relevant IT applications is required. Locating and scoping relevant data sources requires a rigorous and defensible method of structuring the data.	The data fields included in the data model directly impact the information content of process mining, the interpretation of the analyses and the possible evidence that can be obtained.
Understanding entity specific customizations	Due to the heterogeneity of IT landscapes, creating a repository of ERP systems and other relevant IT applications with all configuration and customization options available is not feasible in practice. The data model needs to the adjusted individually to account for entity specific data storage and relevant customizations.	The ability of the data model to appropriately reflect the entity's specific configurations and customizations impacts the reliability of the input data. Understanding the customizations and resulting data requirements required needs to be facilitated by an IT specialist.
Data transformation using relational databases	The information required to create the event log needs to be located, extracted and transferred into a data model suitable for process mining. ERP systems and traditional process mining solutions use relational databases that require a cumbersome preparation and transformation of the data. Any missing links between the data sources need to be manually created by complex SQL queries as necessary for the analyses.	The degree of integration of the data sources in the data model pre-determines the analytical possibilities. If links between data sources are missing, they may not be analyzed in conjunction. As a result, filters are lost when switching between analyses using different data tables as input. Inefficient or redundant data storage in relational databases may further lead to performance issues for larger datasets.
Determining relevant activities	The information on the activities relevant to the process and their source systems is a prerequisite of setting up the process mining analyzer. The relevant activities may vary by auditee even if the same ERP system is used. Each activity needs to be defined from a technical perspective to extract the relevant raw data from the entity's system.	Audit teams need to understand the standard activities covered by the data model to determine additional custom activities that may be required for the specific entity. As there is a trade-off between information content and complexity of the analyses, an activity's relevance to the audit needs to be carefully evaluated.
Determining the process instance and its granularity	In ERP systems like SAP, the header and line item information of documents is stored in different tables. This requires to not only select a process instance but also determine its granularity, i.e., to decide if a case is the entire document or a single line item of the document.	The auditor requires a holistic understanding of the process instance selected and its impact on the analyses that can be performed, on the interpretation of results and on the audit evidence that can be obtained.

[528] With the central development of a standardized process mining application, the case audit firm ensured that a large portion of quality assurance procedures, including a technical review of the application and a certification of the ETL process, could be performed and documented centrally. The audit teams are provided not only with the process mining application but also control templates documenting the data preparation, including the validation of the input data used.

Decision or challenge	**Implications for the setup of the process mining analyzer**	**Implications for the audit team**
Timestamp information	Depending on the activity type and data source, the availability of date and timestamp information may vary for different activities. Challenges include missing timestamp information, the co-occurrence of events and differences between the actual time the event is performed and the time in the event log.	Gaps between the actual event time and the timestamp in the event log or a reconstructed ordering of events that does not reflect reality result in an unreliable sequence of activities. The reliability of date and timestamp information of events directly impact the interpretation of the analyses.
Duration of activities	Events can have multiple timestamps, for example, a separate timestamp for the start, the end and the duration of an event. Event logs may further distinguish between start activities and completion activities. The timestamp-related information available for each activity needs to be considered when designing the analyses.	When information on the start and completion of events are missing, durations of individual activities, throughput times between activities and resource utilization cannot be measured accurately.
Data extraction period	Determining the data extraction period is a fundamental task when using data analytics. Even the most optimized data extraction strategy has its limits with regard to server storage capacity, storage cost and performance of the ETL process.	The data extraction period may impact the completeness of the process data. It is important for the auditor to understand the data extraction strategy and resulting inclusion criteria for cases in the analyzer to avoid systematic errors when addressing completeness risks.
Completeness and accuracy of the data	The completeness of the process model depends on (1) the degree of process digitalization and automation, (2) the information actually recorded and stored in the IT systems and (3) the ETL process. The accuracy of the data is impacted by (1) the appropriateness of user-entered parameters, (2) measures preventing data alteration, (3) manipulations of the input data and (4) the process mining algorithm used.	The completeness and accuracy of the information needs to be evaluated by the auditor. The analyses that can be performed and the evidence that can be obtained with process mining depend on the reliability of the input data.

Table 11: Decisions and challenges in setting up and customizing process mining in practice

Working towards business process management "in the large"

The knowledge, procedures and decisions required for assembling and manipulating the input files for process mining pose risks to the completeness and accuracy of the data on the one hand and reduce the relevance and application areas of process mining in the audit on the other hand. Current process mining techniques further presume that processes are recorded automatically and in a consistent manner, can be summarized in a log file with a standardized structure and have only very limited interrelations with other processes.[529] These assumptions are also referred to as "BPM in the small".[530] In evaluating the use of process mining in an "BPM in the large" context, where the assumptions are lifted, ACCORSI/ULLRICH/VAN DER AALST conclude that concurrent process mining techniques need to be enhanced and developed further to reflect the heterogenous, collaborative and flexible processes in practice.[531]

[529] Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 359.
[530] Cf. ibid.
[531] Cf. ibid.

In contrast to workflow management systems, ERP systems are not structured in separate workflows that are recorded in individual log files and thus easily accessible for an analysis with process mining. However, the financial impact of business transactions and events in an ERP system is recorded in a structured ledger. In this regard, the general ledger may be interpreted as a chronological record of journal entries representing the conjoint impact of all accounting relevant business processes throughout the audit period on the financial statements. With an increasing complexity of organizations' IT landscapes and the increasing use of data analytics not only by internal and external auditors but also for managerial decision making, it may be worth for ERP providers to consider not only logging the financial journal entries but introducing an event ledger.

A process-oriented event ledger chronologically recording the execution of activities in a business process that lead to the recording of a journal entry in the general ledger may overcome many limitations of relational databases. This may include suspending the requirement to obtain an extensive understanding of the process, related IT systems and entity specific configurations and customizations in advance of using process mining.[532] The auditor may no longer require expert knowledge on how data is stored and linked across multiple tables and modules of the ERP system. Further, the auditor may avoid many decisions that are currently required when assembling and scoping the process mining input data but that restrict the scope of subsequent process mining analyses.

The event ledger might either be stored for each process individually or jointly consider the activities performed in all processes relevant to financial reporting. A joint event ledger logging the process data for all of the entity's business processes may be interpreted similar to the general ledger that records the financial data as an outcome of all accounting relevant processes. In combination, both data sources would have the potential to enhance concurrent process mining techniques to reflect the actual heterogeneity and collaboration of today's business processes.[533] A holistic analysis of the entity's business processes would further enhance the application scope of process mining with advanced analyses about the relationships, dependencies and interactions between different business processes. For example, the distribution of a specific good

[532] Cf. Chapter 4.1.2.
[533] Cf. Accorsi, Rafael/Ullrich, Meike/van der Aalst, Wil M. P. (2012), p. 359.

is only possible if the quantity of goods in stock in the warehouse is greater than or equal to the ordered quantity,[534] and resource utilization may only be measured precisely when considering all processes an individual is involved in. Based on the relationships between the activities in the process, activities connecting the initiation, recording, processing and reporting of the same case types may be identified, enabling a data-driven identification of business processes and classes of transactions.

Considering the use of graph databases instead of relational databases

The list-format of the general ledger is only one possible representation of the journal entries reported in the financial statements. MOCHTY first introduces the representation of the double-entry bookkeeping as a graph.[535] The so-called “journal entry network” or “accounting network” illustrated in Figure 52 visualizes the total population of general ledger postings, i.e., the transactions flows between the general ledger accounts.

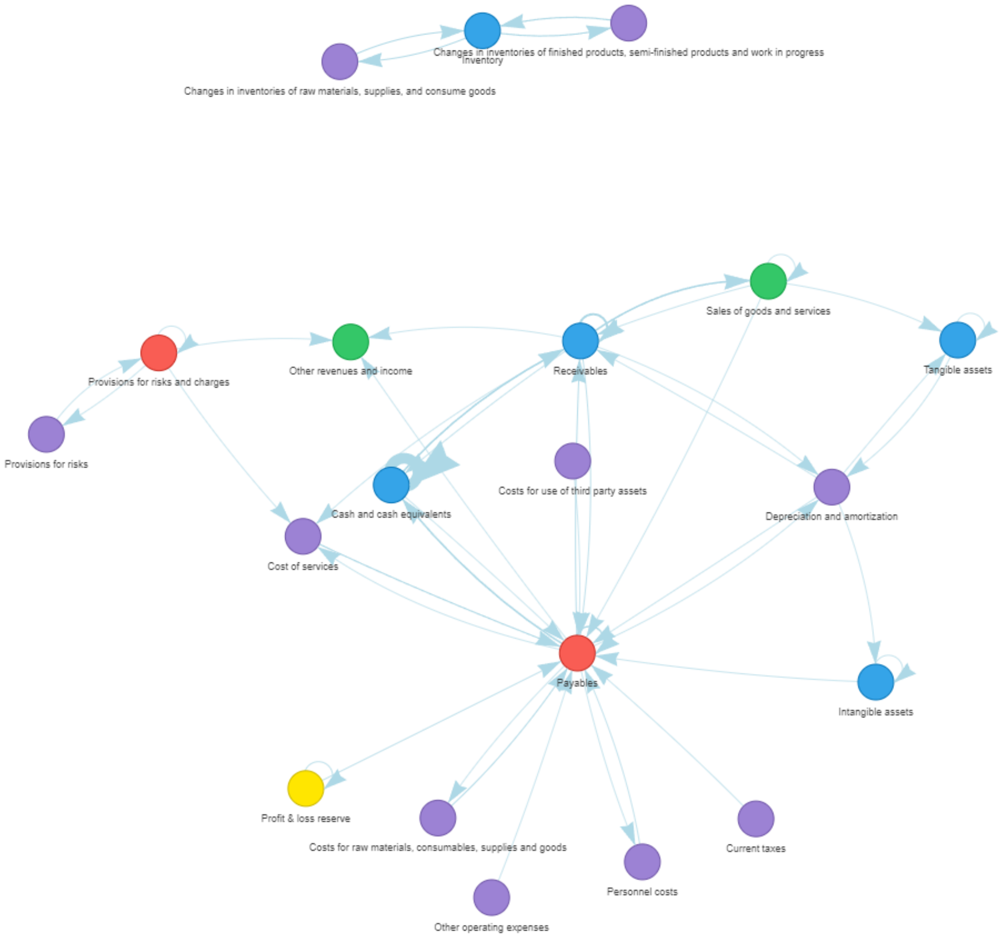


Figure 52: Accounting network visualizing the general ledger as a graph[536]

[534] Cf. MOCHTY, LUDWIG (2015), p. 32.
[535] Cf. MOCHTY, LUDWIG (1985), pp. 2-15.
[536] Cf. MOCHTY, LUDWIG/WIESE, MICHAEL/MAAS, ANNALOUISE (2022).

The nodes of the network represent the account classes summarizing the individual financial statement accounts. The color of a node represents its account type.[537] The edges in the network visualize the total transaction volume between the related account classes. The direction of the edges follows the debit-credit booking systematic, visualized by an arrow from the debit account class to the credit account class. The size of the connections represents the total transaction volume processed between the related account classes.[538]

In analogy to representing the financial data as a graph, the event log (or the event ledger proposed earlier in this chapter) may be represented as a graph as well. Graph databases store information as nodes and edges as opposed to rows and columns[539] in relational databases. Figure 53 shows the same information stored in a graph database and in a relational database and compares the adjustments of both databases that are necessary to add new information.

[537] Green nodes represent revenue account classes, purple nodes represent expenses, blue nodes represent assets, red nodes represent liabilities and yellow nodes represent equity account classes.

[538] Refer to MOCHTY, LUDWIG (1985) and MOCHTY, LUDWIG/WIESE, MICHAEL/MAAS, ANNALOUISE (2022) for further details on the journal entry network and its application in an audit of financial statements.

[539] Cf. Chapter 4.1.2.

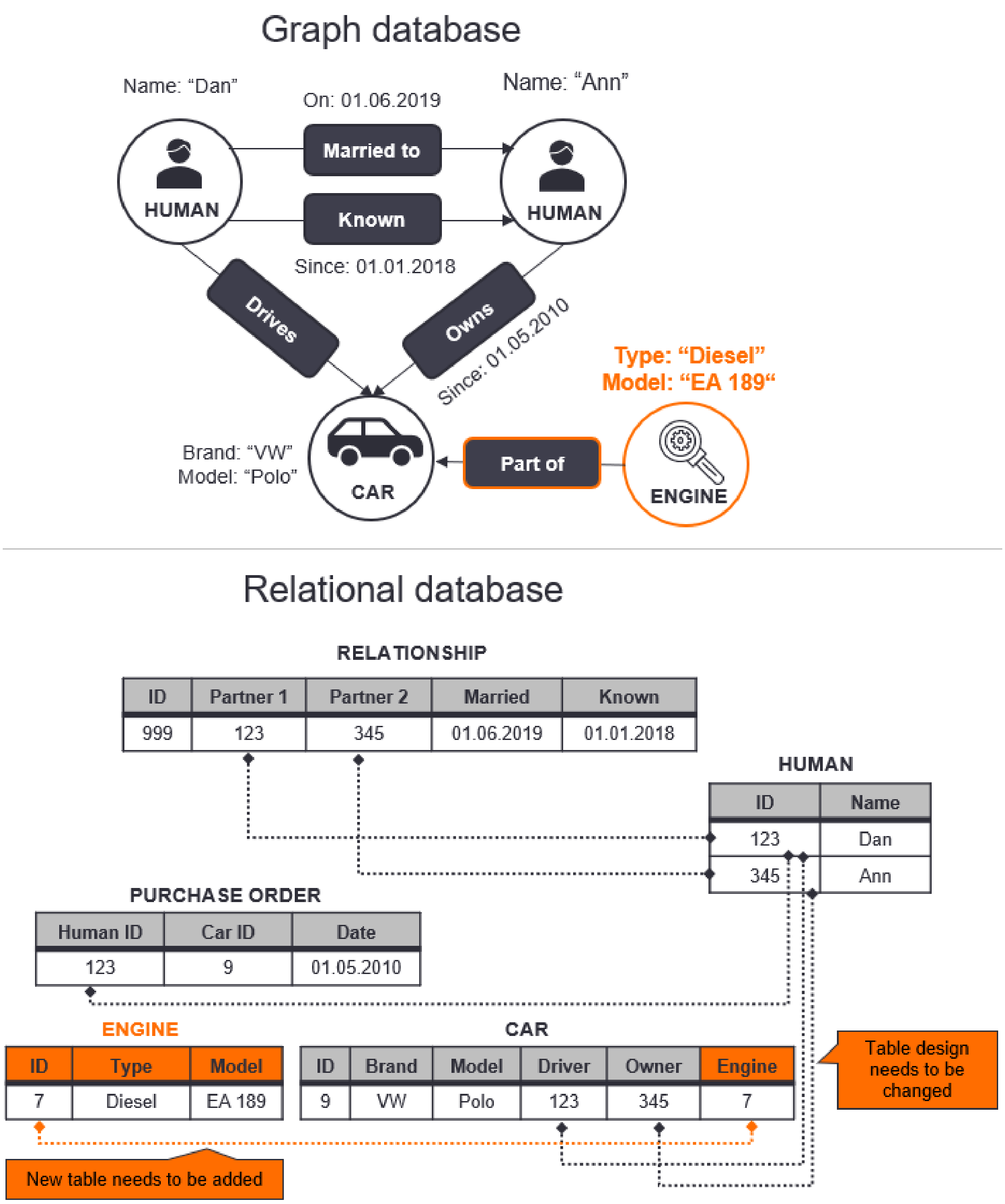


Figure 53: Adding new information to a graph database and a relational database

In the example provided, the graph database stores information on two individuals and a car as nodes and the relationship between these nodes as edges. Both the nodes and the edges have multiple attributes describing their characteristics. In the relational database, the same information is stored in different tables. The relationship between two tables is established by a reference field shared between the tables.[540] Figure 53 illustrates in orange color how new information may be added to both data sources. The engine type and model of the car is added to the graph database by adding a new node

[540] Cf. Chapter 4.1.2.

and edge linking the engine to the car it belongs to. In contrast, in a relational database, any information needs to be stored and modeled for each analysis perspective that is provided to prevent filters getting lost when switching between analyses that use different data sources.[541] That is, a new table needs to be added that stores the information on the engine type and engine model. In the table storing the information about the car, a new column is added that is used to manually establish the link between the car and the engine table based on the unique engine ID. The example illustrates that graph databases offer an easier way to establish links between different data sources. This would not only contribute to the level of data integration within the process mining data model but also facilitate its maintenance. Compared to relational databases, graph databases further offer a higher flexibility, scalability and availability at lower operating cost due to more efficient data storage.[542]

The consideration of using graph databases is not new to the process mining domain but has been introduced earlier under the concept "financial process mining".[543] The basic functionality is described by GEHRKE/MÜLLER-WICKOP.[544] Financial process mining does not rely on log files of process data but reconstructs the process instances based on accounting data from the ERP system to explain how financial entries have been produced.[545] The technique uses the clearing mechanism of open item accounting in an ERP system to establish relationships between accounting documents and items.[546] Figure 54 shows the structure of a purchasing process instance mined with financial process mining.

541 Cf. Chapter 4.1.2.
542 Cf. IBM CLOUDANT (2015), p. 5.
543 Cf. Chapter 2.2; GEHRKE, NICK/MÜLLER-WICKOP, NIELS (2010), p. 2.
544 Cf. GEHRKE, NICK/MÜLLER-WICKOP, NIELS (2010), p. 2.
545 Cf. MÜLLER-WICKOP, NIELS/SCHULTZ, MARTIN (2013), p. 106.
546 Cf. GEHRKE, NICK/MÜLLER-WICKOP, NIELS (2010), p. 1.

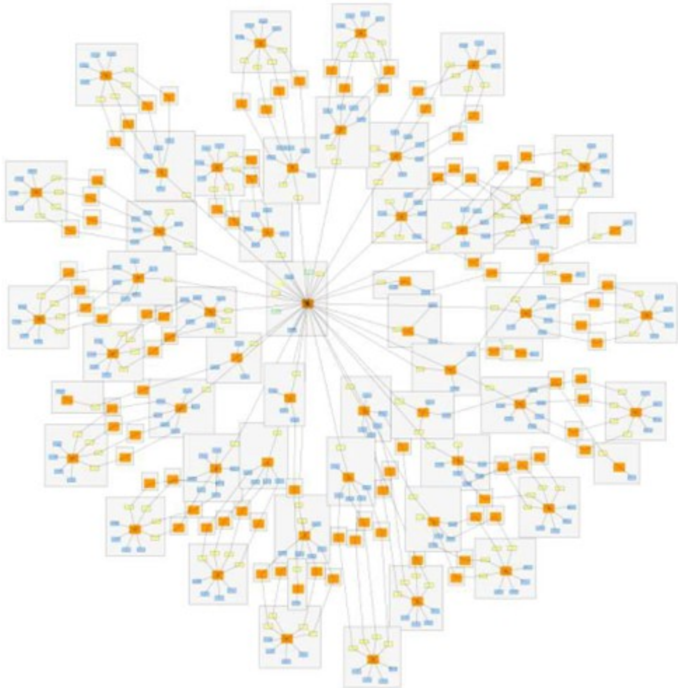


Figure 54: Structure of a process instance mined with financial process mining[547]

The accounting document in the center of the graph represents a payment journal entry clearing several invoices from a specific vendor. To construct the directed graph of documents and items, the financial process mining algorithm uses a single document ID as input and mines the documents and items in the data population to obtain the complete process instance.[548] While financial process mining enables to analyze the reporting of accounting documents in the financial statements, it is limited to preselected accounts that need to be subject to open item accounting.[549] Further, as the same account may appear multiple times, i.e., as often as it is referred to in a journal entry line item, aggregation rules are necessary to reduce complexity of the resulting graph.[550] Although the concept is presented under the team financial "process" mining, the algorithm exclusively considers financial data.

While the process instances resulting from financial process mining constitute directed graphs, the approach is not based on the idea to use graph databases for process mining but to integrate financial information in the event log. In later research, MÜLLER-WICKOP/SCHULTZ propose an approach to transform the mined process instance graphs into the sequential event log format.[551] In summary, the authors aim to make

547 Cf. GEHRKE, NICK/MÜLLER-WICKOP, NIELS (2010), p. 8.
548 Cf. ibid., pp. 3f.
549 Cf. MÜLLER-WICKOP, NIELS/SCHULTZ, MARTIN (2013), p. 109.
550 Cf. MÜLLER-WICKOP, NIELS/SCHULTZ, MARTIN/GEHRKE, NICK/NÜTTGENS, MARKUS (2011), pp. 127f.
551 Cf. MÜLLER-WICKOP, NIELS/SCHULTZ, MARTIN (2013), p. 105.

the process instances derived from accounting data available for traditional process mining techniques that require a timestamped event log.[552] However, reverting back to the event log format to enable traditional process mining analysis does not remove the challenges and decisions in building the event log summarized in Table 11.

While the accounting network and financial process mining do not consider the process perspective, that is, initiating and processing activities within the process, current process mining techniques do not consider the financial perspective, i.e., the resulting journal entries reported in the financial statements. A network representation of both data sources using a graph database has the potential to not only overcome the limitations of relational databases but connecting process and financial data in an effective manner for a jointly investigation. MOCHTY first explores the benefits that enhancing the journal entry network with event log data may provide to the auditor.[553] As will be shown in Chapter 4.2 of this thesis, a connection of both data sources may be established using the information and document references in the accounting document, i.e., the journal entry header and line item information stored in the ERP system. As there is a finite number of journal entries relevant to the audit period under review, theoretically, the related process data relevant to the audit may then be identified and extracted automatically based on their connection to the journal entries, helping to overcome the limitations with regard to the data extraction period.[554] Within this graph database combining the process and financial data, the nodes might represent the journal entries, documents and users in the process, while the edges may indicate their relationships. These relationships may not only include the links between documents and their related journal entries that are established based on the shared accounting reference, but also the activities of the process. For example, the nodes of "user A" and "invoice document 1" may be linked by the edge "invoice entered".

In summary, the missing explorative character of today's process mining solutions leads to significant set-up costs when applying process mining in the field of auditing. These costs are perceived as a blocker for adoption even when compared to the increased quality of the process understanding. As decisions made in data preparation directly impact the audit evidence that may be obtained with process mining and many

---

[552] Cf. MÜLLER-WICKOP, NIELS/SCHULTZ, MARTIN (2013), p. 105.
[553] Cf. MOCHTY, LUDWIG (2015), p. 46.
[554] Cf. Chapter 4.1.6.

challenges relate back to the relational databases used, further theoretical and empirical research is required.

On the one hand, research is required to enhance assembling and extracting relevant process mining input data and establishing required links between related data sources. This research may involve maintaining an event ledger in the ERP system to record process information in a systematic and structured manner, similar to the recording of financial information in the general ledger. Alternatively, advanced algorithms may help in the future to automatically extract data having the nature of an event from the entity's IT systems. The use of a graph database to store and present the event data that automatically links data including the same references may further help to overcome the challenges identified with regard to relational databases.

On the other hand, current process mining solutions are based on relational data models. Although with the increasing adoption of automated tools and techniques in the audit practice, audit firms are developing competencies regarding the preparation and use of data analytics, it is not likely that the future auditor will work with queries to retrieve information from graph databases. Further research will be required to explore designing process mining analyses and audit procedures for evaluating information stored as a graph in an effective and efficient manner. As the transition from descriptive process mining techniques to an explorative "mining" of the entity's IT systems is subject to extensive future research, the considerations in the following chapters will contribute to the further development of concurrent process mining techniques relying on relational databases.

4.2 Missing integration of the process data and the financial data

4.2.1 Integration of general ledger data to assess the coverage of the process

Identification of transactions volumes covered by the process data

As current process mining techniques do not consider the financial aspect of processes,[555] a fundamental task in building a process mining application for use in an audit of financial statement was integrating general ledger data.[556] The integration was performed in two steps. First, the general ledger data, including the trial balance and

[555] Cf. MOCHTY, LUDWIG (2015), p. 31; GEHRKE, NICK/MÜLLER-WICKOP, NIELS (2010), p. 1.
[556] Cf. Chapter 3.1.1 and 3.2.1.

the journal entries exported from the entity's ERP system, has been imported into the process mining analyzer. Secondly, the process data has been linked to the financial data in order to assess the coverage of the transactions and account balances by the analyzed process. This reconciliation becomes even more important provided the challenges regarding the completeness of the event log.[557]

Figure 27[558] shows the "General ledger activity" dashboard that is based on the general ledger data exported from the system. The dashboard shows the balance sheet and income statement accounts and the related debit and credit movements throughout the period that can be reconciled to the balance sheet and income statement prepared by the entity.

With the general ledger data being available in the process mining analyzer, the process data can be linked to the general ledger account movements. Figure 55 shows the coverage of the transaction volumes of the trade payables and inventory account classes by the process data.

GL account type	GL account class	Coverage	Debit	Credit	▼ Net activity (LC)
Liabilities	4-CL2999_TO TRADE PAYABLES	Not covered	125,930,499 (45.33%)	121,121,966 (42.97%)	4,808,532
		Invoice - Without PO	10,119,052 (3.64%)	8,978,406 (3.19%)	1,140,646
		Invoice - With PO	141,743,120 (51.02%)	151,782,608 (53.85%)	-10,039,488
Assets	3-CA3999_TO INVENTORIES	Invoice - With PO	80,603,723 (67.09%)	13,563,648 (10.92%)	67,040,075
		Invoice - Without PO	7,547 (0.01%)	7,547 (0.01%)	0
		Not covered	39,534,626 (32.91%)	110,678,219 (89.08%)	-71,143,593
		Total	397,938,567	406,132,394	-8,193,827

Figure 55: Coverage of the trade payables and inventory account classes by the process data

In contrast to the account coverage analysis introduced in Chapter 3.2.1, the analysis in Figure 55 considers the coverage separately for each case type, i.e., SAP MM and SAP FI transactions. Next to the information on the account type and class, the dashboard shows the debit, credit and net movements of the balance sheet and income statement accounts[559] throughout the period. The part of the debit and credit movements that is covered by the process data is highlighted in yellow. In the example provided, the analyzed process does not cover around 45 percent of the debit activity (43 percent of the credit activity) in trade payables and around 33 percent of the debit activity (89 percent of the credit activity) in inventories. To support investigating the

[557] Cf. Chapter 4.1.7.
[558] Cf. Chapter 3.2.1.
[559] For illustration purposes, in Figure 55 the "Account coverage" dashboard is filtered to two balance sheet account classes only.

root cause, the analysis can be adjusted to include the individual accounts of the account classes investigated (Figure 56).

GL account class	GL account description	Coverage	Debit	Credit	▼ Net activity (LC)
4-CL2999_TO TRADE PAYABLES	GR/IR clearing-others	Not covered	64,550,402 (99.98%)	58,407,655 (99.93%)	6,142,748
		Invoice - With PO	11,833 (0.02%)	41,833 (0.07%)	-30,000
	Trade accounts payable IC ABCC	Not covered	18,111,717 (94.14%)	14,665,285 (79.71%)	3,446,432
		Invoice - Without PO	31,813 (0.17%)	24,705 (0.13%)	7,108
		Invoice - With PO	1,096,522 (5.70%)	3,709,457 (20.16%)	-2,612,935
	Trade accounts payable-third Local	Invoice - Without PO	10,087,239 (12.88%)	8,953,701 (11.43%)	1,133,538
		Not covered	5,025,576 (6.42%)	4,904,808 (6.26%)	120,768
		Invoice - With PO	63,210,171 (80.70%)	64,473,485 (82.31%)	-1,263,315
	GR/IR clearing-Duty	Invoice - With PO	5,959,179 (99.49%)	5,957,045 (99.48%)	2,134
		Not covered	30,846 (0.51%)	30,846 (0.52%)	0
	GR/IR clearing account for trading goods	Not covered	11,381,186 (13.74%)	11,388,562 (12.80%)	-7,376
		Invoice - With PO	71,465,416 (86.26%)	77,600,788 (87.20%)	-6,135,372
	Trade a/cs - Third Manual	Not covered	26,830,771 (100.00%)	31,724,811 (100.00%)	-4,894,040
3-CA3999_TO INVENTORIES	Trading goods	Invoice - With PO	80,603,723 (92.75%)	13,563,648 (14.93%)	67,040,075
		Not covered	6,299,412 (7.25%)	77,261,066 (85.07%)	-70,961,654
	Specific provision	Not covered	1,177,412 (100.00%)	92,168 (100.00%)	1,085,244
	Goods in transit	Invoice - Without PO	7,547 (0.02%)	7,547 (0.02%)	0
		Not covered	32,057,802 (99.98%)	33,324,985 (99.98%)	-1,267,184
		Total	397,938,567	406,132,394	-8,193,827

Figure 56: Coverage of trade payables and inventory accounts by the process data

Lowering the level of aggregation shows that significant parts of trade payables that are not covered by the analyzed process involve, in particular, certain clearings made to the GR/IR account and intercompany transactions. Inventory related transactions that are not covered include provisions, goods in transit and credit entries made to trading goods.

Advantages of analyzing the account coverage of the process data

As a result of applying process mining in practice, all audit teams confirmed that integrating the information on the general ledger accounts' coverage by the process data significantly contributes to the understanding of the flows of transactions as required by ISA 315. As the auditor allocates the overall materiality to individual classes of transactions, account balances and disclosures,[560] the concept of materiality becomes applicable when assessing the account coverage of individual process flows. By this, process mining can support:

- assessing the significance of a class of transaction,
- deconstructing the flow of transactions into different significant classes of transactions that have not been previously identified,[561]

[560] Cf. Chapter 2.3.1.
[561] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A203.

- identifying significant transaction volumes that have been expected to relate to the analyzed process but are not covered by the extracted data and, at the same time,
- challenging the completeness of the event log for the specific entity.

As indicated by the relatively small coverage from cases without a purchase order in Figure 55, the entity does not seem to make extensive use of the SAP FI module. In this scenario, the auditor may decide to not identify direct purchases that do not relate to a purchase order as a significant class of transactions. On the other hand, using traditional audit procedures, some teams identified the purchase to pay process as a significant class of transactions but did not distinguish between purchases related to a purchase order and direct purchases. Based on the data-driven assessment of the individual transaction flows that are running through the SAP MM and FI modules, and the different risks of material misstatement related to these transactions, some teams decided to further deconstruct the significant class of transactions.

In the example provided in Figure 56, intercompany transactions might represent a significant flow of transactions that is not covered by the process at hand.[562] This demonstrates how linking the process data to the general ledger data helps to overcome some of the limitations identified with regard to the completeness of the event log.[563] Based on the degree of automation and the IT systems used to process these transactions, the auditor may either extend the event log to include those transactions or, alternatively, identify related party transactions as a separate class of transactions and design and execute further audit procedures to identify and address any related risks of material misstatement. At the same time, transaction volumes not part of purchase to pay (and thus inherently not covered by the process data) can be identified, like goods in transit that relate to the inventory cycle or credits to inventory that relate to the order to cash cycle.

[562] In the example provided, intercompany transactions are not covered by the process data as they are not processed using the SAP MM or SAP FI module.

[563] Cf. Chapter 4.1.7.

Challenges identified during piloting and suggested solution approaches

Challenges and acceptance problems identified with regard to the integration of general ledger data include:

(1) differences between the audit teams' process narrative and the analyzed process data regarding significant flows of transactions,
(2) inconsistencies in the coverage of accounts (accounts indicated as "covered" by the process mining analyzer were misleading for one specific entity) and
(3) inconsistent filtering when analyzing general ledger and process data in conjunction.

At (1): Differences between the process narrative and the data-driven process analysis

The auditor typically identifies the purchasing process and related accounts as significant class of transactions. Process mining may provide evidence that the process identified by the audit team using traditional inquiry and walkthrough procedures does not cover the entire purchasing process of the entity. For example, using the process data, additional initiating and processing activities different to those documented by the audit team may be identified. When combining the process mining data with general ledger data, the impact of the different flows of transactions on the general ledger accounts can be quantified. By this, the auditor may identify significant transaction volumes or account balances covered by the process that have not been previously identified. For example, some audit teams did not deconstruct the purchase to pay process into direct purchases and purchases related to a purchase order using traditional procedures, although process mining demonstrates that the critical path of these transactions and the bookkeeping practice differs significantly. Depending on the completeness of the event log,[564] there may also be instances where the process mining solution does not cover the entire purchasing process that has been identified and documented by the audit team in prior audit periods. In additional, manual parts of the purchasing process or transaction flows not covered by the event log need to be audited the traditional way despite using process mining. Audit teams criticize that although process mining significantly increases the process understanding, the described scenarios require additional work by the auditor, reducing audit efficiency.

[564] Cf. Chapter 4.1.7.

The described scenarios and the feedback obtained from the audit teams are challenges both research and audit practice will have to deal with.

The data-driven process analysis, together with the integration of process and financial data, provides a transparency that has not been there before. The increased quality of the auditor's process understanding comes with the downside of additional time and cost investment to identify and address risks of material misstatement of significant flows of transactions not previously identified – at least in the first period of application.

As process automation increases, manual activities will decrease. However, there will always be manual procedures performed by the entity that cannot be addressed with any automated tool or technique. These manual procedures may or may not be relevant to the audit. As such, process mining will likely not entirely "replace" the traditional means of analyzing processes, but support focusing these procedures to aspects of the process inherently not evident in the data.

On the other hand, the involvement of IT specialists will be required in improving the completeness of event logs in today's heterogenous IT landscapes.[565] Today's process mining solutions do not support analyzing the processing of significant transaction volumes not covered by the analyzed process. Currently, these process flows may only be analyzed by extending the event log and the definition of the process instance, or (depending on the interrelations between the processes) setting up a separate process mining analyzer. Alternatively, as will be demonstrated in the following two chapters, identifying and analyzing flows of transactions affecting trade payables related accounts can be supported by integrating journal entry posting (i.e., recording) events and the trade payables subledger data into the application.[566] If the transaction flows represent an individual significant class of transactions, the related initiating and processing activities along the critical path may be integrated into the event log. Otherwise, if these transactions are typical postings within the audited process, the related recording events may be incorporated as activities into the process graph and considered as "covered" in the account coverage analysis.

[565] Cf. Chapter 4.1.2.
[566] Cf. Chapters 4.2.2, 4.2.3 and 4.2.4.

At (2): Inconsistencies in the coverage of accounts

As part of the quality assurance procedures performed during piloting, the author of this thesis identified occasions where parts of account balances that are not related to purchase to pay were shown as covered by the process data in the analyzer. These scenarios may occur because the coverage of a transaction by the process data is determined using the header information of the respective documents.

Event table			
Case ID (PO line item)	Event time	Activitiy	Accounting Reference
00010000-00010	Jan 21	Create PO	
00010000-00010	Jan 21	Post INV	5100000590-2021-CH80
00020000-00010	Jan 21	Create PO	
00020000-00010	Jan 21	Post INV	5100000590-2021-CH80

Invoice header	
Document number	Fiscal year
5105625922	2021

Invoice line item					
Document number	Fiscal year	Line item	Amount	PO number	PO line item
5105625922	2021	1	450	00010000	00010
5105625922	2021	2	450	00020000	00010

Journal entry header					
Document number	Fiscal year	Company code	Posting date	Reference key	Coverage
5100000590	2021	CH80	31.01.2021	5105625922-2021	Yes

Journal entry line item							
Document number	Fiscal year	Company code	Line item	Account	Amount	PO number	PO line item
5100000590	2021	CH80	1	Trade accounts payable	-1.000		
5100000590	2021	CH80	2	Input VAT	100		
5100000590	2021	CH80	3	GR/IR clearing for trading goods	450	00010000	00010
5100000590	2021	CH80	4	GR/IR clearing for trading goods	450	00020000	00010

Figure 57: Determination of the account coverage on document header level

Figure 57 provides an example of a journal entry that is covered by the process data. Each event in the event log refers to the respective case (purchase order line item) it belongs to (shown in blue). The same case reference is included for each invoice and journal entry line item. The related journal entry header and the invoice header are linked by a unique reference as well (shown in red). The journal entry is considered as covered by the process data as at least one of the journal entry line items is referencing to a purchase order line item that is included in the extracted event data.

The determination of the coverage of the general ledger data by the process data on the journal entry header level is appropriate only if the relationships between the data sources can be established (i.e., no references are missing) and if there are no journal entries covering multiple business transactions.

A journal entry line item may either miss the reference to the related purchase order line item (because it has not been entered by the employee posting the journal entry) or refer to a purchase order line item that is not included in the event data (for example, because the related purchase order has been created in a period outside the data extraction period). The problem of missing references can be solved by reconciling both data sources.[567] Identifying the affected cases and journal entries will at the same time identify the business activity that is inappropriately classified as "covered" or "not covered" by the process data.

Even if all references to the process data can be established there may be journal entries covering multiple business transactions, as illustrated in Figure 58.

Journal entry number	Fiscal year	Company code	Line item	Account	Amount	PO number	PO line item
5100000590	2021	CH80	1	GR/IR clearing for trading goods	450	00010000	00010
5100000590	2021	CH80	2	GR/IR clearing for trading goods	450	00020000	00010
5100000590	2021	CH80	3	Input VAT	100		
5100000590	2021	CH80	4	Trade accounts payable	-1.000		
5100000590	2021	CH80	5	Trade receivables	14.280		
5100000590	2021	CH80	6	Sales	-12.000		
5100000590	2021	CH80	7	Output VAT	-2.280		

Figure 58: Journal entry covering multiple business transactions

In the example, the purchase and sale of different goods are combined in a single journal entry. As some line items of the journal entry are referring to a purchase order existing in the process data (and the coverage is determined on journal entry header level), the entire journal entry is considered as "covered" by the process data. That is, the account coverage analysis will display the related activity on the trade receivables, sales and output VAT accounts as "covered" by the purchase to pay process.

Journal entries covering multiple business transactions may further be ambiguous. A journal entry is ambiguous if it consists of more than two debit and more than two credit line items. In these scenarios, the transaction flow between the related accounts intended when entering the business transactions into the system cannot be determined anymore.[568]

As journal entries covering multiple business transactions can occur in practice, the analysis of the account coverage needs to be focused on the accounts that relate to the purchasing process. The auditor may leverage evidence on these journal entries to evaluate the entity's adherence to applicable GAAP.

567 A reconciliation between these data sources is discussed in Chapter 4.2.3.
568 Cf. MOCHTY, LUDWIG/WIESE, MICHAEL (2012), p. 484.

At (3): Inconsistent filtering across dashboards using general ledger and process data

Several observations made during the practical implementation of process mining relate to the behavior of global filters in the analyzer. For example, some audit teams tried to filter for a specific account in the trial balance to then switch to the process graph and investigate the related flow of transactions. However, the global filters set for the trial balance accounts were deleted when switching to the process graph. Similarly, when investigating process variations, teams were interested in identifying the general ledger accounts involved in a specific process trace, but the variation was removed from the filter when switching to the trial balance. The behavior of filters is a key element to consider when using process mining in the audit, as it directly impacts the audit procedures that can be performed while in parallel, inconsistent filtering may also go undetected by the auditor.

Due to the different nature of process and financial data, the integration of general ledger data is limited to specific parts of the application, depending on the data source of the specific analysis. As the ERP system uses a relational database,[569] both sources of information have to be linked manually, either in the data model or in the back-end of the process mining application. If a link is not established in both directions this directly impacts the analyses possible for the auditor. In the example provided, it is not possible to filter data in the trial balance and then switch to the process overview, as the trial balances is based on general ledger data and the process graph uses case and event information. As such, understanding the data model behind the process mining application is key in order to design appropriate audit procedures. When using relational databases, there is no solution approach to overcome these challenges despite the further integration of the data model.[570] An approach to further integrate the financial data and the process data is discussed in the following two chapters.

[569] Cf. Chapter 4.1.2.
[570] Cf. ibid.

4.2.2 Integration of journal entry recording events to complete the critical path

Identifying and understanding the critical path of the process

As part of understanding a significant class of transactions, the auditor obtains an understanding of its critical path, i.e., "*how transactions are initiated, and how information about them is recorded, processed, corrected as necessary, incorporated in the general ledger and reported in the financial statements*".[571] Process mining provides information about the initiation activities (for example, the creation of a purchase order), processing activities (for example, the approval of the purchase order or the entry of an invoice) and correction activities (for example, the price change of a purchase order). However, current event logs used in process mining do not inform about events that lead to the recording of a journal entry and, consequently, about the account classes affected in the financial statements.[572] Figure 59 illustrates the parts of the critical path that are covered by traditional event logs by adding the process perspective to the methodology layer introduced in Figure 11.

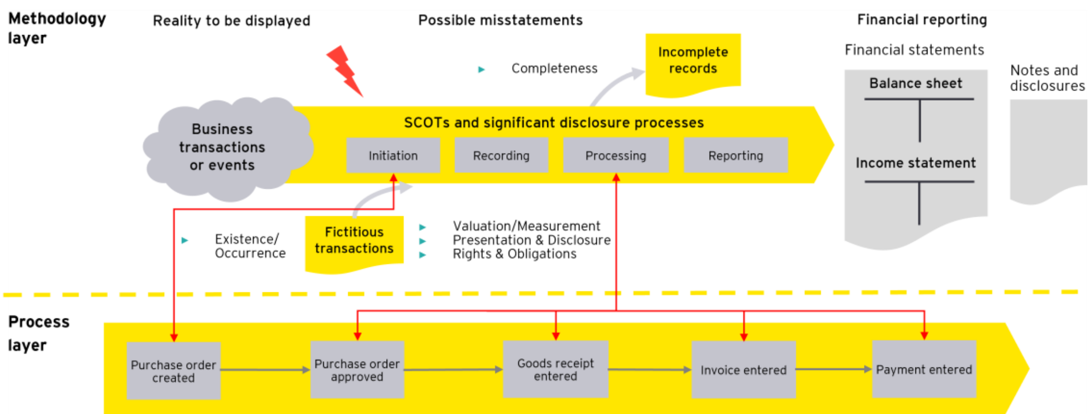


Figure 59: Analyzing elements of the critical path using an isolated process perspective

To benefit from the data-driven process analysis beyond removing the bias from inquiries and manual observations, the auditor needs to be able to obtain evidence about the recording in the general ledger and reporting in the financial statements as well.[573] For example, in purchase to pay, besides information about the entry of a goods re-

[571] IFAC (2021), ISA 315 (Revised 2019), para. 25(a).
[572] Cf. MOCHTY, LUDWIG (2015), p. 32.
[573] MOCHTY highlights that in the audit practice, the documentation of the process understanding often lacks a precise documentation of how the activities in the process lead to the recording of journal entries in the financial accounts, cf. ibid., p. 35.

ceipt, an invoice or a payment in the system, process mining should provide information about their respective recording in the general ledger, including the related financial accounts.

Covering the critical path of the process: Integration of recording events

With the integration of the account coverage analysis[574], the auditor may establish a link between the process data and the balance sheet and income statement accounts of the entity.[575] However, this link is established using the document information extracted from the ERP system. It is not evident from this data if a single case triggers a contiguous chain of debit and credit entries on the respective accounts in the purchase to pay process.[576]

To provide a complete picture of the critical path of the transactions, the posting events in the case audit firm's process mining analyzer introduced in Chapter 3 have been linked with the related journal entries. Figure 60 illustrates the analysis of the critical path of the purchase to pay process from the methodology layer, the process layer and the accounting layer.

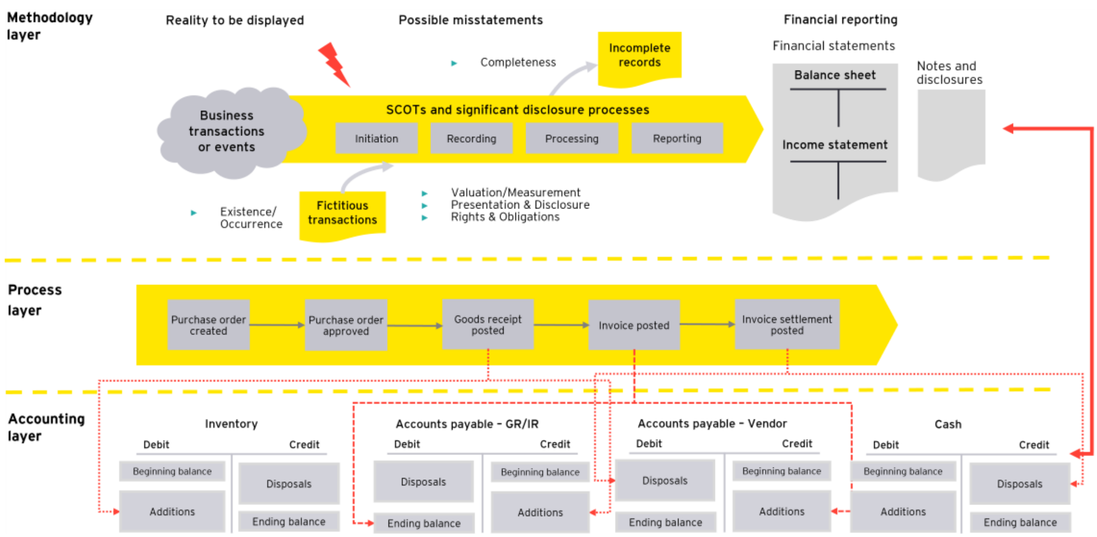


Figure 60: Analyzing the entire critical path of a process by integrating the accounting layer

In the example provided, the activity "goods receipt posted" is related to a journal entry debiting the inventory account class and crediting the GR/IR account in the trade payables account class. The journal entry related to the "invoice posted" activity is debiting the GR/IR account and crediting the respective vendor account in the trade

[574] Cf. Figure 55 and Figure 56.
[575] Cf. Chapter 4.2.1.
[576] Cf. MOCHTY, LUDWIG (2015), p. 32.

payables account class. Once the payment is made, the settlement of the invoice is reflected by a journal entry debiting the vendor account and crediting the bank account.

The analysis of recording activities and the related journal entries is implemented on the "Booking pattern" dashboard[577] introduced in Chapter 3.2.1. By analyzing the bookkeeping systematic behind the events in the event log, the auditor may understand the journal entries related to recording, reversal or clearing activities in the process and determine if the entity's booking behavior within the process complies with applicable GAAP.

Discussion of challenges and acceptance problems identified

In audit practice, the documentation of the business cycles and the activities performed along the critical path occasionally lacks a precise documentation of the related journal entries made to the financial accounts.[578] Accordingly, many audit teams confirm that the booking pattern re-constructed from the process and accounting data facilitates their understanding of the recording activities within the process. Some audit teams further indicate that the information supports deconstructing the significant class of transactions according to the different risks of material misstatement and control procedures related to the processing of transactions. For example, the three-way-match between quantity and/or price between the purchase order, goods receipt and invoice document in SAP is only performed for transactions where all three documents are present. As direct purchases or purchases of services do not include a posting of a goods receipt, additional control activities, such as the business approval of the related invoice, need to be identified to address related risks of material misstatement.

However, due to the limited connection between the financial and process related data sources that are inherent to relational databases,[579] the relationship between the activities displayed in the process graph and the posting of a journal entry is not always clear to the teams. The activities analyzed with process mining usually cover the process perspective, i.e., clicks and actions performed within a software related to the creation, approval or entry of specific documents. When combining these activities with journal entry recordings, specific consideration needs to be devoted to the interpretation of the analyses. For example, does the activity "invoice settled by payment

[577] Cf. Figure 29.
[578] Also see MOCHTY, LUDWIG (2015), p. 35.
[579] Cf. Chapter 4.2.1.

run" relate to the initiation of the payment run or the posting of the related journal entry? In the implementation described, the journal entries have been linked to their relevant activities based on the posting keys. However, the "posting" activities in the process are, in fact, describing the entry of related documents from a process perspective. Journal entry recording events are not integrated in the process graph, i.e., not considered as separate events in the event log. Consequently, they cannot be considered in any process data related analysis. Moreover, there are frequently significant differences between the date a journal entry is entered into the system and the date it becomes effective.[580]

The difficulties with regard to understanding the activities in the process can be addressed by explicitly differentiating between activities that are processing transactions (for example, "invoice entered") and activities that relate to the recording of a journal entry (for example, "invoice JE recorded"). Figure 61 illustrates the integration of the recording events into the process layer.

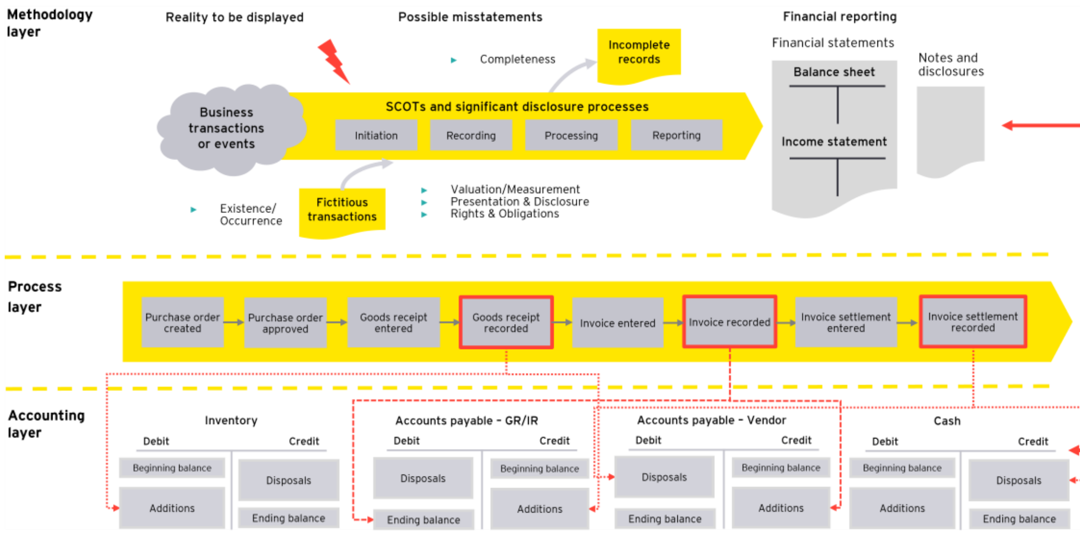


Figure 61: Differentiating between entry and recording activities

The "invoice posted" activity in the event log[581] is replaced by the two separate activities "invoice entered" and "invoice recorded". The "invoice entered" activity is referring to the entry date of the document while the "invoice recorded" activity is based on the effective date of the related journal entry. Similarly, any corrections made to the posting of invoices may be considered by implementing, for example, entry and posting activities for invoice reversals. By integrating the recording activities (i.e.,

[580] Cf. Chapter 4.1.5.
[581] Cf. Figure 60.

journal entry posting events) into the event log, the process data covers the entire critical path of the significant class of transactions. The recording activities can then be used to establish the connection to the account layer, i.e., their respective journal entries and the related general ledger accounts (represented by the booking pattern analysis[582]).

Despite the challenges identified, the approach to analyze an account's coverage by a specific process and to reconstruct the journal entries initiated by activities in the process demonstrates the potential of integrating process data with financial data. In essence, the audit risk model is relying on the double-entry bookkeeping and an incomplete identification of the actual bookkeeping systematic impacts the reliability of audit evidence obtained.[583] Currently, the auditor identifies classes of transactions and related general ledger accounts based on a typified and assumed accounting relationship in the trial balance.[584] For significant classes of transactions, these assumptions are supported by inquiries and confirmed by a walkthrough tracing an individual transaction through the process. However, it is not warranted that the booking systematic assumed (or tested by tracing one individual transaction) reflects the entity's actual bookkeeping practice. The allocation of an account to a class of transactions it does not belong to (or the missing allocation of a significant account to a significant class of transactions) may cause the auditor to obtain more (or less) audit evidence than required.[585] Integrating the information about the processing of transactions with the financial information of the general ledger accounts provides the opportunity of a data-driven approach to reliably confirm the classification of classes of transactions and related accounts.[586]

4.2.3 Integration of subledger data to track the financial amounts related to a case over time

Problem definition

An event in the event log does not contain information on its impact on the value of the process instance it relates to, for example, the invoice amount or purchase order amount. If considered in the analysis, these metrics are usually stored on the level of

[582] Cf. Figure 29.
[583] Cf. MOCHTY, LUDWIG/WIESE, MICHAEL (2012), p. 490.
[584] Cf. ibid., p. 489.
[585] Cf. ibid.
[586] Cf. Chapter 4.1.8.

the case and by this, they keep constant throughout the period.[587] Consequently, current process mining applications do not support to infer from a sequence of events performed to the value of the related process instance at a specific point in time.

This is significantly decreasing the application areas of process mining in the audit, as it has implications for both risk assessment procedures (including procedures to obtain an understanding of internal controls) and substantive procedures. For example, testing controls over the approval of purchase orders typically includes determining if the purchase order has been approved by an authorized person. Most organizations have approval hierarchies in place that assign the authorized individuals to different approval levels and define the level of approval required for specific amount intervals of the document. If the purchase order amount that was entered when creating the purchase order is stored on case level and thus kept constant throughout the period, possible change events related to the purchase order price before or after the approval of the purchase order are not considered when investigating the approval event.[588]

Similarly, many of the challenges identified when using process mining to support substantive audit procedures are caused by the incomplete integration of process and financial data. For example, if the invoice amount related to a case is only available on the level of the case, the total invoice amount related to a specific vendor cannot be identified at a specific point in time. This not only impedes the comparison of monthly invoicing activity but also the inference to the open liabilities of a specific vendor throughout the audit period.[589]

The examples demonstrate that storing the amounts related to a case as an event attribute in the event log and thus tracking their development as the case is processed can increase the relevance of process mining in the audit significantly.

Storing the invoice amount as event attribute

MOCHTY states that provided that (invoice) amounts are captured in the event log, it must be possible to present the monetary flow in the graph of a business process.[590] In the following, the invoice amounts are used as example to propose a concept for capturing the amounts related to an event in the event log.

[587] Cf. Table 1 and Table 2; JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), p. 17.
[588] This does not only affect testing the effectiveness of the approval control but also the evaluation of the magnitude of control deficiencies identified, cf. Chapter 2.3.4 and Chapter 4.5.2.
[589] Cf. Chapter 3.3.2.
[590] Cf. MOCHTY, LUDWIG (2015), p. 40.

A prerequisite for identifying the relevant amounts is understanding the object and table relationships of documents in the ERP system. Figure 62 illustrates the relationship between purchase order documents, invoice documents and journal entries (accounting documents) in SAP. The SAP tables containing the respective header and line item information are shown in brackets.

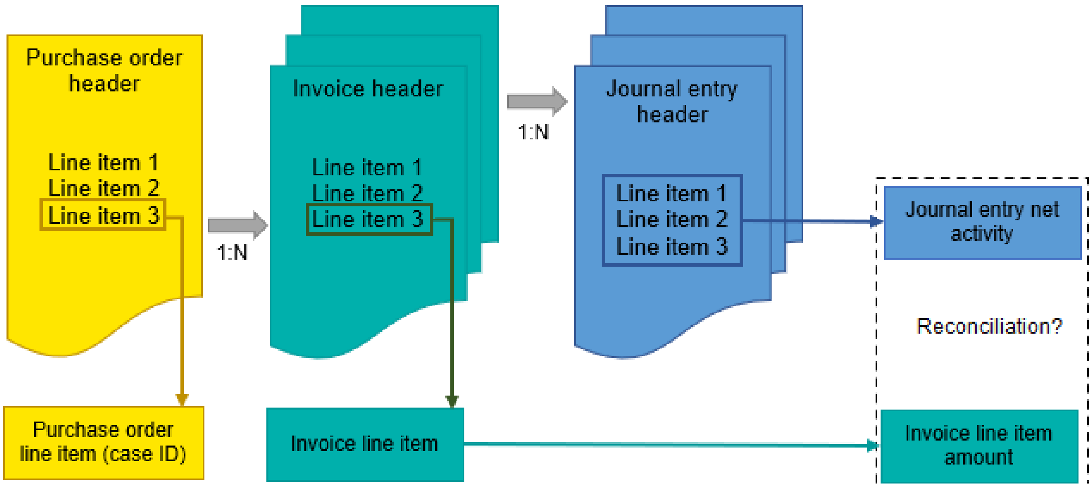


Figure 62: Object and table relationship of purchase orders, invoices and journal entries in SAP

A purchase order line item (that is used as the case ID) can relate to multiple invoice line items. Similarly, an invoice line item can be associated to multiple journal entries and a single journal entry can refer to multiple invoice line items.[591] That is, the exact relationship between the documents cannot be identified on header level.

As the case is defined as a purchase order line item, the invoice related events for a case are on line item level as well, that is, on the level of an individual material or service listed on the invoice. Usually, invoice line items are entered with a reference to the related purchase order line items. Consequently, an approach to overcome the complexity of possible relations of the data sources is storing the invoice line item amount on event level as well. However, the journal entry posting the invoice to the subledger may include other positions in addition to the individual line items (such as VAT or rebates applying to the entire invoice). Here, a reconciliation between the invoice amount of the process data with the net activity of the related journal entries can be performed. The reconciliation provides transparency on the data that can and cannot be reconciled and on any additional journal entries that have been made on invoice header level.

[591] Cf. Chapter 4.1.4.

Figure 63 illustrates the challenge in allocating the invoice amount related to a case on event level based on the purchase order, invoice and accounting document tables in SAP.[592]

Purchase order header (EKKO)

EBELN
450000

Invoice header (RBKP)

BELNR	GJAHR
10000	2020

Purchase order line items (EKPO)

EBELN	EBELP	AMOUNT
450000	10	300
450000	20	200

Invoice line items (RSEG)

BELNR	GJAHR	BUZEI	AMOUNT	EBELN	EBELP
10000	2020	1	300	450000	10
10000	2020	2	200	450000	20

Accounting document header (BKPF)

BELNR	GJAHR	BUKRS	AWKEY	AWTYP
500000	2020	CH00	10000-2020	RMRP

Accounting document line items (BSEG)

BELNR	GJAHR	BUKRS	BUZEI	AMOUNT	EBELN	EBELP
500000	2020	CH00	1	-550		
500000	2020	CH00	2	50		
500000	2020	CH00	3	300	450000	10
500000	2020	CH00	4	200	450000	20

Event table

Case	Activity	Amount
450000-10	Invoice entered	300
450000-10	Invoice posted	550
450000-20	Invoice entered	200
450000-20	Invoice posted	550

Case table

Case	Amount
450000-10	550
450000-20	550

Invoice amount: 1.100

Posted in accounts payable
Link between invoice and accounting document header
Link between purchase order, invoice and accounting document line items

Figure 63: Challenge in allocating the invoice amount on event level

The purchase order 450000 has the line items 10 and 20 with an amount of 300 EUR and 200 EUR, respectively. For the purchase order, a single invoice is received and entered into the system referencing the two purchase order line items (highlighted in yellow in Figure 63).[593] Accordingly, the event table includes the event "invoice entered" with an amount of 300 EUR referencing to the case 450000-10 (line item 10 of purchase order 450000) and a second event "invoice entered" with an amount of 200 EUR referencing to the case 450000-20. The related journal entry posting the invoice

[592] The goods receipt document is not considered in the analysis for simplification purposes. However, similar considerations apply when adding the goods receipt posting information to the process data.

[593] Price variances are not considered in this illustrative example. In SAP, the reference to the purchase order header and line items is performed using the EBELN and EBELP fields, respectively (highlighted in yellow in Figure 63).

is crediting the trade payables account (-550 EUR) and debiting VAT (50 EUR) and the GR/IR clearing account (200 + 300 = 500 EUR).[594]

However, the total invoice amount posted to the trade payables subledger also includes general invoice positions like VAT or rebates that cannot be related to an individual line item and, by this, are not part of the process data. That is, while the "invoice entered" event can be deconstructed to the individual line item, the "invoice posted" event representing the journal entry posted to the trade payables subledger (displayed in green) may relate to multiple line items and include general invoice positions. In addition, as illustrated at the bottom-right of Figure 63, aggregating the recorded invoice amount of a case with multiple "invoice posted" events will result in an incorrect invoice amount on case level.

A solution approach for summarizing the invoice amount on case level is provided in Figure 64.

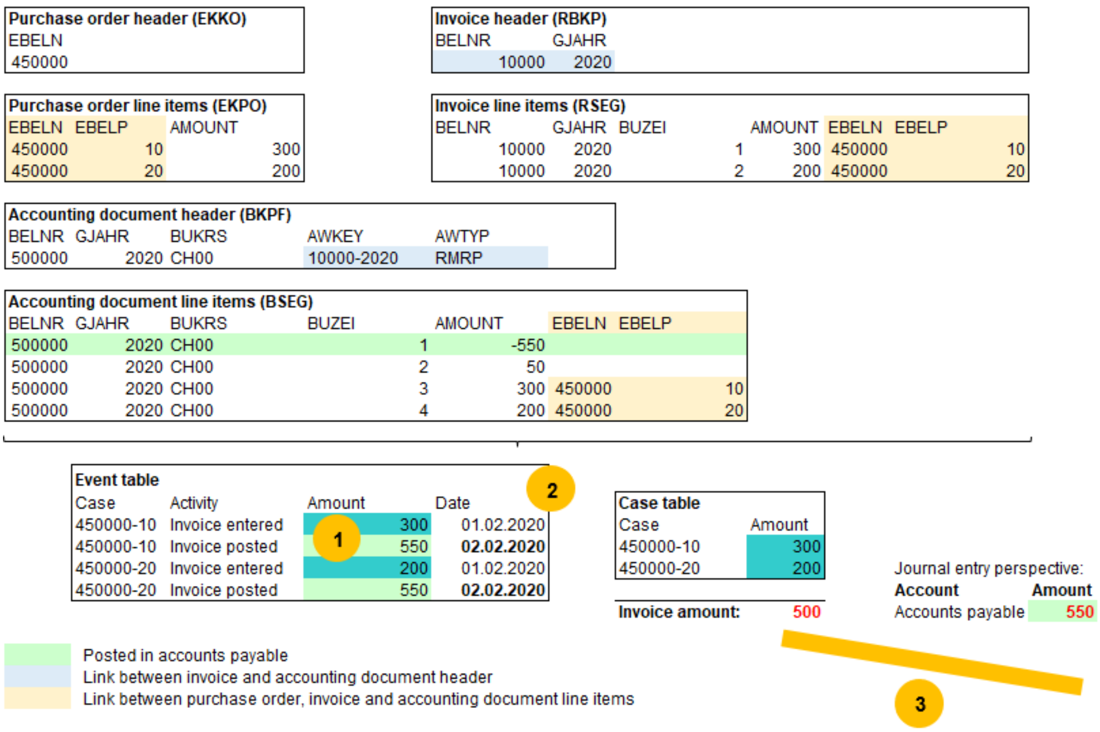


Figure 64: Solution for allocating the invoice amount on event level

[594] The relationship between the invoice and the accounting document in SAP is established on header level using the invoice document number and the fiscal year information (highlighted blue in Figure 63). The accounting document line items again reference to the respective purchase order line items.

As indicated in Figure 64, the allocation is performed in three steps:

(1) In the event table, the invoice amount related to the "invoice entered" event is used for calculating the case metric (highlighted in turquoise), as this amount corresponds to the invoice line item (and not the journal entry posted to the trade payables account).

(2) The effective date of the journal entry is considered for calculating the current invoice amount of a case at a given point in time.

(3) A reconciliation between the invoice amount as per the process data and the invoice amount posted to the trade payables subledger is performed.

The allocation of the invoice amount enables calculating the invoice amount related to a case at a specific point in time.

Event table

Case	Activity	Amount	Date	Accounting reference
450000-10	Invoice entered	300	01.02.2020	500000-2020-CH00
450000-10	Invoice posted	550	02.02.2020	500000-2020-CH00
450000-20	Invoice entered	200	01.02.2020	500000-2020-CH00
450000-20	Invoice posted	550	02.02.2020	500000-2020-CH00

Journal entry

Number	Fiscal year	Company code	Line item	Amount	PO Number	PO Line item
500000	2020	CH00	1	-550		
500000	2020	CH00	2	50		
500000	2020	CH00	3	300	450000	10
500000	2020	CH00	4	200	450000	20

Aggregation based on case ID, accounting reference and cumulated amount per case

Event table Case ID	Event table Accounting reference	Event table Amount	Journal entry Case ID	Journal entry Reference	Journal entry Amount	Control Case ID	Control Reference	Difference Amount
450000-10	500000-2020-CH00	300	450000-10	500000-2020-CH00	300	TRUE	TRUE	0
450000-20	500000-2020-CH00	500	450000-20	500000-2020-CH00	500	TRUE	TRUE	0

Comparison of event and journal entry data fields

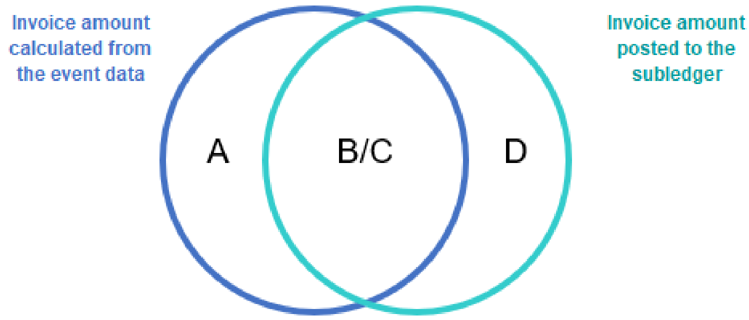
Figure 65: Tracking the invoice amount of a case over time

Figure 65 shows the same event table and journal entry used in the example provided in Figure 64. The information from both data sources can now be aggregated using the case ID and accounting reference. Cumulating the invoice amount of the individual events enables to determine the invoice amount of the case at a specific point in time. Verification is performed by comparing the individual data fields with the source tables.

By separating between processing activities (for example, "invoice entered") and recording activities (for example, "invoice recorded") and storing the invoice amount information on event level, the critical path of the process can be covered by process mining as required by ISA 315. However, there may be differences between the invoice amount as per the process data and the invoice amount posted to the subledger. As introduced in Figure 62 and demonstrated in Figure 64, when the process instance is a purchase order line item, the invoice amount for these cases is based on the individual invoice line items as well. In contrast, the journal entries in the financial data are using the total invoice amount. The total amount of an invoice does not only include the invoice line items but also general positions applying to the entire invoice, for example, VAT, shipping fees or rebates. Consequently, there is a discrepancy between the invoice amount according to the process and the invoice amount posted to the subledger. These differences require a reconciliation between the invoice line item amount and the invoice amount posted to the trade payables subledger.

Reconciliation of the invoice amount between the process and the financial data

When reconciling the invoice amount calculated from the process data to the invoice amount posted to the subledger, there are four different categories available.



A The purchase order reference is missing on the journal entry line item level.
B The invoice amount of the process is reconcilable to the invoice amount posted to the subledger.
C The invoice amount of the process is not reconcilable to the invoice amount posted to the subledger.
D The purchase order referenced in the journal entry is not available in the event data.

Figure 66: Categories available in the invoice amount reconciliation

Category A describes scenarios where the purchase order line item, that is used as the case in the process data, is not referenced in the respective journal entry line item of the invoice posting (Figure 67).

Event table			
Case ID (PO line item)	Event time	Activitiy	Accounting Reference
00010000-00010	Jan 21	Create PO	
00010000-00010	Jan 21	Post INV	5100000590-2021-CH80
00020000-00010	Jan 21	Create PO	
00020000-00010	Jan 21	Post INV	5100000590-2021-CH80

Invoice header	
Document number	Fiscal year
5105625922	2021

Invoice line item					
Document number	Fiscal year	Line item	Amount	PO number	PO line item
5105625922	2021	1	450	00010000	00010
5105625922	2021	2	450	00020000	00010

Journal entry header				
Document number	Fiscal year	Company code	Posting date	Reference key
5100000590	2021	CH80	31.01.2021	5105625922-2021

Journal entry line item							
Document number	Fiscal year	Company code	Line item	Account	Amount	PO number	PO line item
5100000590	2021	CH80	1	Trade accounts payable	-1.000		
5100000590	2021	CH80	2	Input VAT	100		
5100000590	2021	CH80	3	GR/IR clearing for trading goods	450	No PO reference	
5100000590	2021	CH80	4	GR/IR clearing for trading goods	450	No PO reference	

Figure 67: Missing purchase order reference on journal entry line item level

This category can occur when the entity did not maintain the purchase order reference in the journal entry line item. Due to the missing reference, the case in the event log cannot be linked to the related journal entry.

For the cases included in category B, the invoice amount of the process reconciles to the invoice amount posted to the subledger (as illustrated in Figure 65).

In category C, the invoice amount of the event data does not reconcile to the invoice amount calculated from the journal entries (Figure 68).

Invoice line item					
Document number	Fiscal year	Line item	Amount	PO number	PO line item
5105625922	2021	1	49.000	00010000	00010

Journal entry line item								
Document number	Fiscal year	Line item	Posting key	Account	Account type	Amount	PO number	PO line item
5100000590	2021	1	31 - Invoice	Trade accounts payable	Liabilities	-52.920		
5100000590	2021	2	86 - GR/IR debit	GR/IR clearing for trading goods	Liabilities	49.000	00010000	00010
5100000590	2021	3	86 - GR/IR debit	Price variances - GR/IR trading	Expenses	3.920	00010000	00010

Figure 68: Reconciliation difference between the process and the subledger data

In Figure 68, the references can be identified in both data sources, however, the amount identified in the journal entry differs from the amount in the process data. In the example provided, the difference is caused by a price variance that is booked with a reference to the related purchase order line item. As price variances are not represented by an activity in the process data, there is a non-reconciling difference of 3.920 EUR between both data sources.

Category D describes scenarios where the purchase order is not available in the event data extracted from the system (Figure 69).

Journal entry header				
Document number	Fiscal year	Company code	Posting date	Reference key
5100000590	2021	CH80	31.01.2021	5105625922-2021

Journal entry line item							
Document number	Fiscal year	Company code	Line item	Account	Amount	PO number	PO line item
5100000590	2021	CH80	1	Trade accounts payable	-1.000		
5100000590	2021	CH80	2	Input VAT	100		
5100000590	2021	CH80	3	GR/IR clearing	450	00010000	00010
5100000590	2021	CH80	4	GR/IR clearing	450	00020000	00010

Event table				
Case ID (PO line item)	Event time	Activitiy	Accounting Reference	Available in application
00010000-00010	Jan 21	Create PO		Yes
00010000-00010	Jan 21	Post INV	5100000590-2021-CH80	Yes
00020000-00010	Jan 15	Create PO		No
00020000-00010	Jan 21	Post INV	5100000590-2021-CH80	No

Figure 69: Missing purchase order in the event data

Based on the data extraction strategy,[595] these scenarios may occur if there are very old purchase orders (that have been created more than three years before the audit period) but there are still journal entry postings relating to them. Typically, these purchase orders are long-standing and in use for several years, for example, when the rent of a building is entered and maintained through a purchase order in the ERP system.

Categories A and D summarize the cases (purchase order line items) and journal entries where the link between both data sources cannot be established due to missing references. As highlighted in Chapter 4.2.1, the related journal entry might still be considered as "covered" by the process data. These scenarios may occur as the coverage of the general ledger data by the process data is determined on journal entry header level. With the reconciliation of the data sources, the coverage of the financial data by the process data can be determined accurately. Identifying the cases and journal entries in the reconciliation categories A and D and investigating the bookkeeping systematic will at the same time identify the business activity that is inappropriately classified as "covered" or "not covered" by the process data.

After reconciling the invoice line items between both data sources, the additional positions applying to the overall invoice that are included in the subledger postings (such as VAT, rebates or packaging costs) but not in the process data can be identified.

[595] Cf. Chapter 4.1.6.

Figure 70 summarizes the concept of the invoice amount reconciliation using a contiguous example.

ABC Vendor | **XYZ AG**

Invoice 00100 | **6 April 2021**

Line item	Material	PO number	Quantity	Price	Amount
10	Material A	1000-10	10	10	100
20	Material B	1000-20	5	10	50
30	Material C	1000-30	20	10	200
40	Material D	1000-40	1	10	10
50	Material E	2000-10	4	5	20

Net amount	**380**
Rebate 10%	-38
VAT 7,7%	29
Total including VAT	**371**

Company code: **XYZ**
Fiscal year: **2021**

Journal entry 00200

GL Account	Line item	Posting key	Amount	PO number	Vendor
200010 GR/IR clearing account	1	86	100	1000-10	
200010 GR/IR clearing account	2	86	50	1000-20	
200010 GR/IR clearing account	3	86	150	1000-30	
400010 Purchase price differences	4	83	50	missing	
200010 GR/IR clearing account	5	86	10	missing	
200010 GR/IR clearing account	6	86	20	2000-10	
120000 Trading goods	7	99	-38		
240000 VAT	8	40	29		
200000 Trade payables third parties	9	31	-371		ABC Company

Invoice amount reconciliation table

PO number	JE number	Invoice amount	JE net activity	Difference	Reconciliation category
1000-10	00200	100	100	0	B - Reconcilable
1000-20	00200	50	50	0	B - Reconcilable
1000-30	00200	200	150	50	C - Not reconcilable
1000-40	00200	10	0 (missing)	10	A - Missing PO reference on JE line item
2000-10	00200	0 (missing)	20	-20	D - Missing PO in event data

Invoice amount with PO reference	**360**	
Journal entry activity with PO reference		**320**
GR/IR		10
Purchase price difference		50
Rebates		-38
VAT		29
Invoice amount posted to the subledger		-371

Figure 70: The invoice amount reconciliation concept

On the top of the figure, an invoice with five line items, an overall rebate of ten percent as well as VAT of 7,7 percent is displayed. The journal entry in the middle shows the posting of the invoice in the bookkeeping system. For invoice line item 30, there is a purchase price difference of 50 EUR. The table at the bottom illustrates the reconciliation of the invoice amounts between both data sources. The reconciliation is designed in two steps.

First, the invoice line items are compared to the journal entry line items by linking them through their purchase order line item reference. When combined in the invoice amount reconciliation table, the invoice amount of the process data, the corresponding journal entry net amounts as well as any differences are shown in separate columns. If the individual positions match, they are labeled as "B - Reconcilable". For the purchase order line items "1000-10" and "1000-20" in the example provided, the invoice and journal entry line items are reconciled without difference and categorized as "B - Reconcilable". If the reference to the purchase order can be identified in both the invoice line items and the journal entry line items but there is a difference between both data sources, they are categorized as "C - Not reconcilable" (purchase order line item "1000-30"). For invoice line items without a corresponding purchase order reference in the journal entry line items, the label "A - Missing PO reference on JE line item" is used (purchase order line item "PO 1000-40"). The invoice line item 50 is associated to a purchase order that is not available in the event data (for example, because it has been created outside of the data extraction period). Consequently, the connection with the corresponding journal entry line item 6 cannot be established and the related invoice amount is categorized as "D - Missing PO in event data".

In a second step, the journal entry line items that do not have a purchase order reference are summarized. This includes the journal entry line items with a missing purchase order reference as well as general invoice positions not related to a specific invoice line item.

The total journal entry activity with a purchase order reference (320 EUR, identified in the first step) together with the positions identified in the second step sum up to the total invoice value being posted to the trade payables subledger and vendor accounts (371 EUR). Overall, there is a difference of 11 EUR between the invoice amount in the process mining data (360 EUR) and the invoice amount posted to the accounts payable subledger (371 EUR). This difference can be explained with:

- the general invoice positions that are not line item related, namely the overall rebate (-38 EUR) plus VAT (29 EUR) resulting in -9 EUR, plus

- the difference of 20 EUR resulting from reconciliation category D due to the missing purchase order in the event data.

Implementation of the invoice amount reconciliation

In the case audit firm's process mining analyzer, the concept of the invoice amount reconciliation may be integrated as a new dashboard on the "Financial reconciliation" menu page.

Elements include:

- the invoice amount reconciliation table, summarizing the results of the reconciliation using the categories[596] A, B, C and D,
- the reconciliation of the total invoice amount, including the general invoice positions, to the trade payables subledger and
- means to identify and investigate, if necessary, the differences resulting from the non-reconciling invoice line items (category C).

The invoice amount reconciliation table is presented in Figure 71.

Reconciliation category	Process data (INV amount CY)	in %	Journal entries (INV amount CY)	in %	Difference
Invoice amount with PO reference	**69,303,922**	**100%**			
JE/PO reference pair not available in journalentry data	0	0%	0	0%	0
Reconcilable	58,032,151	84%	58,032,151	84%	0
Not reconcilable	11,271,771	16%	11,399,877	16%	128,105
JE/PO reference pair not available in event data	0	0%	3,711	0%	3,711
Journal entry activity with PO reference			**69,435,738**	**100%**	**131,816**

Figure 71: Reconciliation between the invoice and the journal entry line items

The table summarizes the results of the reconciliation for the individual documents (as introduced in Figure 70) over the total population of cases and transactions. The first line of the table ("Invoice amount with PO reference") reconciles with the total invoice amount of the cases included in the analyzer. The last line of the table ("Journal entry activity with PO reference") summarizes the invoice amount of all journal entry line items that are referencing to a purchase order. The four lines in between represent the reconciliation categories A, B, C and D.

Figure 72 shows the reconciliation of the total invoice amount, including the general invoice positions, to the trade payables subledger.

[596] Cf. Figure 66.

GL account type	GL account class	GL account number	GL account description	▼ Net activity
Journal entry activity with PO reference				**69,435,738**
Assets	1300-Other Receivables	0011086000	Input Vat	2,410,149
Expenses	4000-Gross Margin I-Total (Exp)	0055001840	Freight And Insurance (Inbound)	52,719
	4200-Logistics & Distribution	0066105000	Packaging	34,590
		0066100100	Transport-Third Party	1,719
	4100-Selling Expenses	0066007850	Other Selling Expenses	515
	6000-Administrative Expenses	0066316440	Meals (Local)	129
Liabilities	2400-Accrued Expenses / Prepaid Income	0022027500	Accrued Expenses-Freight	33,470
	2100-Other Payables	0022014000	Output Vat	-22,395
Subtotal				**2,510,896**
Calculated invoice amount				**71,946,634**
Actual invoice amount posted to the subledger				**-71,946,634**
Difference				**0**

Figure 72: Reconciliation of the total invoice amount to the subledger

The first line of the table ("Journal entry activity with PO reference") reconciles to the last line of Figure 71 and, again, summarizes the invoice amount of all journal entry line items that are referencing to a purchase order. Subsequently, the table displays the net transaction volume posted to the general ledger accounts that is related to the general invoice positions, such as VAT, freight, insurance and packaging cost. The resulting subtotal is then reconciled to the total invoice amount posted to the subledger to ensure completeness of the analysis.

Figure 73 provides details on the reconciliation differences identified.

Reconciliation to Accounts Payable | Details

Local currency	Posting date	Journal entry number	Company code	Case ID	Reconciliation category	Total: CHF 69,303,922 Process data (INV amount CY)	Total: CHF 69,435,738 Journal entries (INV amount CY)	Total: CHF 131,816 ▼ Difference
CHF	2018-09-07	20185100004819CH80	CH80	300821903006700010	Not reconcilable	1,848,110.00	1,868,579.46	20,469.46
CHF	2018-03-20	20185100001480CH80	CH80	300821903006400010	Not reconcilable	156,575.44	163,594.72	7,019.28
CHF	2018-10-24	20185100005722CH80	CH80	300821903060400040	Not reconcilable	49,000.00	52,920.00	3,920.00
CHF	2018-01-18	20185100000291CH80	CH80	300821903125000010	Not reconcilable	35,430.54	38,609.22	3,178.68
CHF	2018-04-06	20185100001854CH80	CH80	300821902998700010	Not reconcilable	391,295.20	394,321.80	3,026.60
CHF	2018-08-16	20185100004347CH80	CH80	300821903099000010	Not reconcilable	452,992.02	455,992.02	3,000.00
CHF	2018-07-27	20185100004057CH80	CH80	300821903099000010	Not reconcilable	452,992.02	455,992.02	3,000.00
CHF	2018-02-22	20185100000979CH80	CH80	300821903147500010	Not reconcilable	47,416.27	49,147.33	1,731.06
CHF	2018-12-21	20185100006807CH80	CH80	300821903388500010	Not reconcilable	43,854.50	45,343.91	1,489.41
CHF	2018-06-27	20185100003504CH80	CH80	300821903178500010	Not reconcilable	39,965.66	41,164.63	1,198.97
CHF	2018-06-25	20185100003439CH80	CH80	300821903255700020	Not reconcilable	1,385.82	2,385.94	1,000.12
CHF	2018-08-15	20185100004315CH80	CH80	300721000936600020	Not reconcilable	25,027.72	26,008.50	980.78

Figure 73: Summary of identified reconciliation differences

The table lists the journal entries that cannot be reconciled to the process data, including the related invoice amounts and the reconciliation difference. In contrast to the aggregated invoice amount reconciliation table, debit and credit differences are shown separately. The total difference reconciles to the "difference" column in Figure 71.

If there are differences that are material individually or when aggregated with other differences, the auditor may drill-down to the reconciliation details for further investigation (Figure 74).

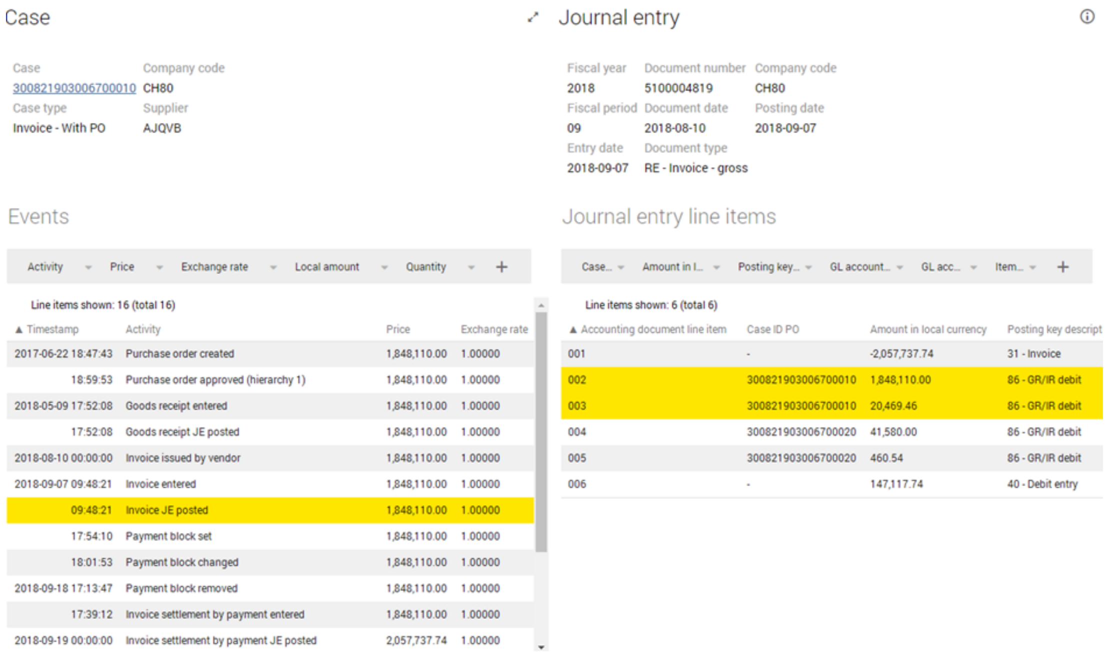


Figure 74: Reconciliation details for root-cause investigation for non-reconciling items

The reconciliation details include the case and event information extracted from the process data at the left side and the journal entry information at the right side. Differences identified between both data sources are highlighted in yellow and support identifying the root-cause of mismatches identified.

Empirical validation of the invoice amount reconciliation

In the following, the proposed approach of reconciling the invoice amount of the process data to the invoice amount posted to the subledger is tested on different real-life datasets. Figure 75 summarizes the testing results.

	Process data (invoice amount)	in %	Journal entries (invoice amount)	in %	Difference
Invoice amount of the process data	**69.303.922**	**100,00%**			
A: PO reference missing on JE line item level	0	0,00%	0	0,00%	0
B: Reconcilable	58.032.151	83,74%	58.032.151	83,58%	0
C: Not reconcilable	11.271.771	16,26%	11.399.877	16,42%	128.106
D: PO referenced in JE is not available in event data	0	0,00%	3.711	0,01%	3.711
Journal entry amount posted to the subledger			**69.435.739**	**100,00%**	**131.817**

Invoice amount of the process data

		Reconciliation category			
Entity	**Year**	**A**	**B**	**C**	**Total**
#1	2018	0,00%	96,53%	3,47%	**100,00%**
	2019	0,00%	94,49%	5,51%	**100,00%**
#2	2018	0,00%	97,12%	2,88%	**100,00%**
	2019	0,00%	97,00%	3,00%	**100,00%**
#3	2017	0,00%	88,91%	11,09%	**100,00%**
	2018	0,00%	88,65%	11,35%	**100,00%**
	2019	0,00%	83,74%	16,26%	**100,00%**
#4	2018	0,00%	1,91%	98,09%	**100,00%**
	2019	0,00%	1,59%	98,41%	**100,00%**
#5	2020	0,00%	96,16%	3,84%	**100,00%**
#6	2017	0,00%	85,05%	14,95%	**100,00%**
	2018	0,02%	87,82%	12,16%	**100,00%**
	2019	0,07%	88,27%	11,66%	**100,00%**
	Average	**0,01%**	**77,48%**	**22,51%**	

Journal entry amount posted to the subledger

		Reconciliation category			
Entity	**Year**	**B**	**C**	**D**	**Total**
#1	2018	96,09%	3,37%	0,54%	**100,00%**
	2019	94,54%	5,46%	0,00%	**100,00%**
#2	2018	96,69%	3,18%	0,13%	**100,00%**
	2019	96,62%	3,38%	0,00%	**100,00%**
#3	2017	88,67%	11,23%	0,10%	**100,00%**
	2018	88,47%	11,48%	0,05%	**100,00%**
	2019	83,58%	16,42%	0,01%	**100,00%**
#4	2018	1,81%	98,19%	0,00%	**100,00%**
	2019	1,51%	98,49%	0,00%	**100,00%**
#5	2020	95,61%	3,76%	0,63%	**100,00%**
#6	2017	87,74%	12,25%	0,01%	**100,00%**
	2018	85,49%	14,51%	0,00%	**100,00%**
	2019	86,22%	13,78%	0,00%	**100,00%**
	Average	**77,16%**	**22,73%**	**0,11%**	

Figure 75: Testing results of applying the invoice amount reconciliation to different datasets

The evaluation includes data from six different entities and covers 13 audit periods in total. For each year, the full period of twelve months is included in the data. The reconciliation overview at the top of the figure uses the data from entity #3 to illustrate how the different elements of the reconciliation table (Figure 71) are considered in the evaluation. The tables at the bottom summarize the reconciliation results for the process data (colored blue) and the journal entry data (colored turquoise), respectively.[597]

The validation results indicate that missing references between both data sources are only occurring rarely in practice. On average, only 0,01 percent of the process' total invoice volume cannot be reconciled to the journal entry data due to a missing reference in the journal entry line items (category A). Missing references in the process data due to the definition of the data extraction period[598] represent, on average, 0,11 percent of the total invoice amount posted to the subledger (category D). The related invoice amounts are not material for any of the entities, indicating that the impact of missing references between the data sources can be expected to be limited in practice.

[597] As reconciliation category D summarizes information not included in the process data, only categories A, B and C are included in the results for the process data. Vice versa, category A contains information not included in the journal entry data and is not included in the related table at the right side of Figure 75.

[598] Cf. Chapter 4.1.6.

For all but one entity, more than 85 percent of the total invoice amount is reconcilable between both data sources (category B). Non-reconciling differences (category D) are most frequently caused by price variances between the goods receipt posting on the GR/IR account and the posting of the invoice.[599] The auditor usually investigates these variances as part of the three-way-match analysis between purchase orders, goods receipts and invoices.[600] The deferral between the invoice amount that can and cannot be reconciled (categories B and D) for entity #4 were caused by an error in loading the data into the analyzer. As such, the reconciliation of the invoice amount between the data sources may be performed as additional procedure in the ETL process, validating the completeness and accuracy of the data. The auditor may be provided with a comprehensive report on non-reconciling items and their root-causes.

The proposed concept of tracking the invoice amount of a case over time enables to fully reconcile the invoice amount between the process and the financial data. The auditor can infer from a sequence of events performed to the value of the related process instance at a specific point in time, significantly increasing the application areas of process mining in the audit. Applications areas include, for example:

- a significant improvement of vendor related analyses (Chapter 4.2.4), supporting alternative procedures for (parts of) vendor accounts not covered by the process data without the need to use additional automated tools or techniques,
- the determination of the status of a case and analytical procedures on the completeness of liabilities (Chapter 4.2.5),
- roll-forward and comparison functionalities, requiring appropriate cut-off to determine if amounts are effective in one period or another (Chapter 4.4),
- evaluating internal controls (Chapter 4.5), including:
    - analyzing control procedures depending on specific amount thresholds such as invoice approval activities,
    - evaluating whether an additional approval is required in case the price has been adjusted and

599 Cf. Figure 68.
600 Cf. Chapter 3.2.1.

- a more detailed investigation of the root-cause and follow-up activities of exceptions from the three-way-match between purchase order, goods receipt and invoice.

4.2.4 Analysis of the creditor structure and reconstruction of open items

Challenges and acceptance problems identified

Various feedback obtained from the audit teams relates to the analysis of the vendor structure and related information.[601] The challenges identified may be summarized as follows:[602]

(1) The "open cases" in the process mining analyzer do not reconcile to the open item list prepared by the entity that displays the individual positions of liabilities by vendor at period end.

(2) The total balance by vendor displayed in the process mining analyzer does not reconcile to the financial vendor balances at period end.

(3) If parts of the vendor accounts related to trade payables are not covered by the process mining analyzer, they need to be addressed by audit procedures outside the analyzer, for example, using automated tools and techniques to analyze the subledger data. However, it is not clear to the audit teams how to identify these transaction flows in an analyzer that does not include the process data.

(4) Analytical procedures with regard to the movement of the invoice amount throughout the period, for example, the comparison of the invoice amounts processed in individual months, are not supported by the process mining analyzer.

(5) Quality assurance procedures showed that in designing the analyses with process mining, specific consideration needs to be devoted to the parties involved in a case. In the purchase to pay process, the vendor from whom a good or service is purchased, the vendor providing the goods or the service, the vendor sending the invoice and the vendor to whom the payment is made can differ.

[601] The implementation of the vendor-related analyses is described in Chapter 3.2.1.
[602] For more details refer to the feedback evaluation provided in Chapter 3.3.2.

Implementation of vendor-related analyses based on process and financial data

In Chapter 4.2.3, the concept of storing the invoice amount related to a case on event level together with the invoice amount reconciliation to the subledger have been introduced. By integrating the subledger data into the process mining application and linking it with the event and case data, the vendor related data can be reconstructed accurately. The proposed approach of analyzing vendor information based on the jointly consideration of process and financial data includes:

- reconstructing vendor balances and open items at a selected point in time and reconciling the results to the trial balance and open item list prepared by the entity (addressing challenges (1) and (2)),
- differentiating between transaction volumes related to trade payables that are covered by the process data and additional postings made to the subledger without a reference to a process instance (addressing challenge (3)),
- accurately displaying both the vendor balances at period end and the debit, credit or net transaction movements on a monthly basis, supporting analytical procedures (addressing challenge (4)) and
- identifying and appropriately reporting cases where the vendor varies between different events (addressing challenge (5)).

Linking information on open items with the process data

Presuming the subledger data has been added to the process mining analyzer and the invoice amount is stored on the level of the individual events as proposed in Chapter 4.2.3, the auditor can identify cases that represent open liabilities at a specific point in time. Table 12 contains an excerpt of a typical list of open liabilities by vendor.[603]

[603] The information included in the table can be exported from the ERP system. The age of the open liability as of period end is calculated using the balance sheet date and either the document date or posting date, depending on the purpose of the procedure. If an open item is already paid within the data extraction period, the clearing date can be added to the analysis to support auditing events and conditions subsequent to the balance sheet date.

Vendor name	Accounting document	Posting date	Clearing date	Document date	Document type	Posting key	Age (days)	Amount
SSGV	5100006895	28.12.2018	15.02.2019	06.12.2018	RE - Invoice - gross	31 - Invoice	25	-188.894,54
ABC	1900002636	21.12.2018	03.01.2019	13.12.2018	KR - Vendor invoice	31 - Invoice	18	-65.558,00
AAIHP	0700000071	30.11.2018	25.01.2019	30.11.2018	KD - Vendor debit memo	31 - Invoice	31	-64.698,01
AAIHP	0700000076	19.12.2018	22.02.2019	31.12.2018	KD - Vendor debit memo	31 - Invoice	0	-21.765,88
AADTH	1900002709	28.12.2018	03.01.2019	14.12.2018	KR - Vendor invoice	31 - Invoice	17	-7.746,20
AAZAI	5100006637	13.12.2018	03.01.2019	30.11.2018	RE - Invoice - gross	31 - Invoice	31	-644,90
AAZAI	1900002588	13.12.2018	03.01.2019	30.11.2018	KR - Vendor invoice	31 - Invoice	31	-500,35
AAZAI	1900002703	28.12.2018	24.01.2019	19.12.2018	KR - Vendor invoice	31 - Invoice	12	-500,35
AAAZB	5100006905	28.12.2018	10.01.2019	21.12.2018	RE - Invoice - gross	31 - Invoice	10	-558,37
AAGAV	5100006958	31.12.2018	17.01.2019	17.12.2018	RE - Invoice - gross	31 - Invoice	14	-556,25

Table 12: Open item list for trade payables at period end

The individual open items are linked to the related journal entries by the accounting document number. Table 13 shows the journal entry header and line item information for the accounting document "5100006895" related to vendor "SSGV".

Journal entry header				
Document number	Fiscal year	Document date	Posting date	Entry date
5100006895	2018	06.12.2018	28.12.2018	28.12.2018

Journal entry line item						
Document number	Line item	Posting key	Account	Account type	Amount	Vendor
5100006895	1	31 - Invoice	Trade accounts payable	Liabilities	-188.895	SSGV
5100006895	2	86 - GR/IR debit	GR/IR clearing for trading goods	Liabilities	188.515	-
5100006895	3	83 - Price difference	Price variances - GR/IR trading	Expenses	379	-

Table 13: Journal entry details related to the open item of vendor „SSGV"

As the journal entry data can be linked and reconciled to the process data, the related cases (i.e., purchase order line items) can be identified. In the example provided, the posting from trade accounts payable to the GR/IR clearing account is covered by the "invoice JE posted" event. Figure 76 shows the related booking pattern for the selected journal entry.

Event name	Debit/credit indicator	GL account class	Local currency	Posting key description	Net activity
Invoice JE posted	Credit	2000-Trade Payables	CHF	31 - Invoice	-188,895
	Debit	2000-Trade Payables	CHF	86 - GR/IR debit	188,515

Figure 76: Booking pattern related to the invoice posting activity for the selected journal entry

The purchase price difference booked to the expense account is not covered by the process data and identified by the invoice amount reconciliation introduced in Chapter 4.2.3.

While the vendor related open items cannot be identified based on whether a case includes a payment activity or not, the example demonstrates that they can be reconstructed by combining the process information with the subledger data. Without relying on a predefined "case status" filter that may not be suitable for the process or dataset at hand,[604] the auditor can identify process instances that result in open liabilities, be it at period end or at another point in time throughout the period.

Table 14 shows the table of open items implemented in the process mining analyzer.

Vendor name	Accounting document number	Age	Posting date	Clearing date	Document date	Document type description	Posting key description	Amount in local currency
AAAZB	5100006905	3	2018-12-28	2019-01-10	2018-12_21	RE - Invoice - gross	31 - Invoice	-558.37
AADTH	1900002709	3	2018-12-28	2019-01-03	2018-12_14	KR - Vendor invoice	31 - Invoice	-7,746.20
AAGAV	5100006958	0	2018-12-31	2019-01-17	2018-12_17	RE - Invoice - gross	31 - Invoice	-556.25
AAHZF	1900002065	87	2018-10-05	2019-04-11	2018-10_03	KR - Vendor invoice	29 - Down Payment Paid	98,248.45
AAIHP	0700000071	31	2018-11-30	2019-01-25	2018-11_30	KD - Vendor debit memo	31 - Invoice	-64,698.01
	0700000076	12	2018-12-19	2019-02-22	2018-12_31	KD - Vendor debit memo	31 - Invoice	-21,765.88
AAQOC	0100028585	0	2018-12-31	2019-01-01	2018-12_31	SA - G/L account document	31 - Invoice	-77.40
	1700000133	70	2018-10-22	-	2018-10_03	KG - Vendor credit memo	21 - Credit memo	77.40
AAVDL	1900002569	21	2018-12-10	2019-01-10	2018-12_05	KR - Vendor invoice	31 - Invoice	-292.84
	1900002587	18	2018-12-13	2019-01-10	2018-12_05	KR - Vendor invoice	31 - Invoice	-346.14
	1900002602	18	2018-12-13	2019-01-10	2018-12_05	KR - Vendor invoice	31 - Invoice	-266.63
	1900002654	7	2018-12-24	2019-01-10	2018-12_05	KR - Vendor invoice	31 - Invoice	-63.02

Table 14: Integration of vendor open items into the process mining analyzer

The auditor may investigate the details of the related journal entries (Figure 77) using the accounting document reference of individual open items included in Table 14.

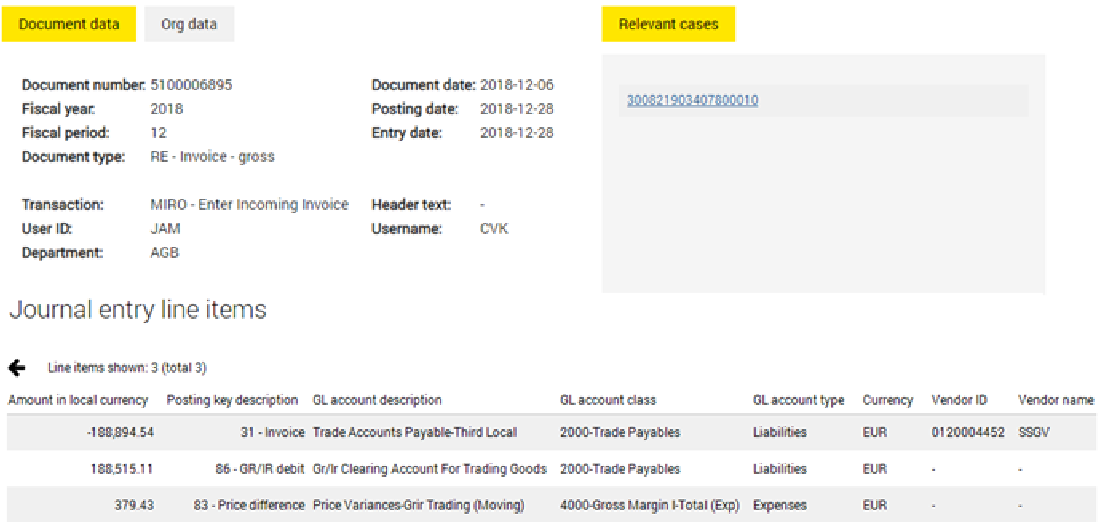


Figure 77: Journal entry details of an open item including information on related cases

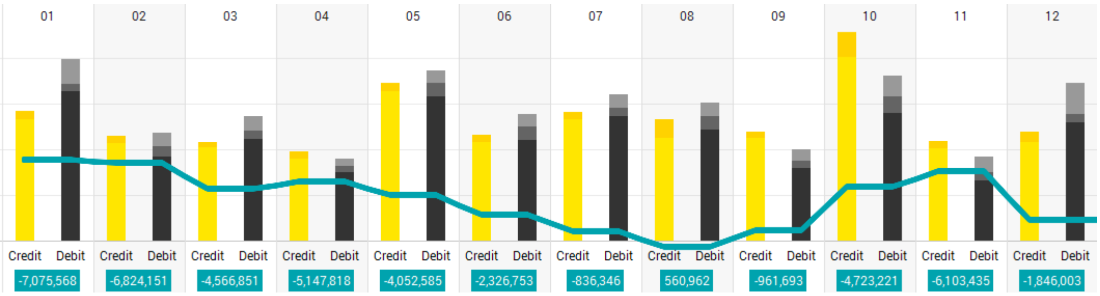
As the financial data has been integrated with the process data, the journal entry details not only include the accounting document header and journal entry line item information, but a reference to any cases affected by the journal entry. Hence, the auditor

[604] Cf. Chapter 3.3.2.

may identify and filter any cases in the process data that relate to a specific open item for further investigation in the process model.

Calculating the vendor balances and transaction volumes using the subledger data

Provided the opening balance of a vendor account at the beginning of the financial period, aggregating the open items at a specific point in time results in the vendor's ending balance. The subledger data further enables to visualize the periodical debit and credit movements on a specific vendor account. Figure 78 summarizes the development of the vendor balances throughout the period and at period end.



Vendor ID	Vendor name	▲ Ending balance CY	%	Invoice volume CY	%
0000410101	BOCDT	-4,584,746	248%	-2,113,910	3%
0120003495	AGUAL	-1,883,957	102%	-5,947,268	7%
0120003540	BRIDT	-214,437	12%	-2,736,163	3%
0120002455	BVCEY	-202,043	11%	-1,516,435	2%
0120006355	BXAIP	-196,661	11%	-2,347,057	3%

Figure 78: Vendor structure showing the development of vendor balances throughout the period

The column chart at the top of the figure displays the monthly debit (black) and credit (yellow) transaction volumes on the vendor accounts. The bars are stacked by the posting key information (differentiating, for example, invoicing, payment and credit memo activity) supporting the auditor to initially assess the composition of the debit and credit transactions. The turquoise graph indicates the total ending balance by month. The table at the bottom of the figure summarizes the ending balance and the total invoice volume related to a specific vendor in the audit period.[605] The total invoice volume summarizes the invoice amount of the invoice related activities in the process data, including the recording of invoices and invoice reversals. For validation

[605] In Figure 78, the audit period is abbreviated with "CY" (current year).

purposes, the recalculated vendor ending balances at period end can be reconciled to the subledger data and, in total, to the entity's trial balance.

With the integration of the subledger data, process mining may support the auditor in obtaining an initial understanding of the vendor structure and changes therein to identify risks of material misstatement.[606] For example, the auditor may filter key vendors in the table at the bottom of Figure 78 to investigate the monthly development of credit and debit activity and compare changes in the structure or composition of vendors with the previous period. The vendor-related information may be analyzed irrespective of the coverage of the related transactions by the process mining analyzer, i.e., without the need to use additional automated tools or techniques based on financial data.

Establishing the link between the subledger and the process data to assess coverage

The consideration of the subledger data and the reconciliation to the process data further enables to differentiate between the trade payables transaction volume that is covered by the process data and any additional postings made to the subledger. Table 15 shows the vendor balances and coverage of the debit and credit transaction volumes by the process data for two selected vendors.

Local currency	Company code	Vendor name	Activity level	Coverage	▼ 2018 (CY)
CHF	CH80	BXAIP	Beginning balance		0
			Debit	Invoice - With PO	2,150,396
				Not covered	32,317
			Credit	Invoice - With PO	-2,347,057
				Not covered	-32,317
			Ending balance		**-196,661**
			Balance change in %		-
		BVCEY	Beginning balance		-3,867
			Debit	Invoice - With PO	1,319,230
				Invoice - Without PO	1,942
				Not covered	1,020,857
			Credit	Invoice - With PO	-2,365,946
				Invoice - Without PO	-1,941
				Not covered	-172,317
			Ending balance		**-202,043**
			Balance change in %		*+5,124.6%*

Table 15: Coverage of individual vendor transaction volumes by the process data

606 Cf. Chapter 3.2.1.

For each vendor, the data is presented as follows:

$$\begin{aligned} & \textit{Beginning balance (sum of all open items at the start of the period)} \\ + \; & \textit{Debit postings within the period} \\ - \; & \textit{Credit postings within the period} \\ = \; & \textit{Ending balance (sum of all open items at the end of the period)} \end{aligned}$$

In the example provided in Table 15, the debit and credit postings are differentiated by their coverage by the process data. If the link between the amounts posted in the related journal entries and the activities in the process can be established,[607] the "coverage" column summarizes the related debit and credit transaction volumes by case type. These transaction volumes have been reconstructed from the process data using the principle of the invoice amount reconciliation introduced in Chapter 4.2.3. If any additional postings are made to the subledger without a reference to a process instance, they are presented as "not covered". These transaction volumes include, for example, postings of credit memos to a vendor that are made without a reference to a specific document.

To support reconciliation of the subledger data with the general ledger, the table may be adjusted to demonstrate how the transactions are recorded on the vendor-related general ledger accounts. Table 16 displays the total vendor activity by general ledger account.

Local currency	Company code	GL account class	GL account description	Activity level	Coverage	▲ 2018 (CY)
CHF	CH80	2000-Trade Payables	Trade Accounts Payable-Third Local	Beginning balance		-5,975,983
				Debit	Invoice - With PO	58,847,157
					Invoice - Without PO	9,994,441
					Not covered	9,481,387
				Credit	Invoice - With PO	-64,562,193
					Invoice - Without PO	-8,913,632
					Not covered	-4,856,169
				Ending balance		**-5,984,992**
			Trade Accounts Payable Ic	Beginning balance		-5,586,254
				Debit	Invoice - With PO	952,545
					Invoice - Without PO	22,506
					Not covered	18,265,001
				Credit	Invoice - With PO	-3,689,072
					Invoice - Without PO	-338,862
					Not covered	-14,371,512
				Ending balance		**-4,745,648**

Table 16: Reconciliation of the subledger data to the general ledger accounts

[607] Cf. Chapter 4.2.3.

Starting from the general ledger account view illustrated in Table 16, the auditor may decide to further investigate the transaction volumes related to trade payables that are not covered by the process data. In Table 17, the data is filtered to the transaction volume on the intercompany account[608] that is not covered by the process data. The transaction volume is displayed by the document type[609] used for posting the journal entries.

Local currency	Company code	GL account class	GL account description	Activity level	Document type description	▲ 2018 (CY)
CHF	CH80	2000-Trade Payables	Trade Accounts Payable Ic	Beginning balance		-104,711
				Debit	AB - Accounting document	66,229
					KA - Vendor document	90,250
					KZ - Vendor payment	26,400
					RE - Invoice - gross	143,977
					SA - G/L account document	17,909,740
					ZP - Payment posting	28,406
					Sum of Debit	18,265,001
				Credit	AB - Accounting document	-618,815
					KA - Vendor document	-26,400
					KZ - Vendor payment	-13,511,541
					RE - Invoice - gross	-20,385
					SA - G/L account document	-175,272
					ZP - Payment posting	-19,099
					Sum of Credit	-14,371,512
				Ending balance		**3,788,777**

Table 17: Non-covered intercompany transaction volumes by document type

In the example provided, a significant part of the transaction volume not covered by the process data relates to the document type “KZ - Vendor payment” on the credit side of the journal entry. In contrast, the transactions covered by the process data only include the document type “ZP - Payment posting” for recording payments. In SAP, the document types “KZ” and “ZP” are both used to record payments made to vendors. However, while the document type “ZP” is used to record automatic payment runs (transaction code F110), the document type “KZ” refers to manual payment postings using the transaction codes F-53 or F-58. The process data in the example does not include an activity for the posting of manual payments, leading to a significant transaction volume that is not covered by the process flow.

By separating the processing and recording events, tracking the invoice amount on event level and incorporating the subledger data, additional procedures for “not

[608] In Table 16, the trade payables related intercompany (IC) account is labeled as “Trade accounts payable IC”.

[609] In SAP, each line item of a journal entry is related to a document type. The document type is used to differentiate between different business transactions and to determine to which account type a specific document can be posted to (for example, vendor, customer or general ledger accounts).

covered" amounts can be designed directly in the process mining application without the need to use additional subledger analyzers.

Substantive analytical procedures on vendor-related activity throughout the period

In Figure 78, the vendor-related information is displayed at a high level of aggregation to support the auditor in obtaining an initial understanding of the vendor structure and identifying risks of material misstatement. As the approach of integrating the process data with the subledger data enables to identify the periodical debit and credit transaction volumes on a specific vendor account, analyses based on disaggregated data may be incorporated that support the auditor in performing substantive procedures to address identified risks.

In Table 18, the analysis of vendor balances introduced in Table 15 is extended to display the balances and transaction volumes monthly throughout the period.

Vendor name	Activity level	2018-01	2018-02	2018-03		2018-11	2018-12
AJQVB	Beginning balance	1,815,412	2,246,896	2,264,674		119,170	107,884
	Net activity	431,484	17,778	134,088		-11,286	-43,840
	Ending balance	**2,246,896**	**2,264,674**	**2,398,761**		**107,884**	**64,044**
AGUAL	Beginning balance	-761,038	-1,223,755	-1,028,273		-884,199	-978,262
	Net activity	-462,717	195,482	377,814		-94,063	-905,696
	Ending balance	**-1,223,755**	**-1,028,273**	**-650,459**		**-978,262**	**-1,883,957**
BUJQL	Beginning balance	947,456	-742,336	-575,448		-90,113	-132,642
	Net activity	-1,689,792	166,888	370,524		-42,529	-22,539
	Ending balance	**-742,336**	**-575,448**	**-204,924**	...	**-132,642**	**-155,181**
BVCEY	Beginning balance	-3,867	-2,414	-559		-369,849	-198,291
	Net activity	1,453	1,855	-4,799		171,558	-3,752
	Ending balance	**-2,414**	**-559**	**-5,358**		**-198,291**	**-202,043**
BXAIP	Beginning balance	0	-101,637	-31,730		0	-87,495
	Net activity	-101,637	69,907	21,031		-87,495	-109,166
	Ending balance	**-101,637**	**-31,730**	**-10,699**		**-87,495**	**-196,661**
GUJB	Beginning balance	64,530	-58,150	-38,710		-86,793	-122,881
	Net activity	-122,681	19,441	-89,984		-36,089	40,741
	Ending balance	**-58,150**	**-38,710**	**-128,694**		**-122,881**	**-82,141**

Table 18: Monthly development of vendor balances and transaction volumes

With the implementation suggested, the analysis of subledger data may be performed as with traditional automated tools and techniques that are using subledger data. Based on the objective of the audit procedure, additional attributes may be added to the analysis to investigate, for example, unusual debit or credit movements throughout the period or transaction codes, document types and posting keys used for recording transactions from a particular vendor. The auditor may select a specific vendor and/or month to drill down to the related journal entries and open items for further investigation. However, in contrast to traditional analyzers based on financial data only, the proposed concept enables the auditor to integrate the process perspective into the analysis. For the transaction volumes covered by the activities in the process, the auditor may switch to the process-related analyses to investigate, for example, the activity sequence and users involved in the initiation and processing of transactions related to a specific vendor.

Analyzing the vendors involved in a case

The analyses on the vendor balances and transaction volumes presented in this chapter are based on the journal entry postings to the related vendor accounts, i.e., the invoices and payments posted are related to the vendor included in the respective journal entry information of the subledger. On the other hand, the case table presented in Table 2 focuses on the process perspective only and stores the vendor the respective purchase order has been sent to as a case attribute. As multiple vendors may be involved in a single process instance,[610] the vendor information in both data sources may differ. Multiple vendors related to a case cannot be accurately presented with process mining if the vendor-related information is stored on the level of the case rather than as event attribute.

Figure 79 integrates the concept of multiple vendors introduced in Figure 45 with the process perspective.

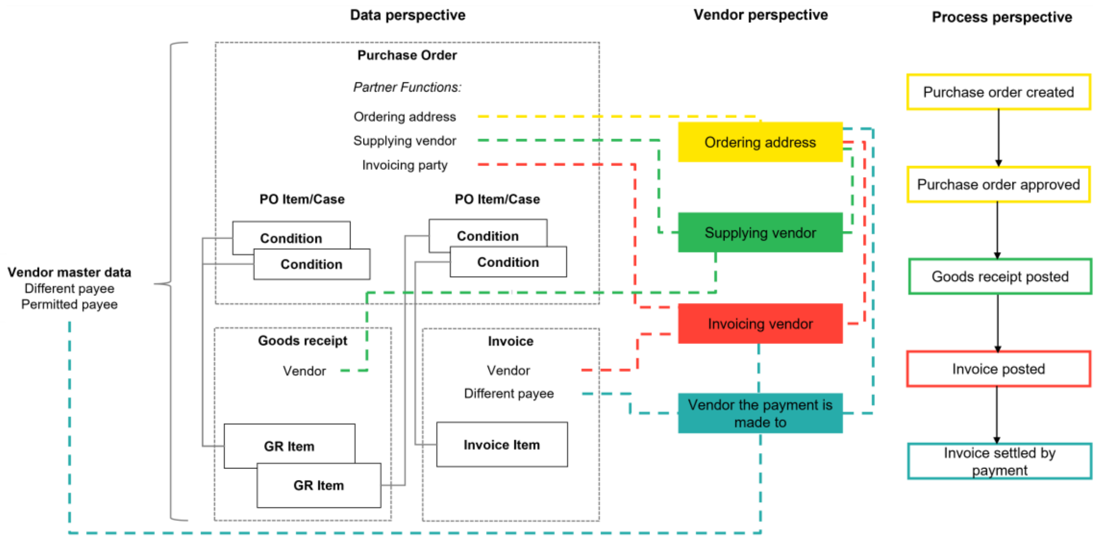


Figure 79: Integrating the concept of multiple vendors with the process perspective

To ensure an appropriate presentation of the purchase amounts, any events in the event log that relate the to the purchase order document need to store the vendor related to

[610] Cf. Chapter 3.3.2.

the ordering address as event attribute. Similarly, the events related to the posting of a goods receipt, an invoice and a payment need to be related to the supplying vendor, the invoicing vendor and the vendor the payment is made to, respectively. Allocating the vendor information as an attribute in the event log ensures that information is displayed appropriately for cases that include events related to different vendors. To support the auditor's process understanding, a separate analysis of the vendors related to a case may be provided, summarizing the number and roles of different vendors involved in a case.

4.2.5 Analytical procedures addressing the completeness of trade payables

Auditing the completeness of purchases

A common fraud scheme related to trade payables involves holding back an invoice by not entering it into the system in the audit period it belongs to and thus shifting the expenses into the subsequent period.[611] Not recording purchases that have been received leads to an understatement of trade payables at period end and may lead to overstated income and equity.[612] Hence, the completeness of liabilities and related expenses is a primary risk in purchase to pay. As performing tests of details of account balances may be difficult to determine if potentially unrecorded transactions exist, the auditor usually relies on tests of controls and substantive tests of transactions to address the completeness assertion.[613] Effective internal controls over the completeness of purchasing transactions, for example, the sequential prenumbering of received invoices,[614] may reduce the risk that invoices are missing. Common substantive tests of transactions include a search for unrecorded liabilities that is traditionally performed by physically tracing significant vendor invoices received after the balance sheet date to the acquisitions journal.[615] If the related goods or services have been received before the balance sheet date, the auditor determines if the goods or services have been recorded and the respective amounts have been included in the accrual for outstanding invoices.

However, this approach presumes that all invoices received after the balance sheet date are made available to the auditor. This includes both invoices that have been entered

---

611 Cf. ACFE (2018), p. 3626.
612 Cf. ELDER, RANDAL J. et al. (2020), p. 574.
613 Cf. ibid.
614 Cf. ibid., p. 573.
615 Cf. ibid.

into the accounting system or an invoice register for unentered invoices after the balance sheet date, and invoices that have not yet been entered but are physically stored somewhere at the entity. Frequently, if it may be reasonably expected that this presumption is not met by the auditee, the auditor considers whether external confirmations of trade payables performed as substantive procedures according to ISA 330 are more appropriate.[616] The manual review of vendor invoices received after the balance sheet date or obtaining external confirmations are especially challenging if the audit opinion is planned to be issued shortly after the balance sheet date, as invoices or requested confirmations related to the audit period may not have been received yet.

Analytical approach to address the completeness of liabilities and expenses

When the risk of material misstatement from unrecorded liabilities is not significant, a properly designed substantive analytical procedure may adequately address the risk of materially understated liabilities.[617] A benefit of enhancing the data used in an audit of financial statements by process data is the evidence on the initiating activities of the process, for example, the creation of a purchase order in purchase to pay. The availability of the purchase order amount and the approval date of the purchase order, together with the ability to track the metrics related to a case over time[618] enable an analytical approach to address the completeness of purchases during the period and liabilities at period end.

The analysis relies on the information on purchase order volumes throughout the period, together with the information on (partial) goods and/or invoices received for a case. While most ERP systems support the match between prices or quantities on the purchase order, goods receipt and invoice documents, the reconciliation is only performed if all documents are present for a case. The status of a purchase order at a specific point in time is not tracked in SAP. With process mining, provided the metrics related to a case are tracked on the level of the individual events, the purchase order amount that is still open at a specific point in time may be determined. For purchase orders where no goods[619] or invoices have been recorded yet, the open purchase order

616 Cf. IFAC (2021), ISA 505, para. 2; IFAC (2021), ISA 330, para. 7(b).
617 Cf. LEVY, HOWARD B. (2021).
618 Cf. Chapter 4.2.3.
619 For simplification purposes, the approach presented in this chapter is limited to the purchase of goods. However, a similar analysis may be designed for the purchase of services using the service entry sheet information available in SAP.

amount is calculated using the purchase order amount from the "purchase order created" event and considering any change events related to the price or quantity of the purchase order that may have occurred. For cases where a (partial) delivery was received, the value of the related goods receipt posting is subtracted from the open purchase order amount. Similarly, if an invoice has been received, the open purchase order amount is reduced by the related invoice amount. If both a goods receipt and an invoice are available, the higher of the two amounts is used for the calculation. In this scenario, any differences between the documents related to the case are identified in the three-way-match between purchase order, goods receipt and invoice.[620]

The analysis of open purchase order balances may be designed in two steps and includes:

(1) performing substantive analytical procedures on aggregated open purchase order balances and
(2) performing further substantive procedures by disaggregating the data to investigate significant open purchase order volumes identified.

At (1): Analytical procedures on aggregated open purchase order balances

Table 19 shows the calculated open purchase order balances as of period end.

PO has deletion flag	Has goods receipt	Has invoice	< 2017-01	2017-12	2018-12
No	No	No	7,429	62,934	9,542,145
		Yes	0	0	33,499
	Yes	No	8,398	33,796	40,803
		Yes	986,154	2,100,899	181,191
Yes	No	No	40,231	883,320	1,784,288
	Yes	No	0	28,480	252,656
		Total	1,042,213	3,109,429	11,834,582

Table 19: Aggregated analysis of the calculated open purchase order amount at period end

The purchase order amounts are categorized based on whether the related purchase order is currently flagged for deletion and whether goods or invoices have already been received. In the example provided, a total purchase order volume of around 1 million EUR from periods before 2017, 3,1 million EUR from 2017 (previous period) and 11,8

[620] Cf. Chapter 3.2.1.

million EUR from 2018 (audit period) is still open at the balance sheet date (31 December 2018), resulting in a total open purchase order balance of around 15,9 million EUR.

The aggregated categorization of open purchase order volumes may support the auditor in assessing related risks of material misstatement and selecting subsets of the population for further investigation:

- *Cases without deletion flag and without any other document:* Based on the understanding obtained of the business, the vendor structure and changes therein compared to the prior period, the auditor may determine if the number and amount of purchase orders without a goods receipt or invoice is in line with the expectation. As a common reason for purchase orders without any other documents may be a missing approval, the analysis may be focused on those purchase orders that have been approved (i.e., that include a purchase order approval event). In case a significant open purchase order amount is remaining at period end, the auditor might expect that these cases have been initiated near period end and select the related cases for further investigation.[621]
- *Cases without deletion flag, with goods receipt but without invoice:* Transactions where goods have already been received but the invoice has not yet been received at period end are usually subject to an accrual for outstanding invoices. If a significant open purchase order balance is remaining at period end, the auditor may perform procedures to confirm that the accrual for outstanding invoices is reasonable, for example, by investigating whether the open purchase orders are considered within the accrual for outstanding invoices.
- *Cases without deletion flag where goods and invoices have already been received:* This category summarizes transactions where only partial deliveries and invoices have already been recorded and/or transactions where there are differences in the three-way-match[622] between purchase order, goods receipt

[621] The procedures that may be performed for further investigation are described in the second part of the analysis introduced later in this chapter.
[622] Cf. Chapter 3.2.1.

and invoice. Depending on the amounts of the related invoices[623] and the results of the three-way-match testing performed, the auditor may determine if further investigation is necessary.

- *Cases without deletion flag, without goods receipt but with invoice:* As the risk of loss or damage to the goods is usually transferred from the seller to the buyer at the time of the delivery, the auditor would usually not expect scenarios where the invoice is posted before the goods receipt. In the standard configuration of the SAP MM module, an indicator that prevents the recording of an invoice prior to the recording of a goods receipt is configured in the vendor master data. However, there may be situations when the invoice is posted before the goods receipt, for example, when the International Commercial Term (Incoterm) CAD (Cash Against Documents)[624] is used and the shipment takes a significant amount of time. In this scenario, the balance on the GR/IR account waiting for the goods receipt clearing is considered to be stock in transit. If the purchase order volume related to cases with invoice but without goods receipt is significant, the auditor might select the cases for further investigation.
- *Cases with deletion flag and without any other document:* The auditor may determine that purchase orders that are flagged for deletion and are not related to a goods receipt or invoice posting do not pose a risk with regard to the completeness of liabilities. If the related purchase order amount is significant, the auditor may discuss with the auditee to what extent such situations happen and evaluate potential recommendations in the management letter.
- *Cases with deletion flag but with goods receipt and/or invoice event:* While no transactions can be processed against a purchase order that is currently flagged for deletion, the deletion flag may be set or reverted as part of the purchase order change process. If there is a material amount of open purchase order volumes marked for deletion where goods and/or invoices have been received, the auditor may analyze the activity flow for these cases to understand the reasoning for setting a deletion indicator. In the example provided, no invoice has

---

623 The analysis in Table 19 is displaying the calculated open purchase order balances. The metric may be changed to investigate, for example, the related invoice or goods receipt activity in the respective categories.

624 The incoterm CAD is explicitly used to ensure the purchasing party pays for the goods before receiving the actual delivery. The supplier maintains the ownership rights of the goods until the purchasing party conducts the payment. These transactions usually involve a bank as an intermediary who holds the related documents from the supplier and the payment form the purchasing party. After the goods are received and accepted by the buyer, the payment is released to the supplier.

been received for cases with a deletion flag yet, but there are some cases where goods have been received. The auditor may decide to select these cases for further analysis to identify any potential remaining obligations.

In analyzing the different categories, the auditor may consider the planned delivery dates specified in the purchase orders.[625] Table 20 summarizes the part of the open purchase order volume as of period end (presented in Table 19) with a scheduled delivery date after the balance sheet date.

PO has deletion flag	Has goods receipt	Has invoice	< 2017-01	2017-12	2018-12
No	No	No	0	53,615	9,483,089
		Yes	0	0	33,437
	Yes	No	0	0	6,767
		Yes	0	0	31,874
Yes	No	No	0	15,649	213,091
	Yes	No	0	0	24,104
		Total	0	69,264	9,792,362

Table 20: Open purchase order amount with a scheduled delivery date after the balance sheet date

From the open purchase order volume of 15,9 million EUR at the balance sheet date,[626] around 9,9 million EUR have a scheduled delivery date in the subsequent period. The auditor may identify from the data if the auditee customized the standard configuration of the scheduled delivery date in SAP and makes consistent use of the functionality, for example, by comparing the purchase order creation dates with the scheduled delivery dates, identifying any missing delivery dates and investigating purchase order change events that refer to the scheduled delivery date. For those purchases where goods have already been received, the planned delivery date may be compared to the actual date of the goods receipt. In the example provided, the scheduled delivery dates in the process data range from 2017 to 2019 without any missing values. However, as presented in Table 20, there are purchase orders with a delivery scheduled after the balance sheet date but with goods and/or invoices received prior to the balance sheet date. The auditor may investigate if these purchase orders relate to a significant volume

[625] Purchase orders in SAP include a field for the planned delivery date of the ordered goods. In the standard SAP configuration, the scheduled delivery date may not be blank or set to a date in the past. The default date is configured to seven days after the date of the purchase order creation and may be changed upon or after the purchase order creation.

[626] Cf. Table 19.

of goods receipts and invoice amounts in the audit period and inquire with the entity as necessary.[627]

If the delivery dates are determined to be a reliable indicator for the expected date of the goods receipt, the auditor may focus on those purchase order volumes that are open as of period end and have a scheduled delivery date before the balance sheet date. Table 21 summarizes the related open purchase order amounts.

PO has deletion flag	Has goods receipt	Has invoice	< 2017-01	2017-12	2018-12
No	No	No	7,429	9,320	59,056
		Yes	0	0	62
	Yes	No	8,398	33,796	34,036
		Yes	986,154	2,100,899	149,317
Yes	No	No	40,231	867,671	1,571,197
	Yes	No	0	28,480	228,552
		Total	1,042,213	3,040,165	2,042,219

Table 21: Open purchase order amount with a scheduled delivery date before the balance sheet date

In the example provided, significant open purchase order volumes from the current audit period relate to cases with a deletion flag where no goods or invoices have been received (1,6 million EUR). Further, in this category, a significant purchase order amount from the prior period (0,9 million EUR) is still open at the end of the audit period. Although these transactions (colored yellow in Table 21) may not pose a risk of incomplete liabilities, the auditor may discuss with the auditee to what extent such situations happen and evaluate potential recommendations in the management letter.

The second group of significant open purchase order volumes at period end (colored orange in Table 21) relates to cases without deletion flag where goods and invoices have already been received. These transactions may either only have partial deliveries or invoices and/or there are differences in the three-way-match between purchase order, goods receipt and invoice. As the related volume of open purchase order amounts is significant, further investigation is necessary.

[627] In confirming the consistent use of scheduled delivery dates, the auditor may further filter the significant open purchase order amount of 9,5 million EUR to understand the root-cause of a planned delivery date after the balance sheet date. For example, the purchase orders might have been created near period end.

At (2): Investigating significant open purchase order balances identified

The second part of the procedure involves using disaggregated data to further investigate any significant open purchase order volumes identified. The procedures described in this section focus on purchases where goods or invoices have not been recorded at the balance sheet date, as the related open purchase order balances are of particular interest when addressing the risk of unrecorded liabilities.[628]

Cases without deletion flag and without any other document

If there is a significant amount of open purchase orders where no goods receipts or invoices are available yet at the balance sheet date, the auditor may usually expect that these cases have been initiated near period end, i.e., within the usual business cycle time for purchases from the specific vendors involved. The auditor may disaggregate the analysis for these purchase order volumes by vendor and month to analyze the monthly development of purchase order, goods receipt and invoice volumes. Table 22 shows the open purchase order amounts by vendor and month for those purchase orders that are not flagged for deletion and where no goods or invoices have been received.[629]

Vendor	< 2018-01	2018-01	2018-02	2018-03	2018-04	2018-05	2018-06	2018-07	2018-08	2018-09	2018-10	2018-11	2018-12
BAFSJ	0	0	0	0	0	0	0	3.220.500	3.220.500	3.220.500	3.220.500	3.232.016	3.232.016
BKBKE	0	0	0	0	0	0	0	1.253.201	830.201	830.201	1.294.696	1.295.191	1.295.191
BXAIP	0	0	0	0	0	0	0	0	0	26.042	640.424	807.555	807.555
BOCDT	0	0	0	0	0	0	0	181.615	181.615	197.246	622.562	700.133	700.133
ARILO	1.002	1.002	1.002	1.002	1.002	1.002	1.002	1.002	1.002	1.002	478.271	535.267	535.267
ACDMC	0	0	0	0	0	0	0	0	0	0	263.950	355.612	355.612
BTKHS	0	0	0	0	0	327.601	327.601	327.601	327.601	327.601	327.601	327.601	327.601
AZTLV	0	0	0	0	0	0	0	12.314	12.314	54.299	207.894	274.440	322.272
AAHZF	0	0	0	0	0	0	0	0	0	0	309.135	309.135	309.135
AJQVB	0	0	0	0	0	0	0	105.807	105.807	258.321	258.321	264.654	264.654
AMRBC	0	0	0	160	160	160	236.022	236.576	237.387	237.387	237.387	237.387	237.387
AKRTJ	0	0	0	0	0	0	0	0	0	22.311	90.431	194.205	194.205
EALP	0	0	0	0	0	0	0	0	0	0	96.838	142.523	142.523
GUJB	474	474	699	888	967	967	967	967	967	967	1.011	138.281	138.281
BXJQQ	0	0	0	0	0	0	0	0	0	0	35.726	126.867	126.867
IFHQ	0	0	0	0	0	0	0	0	31.654	59.180	126.598	126.598	126.598
ADHDV	0	0	0	0	0	0	0	13.129	13.129	13.129	13.129	114.259	114.259
ADLAG	0	0	0	0	0	0	0	0	0	0	0	85.525	85.525
ASZHO	53.615	53.615	53.615	53.615	53.615	53.615	53.615	53.615	53.615	53.615	53.615	56.623	56.623
BRCRY	0	0	0	0	0	0	0	0	0	0	0	49.466	49.466
BUJQL	414	414	414	414	414	414	414	414	414	414	1.288	22.999	22.999
ATAVY	0	0	0	0	0	0	0	0	0	0	0	22.798	22.798
BPVQP	0	0	0	0	0	0	0	0	0	0	0	21.910	21.910
XTWA	0	0	0	0	0	0	0	0	0	0	0	14.611	14.611
BGAEV	0	0	0	0	12.016	12.016	12.016	12.016	12.016	12.016	12.016	12.016	12.016
RGGX	7.429	7.429	7.429	7.429	7.429	7.429	7.429	7.429	7.429	7.429	7.429	7.429	7.429
SSGV	0	0	0	0	0	0	0	7.427	7.427	7.427	7.427	7.427	7.427
AJXWP	0	0	0	0	0	0	0	0	0	0	1.173	5.992	5.992
BNEYX	0	0	5.702	5.702	5.702	5.702	5.702	5.702	5.702	5.702	5.702	5.702	5.702
BPFMG	0	0	0	90	90	90	90	90	90	90	90	90	90
Total	**62.934**	**62.934**	**68.862**	**69.301**	**81.396**	**408.997**	**644.858**	**5.439.405**	**5.048.870**	**5.334.880**	**8.313.213**	**9.494.313**	**9.542.145**

Table 22: Open purchase order amount without any documents received by vendor and month

[628] These purchases include cases without deletion flag and without any other document, cases without deletion flag, with goods receipt but without invoices and cases without deletion flag where goods and invoices have already been received. Examples for the nature and extent of further substantive procedures for significant open purchasing volumes in the other categories are included in the previous section.

[629] The total open purchase order volume of 9.542.145 EUR reconciles to the respective category in Table 19.

The auditor may filter the data to significant vendors identified and determine if further investigation is needed. If the auditee makes use of the scheduled delivery date, the investigation of significant vendor balances may be limited to those cases where a delivery is expected within the audit period. In the example provided, the total open balance in this category would be reduced to 59.056 EUR[630] and will likely not indicate a risk of material understatement of liabilities.

However, process mining also supports alternative procedures in case the scheduled delivery date is not used consistently in SAP. In the example provided, the open purchase order volume of around 3,2 million EUR for vendor “BAFSJ” will likely be material. A significant portion of the related purchase orders has been initiated in July 2018, i.e., six months before the balance sheet date. With process mining, the auditor may build an expectation on the throughput time between purchasing, delivering and invoicing activities for this vendor. For example, the auditor may analyze the throughput time between purchase order approvals and goods receipt posting (or invoice posting, respectively) on vendor level for the entire period and then confirm the expected throughput time with the development of open purchase order balances for this vendor. To understand the nature of significant open transactions for specific vendors, the analysis can be expanded by adding further information depending on the nature of the auditee’s business. Based on the characteristics of the purchases related to the open purchase order amount, the auditor may then further refine the expectation, for example, by considering the throughput time for certain invoice amount intervals or specific material groups only. If the auditor expects that based on the procedures performed, goods or invoices should have already been received for (parts of) the open purchase order balances, he or she may drill-down to the document and journal entry data to obtain the detailed case information and follow-up with the entity.

Cases without deletion flag, with goods receipt but without invoices

Transactions where the goods have been recorded but no invoice has been received as of the balance sheet date are subject to an accrual for outstanding invoices. If a significant open purchase order balance is remaining at period end, the auditor may disaggregate the data to include information on the vendors related to these purchases. For vendors with significant open purchase order volumes, the auditor may drill-down

[630] Cf. Table 21.

to the individual case and goods receipt document information. In comparing the transactions to the items considered in the accrual for outstanding invoices, the information may be leveraged for confirming the reasonableness of the accrual and identifying any transactions that may, individually or when aggregated, result in a remaining risk of unrecorded liabilities.

By this, process mining enables a data-driven identification of outstanding invoices, without requiring the auditor to manually search for significant invoice receipts after the balance sheet date that may or may not have already been recorded.

Cases without deletion flag where goods and invoices have already been received

Transactions where both goods and an invoice have already been recorded at period end might still result in open purchase order balances. These transactions may include (1) cases that have already been fully delivered and invoiced as of the balance sheet date and (2) cases where only partial deliveries and invoices have been recorded.

For the first group of transactions, the remaining open purchase order balance may be caused by differences identified in the three-way-match between purchase orders, goods receipts and invoices. The purchase order volume in this sub-population can be identified by filtering the data to those transactions subject to the three-way-match. If the auditor has reperformed the three-way-match and confirmed that appropriate follow-up procedures have been performed for exceptions outside tolerances, the open purchase order balance related to these cases does not indicate a remaining risk of materially understated liabilities.

The three-way-match in SAP is only performed if all related documents (i.e., purchase order, goods receipt and invoice) are available. As such, the second group of transactions, where only partial deliveries and invoices have been recorded, is not subject to the three-way-match. The auditor may compare the related purchase order, goods receipt and invoice amounts for these open purchase orders[631] to identify the partial goods receipt and/or invoice activity that has not yet been recorded. If the related open purchase order amount is significant for any of the vendors, the auditor may determine that further investigation is necessary. The investigation may be performed similar to

[631] In the analysis presented in Table 22, the purchase order, goods receipt and invoice volumes related to the open purchase order balances may be analyzed by changing the amount metric of the table, respectively.

the analysis of cases without invoice and goods receipt and cases with goods receipt but without invoice, as described earlier in this section.

To summarize, the analysis of open purchase order balances is designed to respond to the risk that purchases have occurred but no goods receipts or invoice receipts have been recorded, resulting in incomplete liabilities. Considering the understanding obtained of the business, the vendor structure and changes therein, the auditor may determine if the number and amount of open purchase orders is in line with his or her expectation when compared to the prior period. The analysis supports a data-driven identification of outstanding invoices and is facilitated by focusing on vendors with significant purchasing volumes or open purchase order balances. For the purchase orders selected for those vendors, data-enabled root-cause analyses may be performed to identify why those purchase orders are still open at period end. If the explanation for goods receipts or invoice receipts being delayed does not meet the auditor's understanding of the business, the expectation can be refined by analyzing the throughput time between activities evident in the historical data of the specific vendor and material. If these procedures do not provide sufficient appropriate evidence for the completeness of transactions related to a purchase order, additional tests of details may be performed to determine if a misstatement has occurred. However, compared to the traditional search for unrecorded liabilities, the auditor may focus any further procedures on those subsets of the data with a remaining risk of material misstatement due to unrecorded liabilities.

4.3 Multitude of process variations and their evaluation in an audit of financial statements

4.3.1 Emergence of process variations in the practical application

Even if processes are expected to follow a designed target process, today's ERP systems allow for some flexibility in the process execution. Deviations from the designed target process are necessary, for example, to allow for exceptional transactions or changes in the organization that occur in the daily business but are not reflected in the initial process model. The practical application of process mining confirms that process mining increases the quality of the process understanding obtained by the auditor. The transparency achieved over all the different process paths available in the data demonstrates that the manual walkthrough procedures for individual transactions do not lead to an appropriate representation of the actual process execution. The flowchart describing the flow of transactions through the process derived from traditional procedures frequently only reflects the most frequent variation of the process.[632] However, due to the multitude of process variations they cannot be evaluated in their entirety in practice. The evaluation of piloting results reveals the audit teams' uncertainty about (a) the degree to which variations need to be analyzed and (b) the qualitative and quantitative criteria applicable to the analysis. These findings give reason to further explore methods and instruments supporting the auditor in evaluating process variations.

Process variations result from different sequences of activities used for the processing of cases throughout the observation period. Factors that may influence the number of variations include the length of the observation period, the number of cases (and thus the size of the entity and the number of business units considered in the analyzer) and the number of activities and events. Although many researchers describe process mining as a topic fairly well researched over the last two decades,[633] little attention is devoted to field studies exploring the application of process mining across different industries and related practical challenges.[634] As a result, only little real-life data is available for research,[635] and information on the characteristics of the data underlying the theoretical descriptions are limited. Table 23 provides an overview on selected process characteristics identified in process mining related literature.[636]

No.	Author(s)	Process	Cases	Activities	Events	Variations
#01	JANS (2009)	Purchase to pay	10.000	7	61.562	170
#02	JANS et al. (2010)	Purchase to pay	10.000	7	61.562	161
#03	JANS/DEPAIRE/VANHOOF (2011)	Purchase to pay	26.185	7	181.845	304
#04	JANS/ALLES/VASARHELYI (2012)	Purchase to pay	26.185	7	181.845	304
#05	JANS/ALLES/VASARHELYI (2014)	Purchase to pay	26.185	7	181.845	304
#06	CHIU (2018)	Purchase to pay	26.185	7	181.845	980
#07	HOBECK et al. (2021)	Betting platform marketplace	2.735	11	22.772	414
#08	BECKER/BUCHKREMER (2019)	Annual account review	309	N/A	4.506	6
#09	MANNHARDT et al. (2019)	Sepsis patient	1.050	16	N/A	846
#10	MANNHARDT et al. (2019)	Road traffic fine	10.000	10	N/A	69
#11	DE WEERDT et al. (2013)	IT helpdesk process	24.770	18	124.217	1.174
#12	DE WEERDT et al. (2013)	CRM process	956	22	11.218	212
#13	DE WEERDT et al. (2013)	CRM process	17.812	42	83.286	1.908
#14	DE WEERDT et al. (2013)	Incoming document handling	12.391	70	65.653	1.411

Table 23: Selected process characteristics in literature dealing with analyzing process variations

632 Cf. Figure 41.
633 Cf. VOM BROCKE, JAN et al. (2021), p. 483; WERNER, MICHAEL/GEHRKE, NICK (2019), p. 208; KERREMANS, MARC (2019), p. 4; MAITA, ANA R. C. et al. (2017), p. 506.
634 Cf. Chapter 2.2.
635 Also see CHUI, TIFFANY (2018), p. 3.
636 The analysis is limited to literature providing information on the number of resulting process variations.

Analyzing the limited information available on datasets used in literature shows that many examples provided and conclusions drawn are based on data obtained from a single case study or on sanitized datasets that are publicly available. For example, all of the studies summarized in Table 23 that relate to the application of process mining in internal or external auditing (rows #01 to #06) are based on data obtained from the same European bank.[637] MANNHARDT et at. (rows #09 and #10) analyze sanitized data publicly available in an online repository for process mining.[638] Further, current research frequently only considers a limited population of data, for example, a random sample from the full population[639] or process instances that lead to an invoice in a selected month.[640] The highest number of process variations can be observed for three of the real-life datasets used by DE WEERDT et al. (rows #11, #13 and #14). When compared to the other datasets, these populations include a considerable amount of process instances and activities.[641] However, due to the different research objectives, observation periods, business processes analyzed and process mining applications used in the studies provided, little comparative information about the process characteristics and resulting process variations may be derived. In addition, differences in the process mining software used may result in a different number of process variations even if the analysis is based on the same dataset.[642] JANS/ALLES/VASARHELYI conclude that whether the number of resulting process variations in their case study is usual for a company of that size and complexity is a question yet to be answered.[643]

When compared to the case studies conducted in literature, the number of process variations resulting in audit practice is significantly higher. Table 24 provides an overview of the process characteristics, including the number of resulting process variations, for selected entities in Germany, Switzerland and Austria.

---

[637] Cf. Chapter 2.2. Although the authors included in rows #03 to #06 describe the population of data differently as "*payments made in January 2007*" (JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2014), p. 1758), "*invoices paid in January 2007*" (JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2012), p. 14), "*all invoices of January 2007*" (JANS, MIEKE/DEPAIRE, BENOIT/VANHOOF, KOEN (2011), p. 33) and "*purchases that led to an invoice in January 2007*" (CHUI, TIFFANY (2018), p. 18), the process characteristics indicate that these studies are based on the same dataset. The studies included in rows #01 to #02 only consider a random sample of 10.000 cases from the same financial institution, cf. JANS, MIEKE (2009), p. 122; JANS, MIEKE et al. (2010), p. 10. The varying number of resulting process variations is caused by changes in the process mining software "ProM", cf. CHUI, TIFFANY (2018), p. 8.

[638] Cf. MANNHARDT, FELIX et al. (2019), p. 616.

[639] Cf. ibid.; JANS, MIEKE et al. (2010), p. 10; JANS, MIEKE (2009), p. 122.

[640] Cf. JANS, MIEKE/DEPAIRE, BENOLT/VANHOOF, KOEN (2011), p. 3.

[641] No information on the length of the observation period is provided by the authors.

[642] Cf. rows #05 and #06 in Table 23; CHUI, TIFFANY (2018), p. 8.

[643] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2014), p. 1759.

Entity no.	Observation period	Number of cases	Number of activities	Number of events	Number of variations
#01	12 months	881.186	21	3.838.398	5.674
#02	6 months	807.306	22	1.847.607	2.743
#03	12 months	747.792	26	4.051.684	30.613
#04	9 months	623.893	55	4.318.209	17.458
#05	4 months	305.421	30	1.710.517	4.582
#06	12 months	196.350	34	1.652.191	9.740
#07	12 months	147.115	54	1.206.127	6.231
#08	12 months	135.853	30	3.066.125	11.946
#09	9 months	132.252	24	1.541.183	10.238
#10	9 months	117.781	38	2.227.228	10.140
#11	9 months	109.915	44	2.993.866	10.748
#12	12 months	104.996	32	381.794	2.714
#13	12 months	61.148	18	342.645	737
#14	12 months	38.927	23	234.870	1.355
#15	12 months	15.192	23	123.676	1.954
#16	8 months	8.906	40	100.108	1.749

Table 24: Overview of the process characteristics of selected entities

All examples are derived from engagements using the same version of the process mining application presented in Chapter 3. Still, the analysis reveals no explicit pattern or relationship between the process characteristics and the number of process variations. For example, entity #01 has the largest number of cases and a large number of events, however, the number of process variations is comparatively low. The largest number of variations is observed for entity #03 that only includes a small number of activities. In contrast, entity #07 has the largest number of activities but a comparatively low number of process variations. Entities #09, #10 and #11 have the same length of the observation period, a comparable number of cases and resulting process variations, but differ in the number of activities and events. Although the length of the observation period and the number of process instances, activities and events impact the number of resulting process variations, there is a range of additional influence factors that may not be identified by analyzing the key characteristics of the process only. Process variations result from repetitions of individual activities or sequences of activities as well as from missing or additional activities when compared to the most common process path. Process variations may further have technical root-causes including the process mining algorithm used, the scope and size of the event log and decisions made as part of the data preparation, such as the definition and granularity of the selected process instance.[644]

The examples demonstrate that the question raised by JANS/ALLES/VASARHELYI, whether the number of identified process variations is usual for a company of that size

[644] Cf. Chapter 4.1.1.

and complexity[645] may likely never be answered. Considering the variety in resulting process variations even for entities of comparable size operating in the same industry and analyzed with the same process mining application, it may not be possible for the auditor to develop an expectation about the number of process variations. As such, establishing a database supporting the comparison of process characteristics with a peer group[646] may not be an appropriate instrument for assessing the reasonableness of the number of process variations. Moreover, a large number of process variations does not necessarily indicate a higher risk of material misstatement for the class of transactions and related accounts and disclosures.[647] As such, the auditor is not particularly interested in each resulting process variation but rather in the root-causes of the process variations. For example, the competence of the users involved in the process, the types of transactions processed and existing policies or business practices for processing these transactions might represent audit relevant influence factors on the number of distinct process executions, as they may give rise to or address a risk of material misstatement. As process variations may not be audited in their entirety, the auditor needs to be supported with appropriate techniques to support identifying, assessing and addressing any risks of material misstatement resulting from the different ways the process is executed.

4.3.2 Evaluation of process variations in an audit of financial statements

Shortcomings of evaluation approaches in audit-related process mining literature

Most authors providing information on the number of resulting process variations do not analyze the variations in detail or provide guidance on the nature and extent of procedures necessary to handle these variations in practice.

In an example provided by the AICPA, the auditor determined that all process variations need to be understood and investigated, as they may provide evidence, for example, on (a) an incomplete process understanding that needs to be updated, (b) changes in the process or (c) a new risk of material misstatement.[648] The example is based on data from an order to cash process[649] that in general involves more complex transaction processing when compared to the purchase to pay process. Still, there is only a very

[645] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2014), p. 1759.
[646] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2012), p. 17.
[647] Also see AICPA (2017), para. A.65.
[648] Cf. ibid., para. A.67.
[649] Cf. ibid., para. A.60.

limited number of resulting process variations that is evaluated by the auditor through "*inquiry and other procedures*"[650]. However, the multitude of variations observed in practice implies that they cannot be evaluated in their entirety over the course of an audit.

JANS/ALLES/VASARHELYI only analyze the six most frequent process variations in detail and most of these variations already require further investigation by the auditor.[651] To support the auditor in evaluating the remainder of 298 process variations, the authors propose to develop and disseminate a knowledge base of compliant and anomalous patterns, i.e., sequences of activities in a process.[652] However, due to the variety of industries, processes, activities, IT systems and configurations, a repository of process variations does not seem feasible in practice. Whether a specific process trace is acceptable may further depend on additional factors that may or may not be evident in the data, such as the entity's procedures, policies and effectiveness of internal controls, the applicable GAAP and other national laws and regulations.

In order to reduce the number of process variations to a realistic number that may be investigated in detail, JANS suggests applying the pareto principle to separate routine process paths from infrequent process executions.[653] When applied to process mining, the pareto rule implies that 80 percent of the event log may be described with 20 percent of the variations.[654] However, especially the infrequent process paths including cases that are processed differently from routine transactions are of special interest for the auditor, as they may indicate a risk of material misstatement due to fraud according to ISA 240.[655]

From the studies listed in Table 23, only CHUI is examining the entire population of 980 distinct process variations. The variations are manually classified into standard and non-standard process paths with 95 percent of the variations (931 out of 980) being categorized as non-standard. As part of the analysis, a large portion of the non-standard variations is categorized as unusual and a total of 6.987 cases is handed over to the auditor for further investigation.[656] However, details on the procedures that may be

[650] AICPA (2017), para. A.71 and A.72.
[651] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2012), pp. 16f.
[652] Cf. ibid., p. 17.
[653] Cf. JANS, MIEKE (2009), p. 107.
[654] Cf. ibid.
[655] Cf. IFAC (2021), ISA 240, para. 33(c); JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2012), p. 17; JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), p. 25.
[656] Cf. CHUI, TIFFANY (2018), p. 24.

performed for further investigation and on the means with which process mining might support this investigation are not provided.

Overall, existing approaches for evaluating individual process paths do not represent an appropriate instrument for addressing the practical challenges in dealing with the multitude of resulting process variations. The examples derived from related literature suggest that either a more feasible approach for evaluating variations is required, or, alternatively, an approach for understanding the processing of transactions and assessing related risks of material misstatement that is not relying on individually examining the extensive number of resulting process variations.

Considering the materiality of individual process flows

Current approaches in audit-related process mining literature focus on qualitative criteria to evaluate variations, for example, by analyzing if the sequence of activities complies with the auditor's understanding of the process, if control activities relevant to the audit have been performed and if segregation of duties between these activities is maintained.[657] The only quantitative criterium incorporated in the majority of today's process mining solutions is the number of cases within a variation that is used to distinguish routine and non-routine transaction flows.

As an audit of financial statements is performed considering the materiality of potential misstatements,[658] it seems natural to apply the concept of materiality to determine when a particular process variation deviating from the routine process path requires further investigation. An individual execution of the process, i.e., a case within the process mining analyzer, may involve multiple transactions posted to the financial accounts. However, as current process mining techniques focus solely on the process perspective and do not consider the financial aspect of processes,[659] a process execution itself does not provide insights into materiality.[660] In discussing the relation between a case and materiality, JANS concludes that additional information needs to be added to the process, including the period covered by the variations and the monetary amount related to the cases, which might require developing new algorithms capable of joining the different perspectives on the data.[661]

[657] Cf. CHUI, TIFFANY (2018), pp. 17ff.
[658] Cf. ELDER, RANDAL J. et al. (2020), p. 210.
[659] Cf. Chapter 4.2.1.
[660] Cf. MIEKE, JANS (2012), p. 2.
[661] Cf. ibid.

In a first approach to quantify a variation in terms of risk, besides the number of cases, the total invoice amount related to the cases in a variation is integrated into the process mining application introduced in Chapter 3.2.1. The invoice amount related to a case provides an indication of the monetary value of the transactions involved in the case. While information on the number of cases supports identifying the critical path of the process, the invoice amount enables identifying non-routine process paths used for processing high-volume transactions that may be subject to increased risks of material misstatement. However, the empirical evaluation indicates that the use of the invoice amount for reducing the extent of variations to investigate in detail is limited. The jointly investigation of the number of cases and invoice amount processed over the process variations rather increases the number of variations identified for a detailed evaluation. For example, for five of the engagements presented in Table 24, the most frequent process variation covers between 20 and 60 percent of all cases in the analyzer, but only less than one percent of the total invoice amount processed in the audit period. In parallel, hundreds of seldom-used process paths may be identified that are used to process a small number of high-volume invoices. As a result, despite the information on the number and volume of cases within each variation, the auditor needs to apply professional judgment in determining which variations to investigate in detail.

In search for a fit between the process perspective and materiality, JANS raises the question if the materiality of a process execution may be determined based on a certain threshold of cases following the variation, or a threshold on the monetary amount involved.[662] Similar feedback is provided by some audit teams challenging if there is a "process materiality" that may be applied to the number of volume of cases following a process trace.[663]

As the document amount of the invoice line items related to a case is stored as a case attribute and not linked to the individual events of the case, it may only provide a broad indication of the monetary amount that is posted to the financial accounts. Consequently, the connection to the journal entries recorded in the financial statements, and thus the materiality thresholds determined for the related accounts, may not be established.[664] However, with the integration of the process data and the financial data proposed in Chapter 4.2 of this thesis, a connection between the cases and events in

[662] Cf. MIEKE, JANS (2012), p. 2.
[663] Cf. Chapter 3.2.2.
[664] Cf. Chapter 4.2.3.

the process data on the one hand and the related transactions leading to debit and credit movements on the financial accounts on the other hand may be established. As such, the monetary impact of a process variation on each related financial account can be quantified. Consequently, the concept of materiality may be applied without the need of a dedicated "process materiality" for the quantitative evaluation of individual process executions.

Qualitative assessment of process variations

Identifying significant process variations based on the impact on the financial statements supports the auditor in focusing the audit procedures on those variations that may lead to a material misstatement. However, provided the extensive number of variations resulting in practice, besides the volume of transactions following a specific process path, qualitative aspects need to be considered.

For example, if the auditor identifies that the activities performed within a process variation that is material to the financial statements are in line with the understanding of the process and obtained evidence on the design and operating effectiveness of integrated controls, the auditor may determine that there is no remaining risk of material misstatement related to the variation. On the other hand, as the primary risk in purchasing is the completeness of purchases throughout the year and the completeness of liabilities as of period end, a variation including cases that are not (significantly) affecting the financial statements in the audit period under review might still indicate a risk of material misstatement if transactions that should have been recorded are not recorded. While the concept of materiality may support the auditor in determining the magnitude of misstatements, i.e., the impact of any non-compliant process executions identified on the financial accounts, the question on the extent of variations to investigate is yet to be answered.

To support the auditor in performing the audit efficient and effective, i.e., to avoid both over- and under-auditing of individual process paths, in the following, different approaches to enhance or replace the manual evaluation of distinct process variations are discussed. Mechanisms to facilitate a qualitative assessment of different process flows may include:

- identifying and addressing risks by analyzing sub-populations of transactions,
- altering the definition of a variation based on the objective of the audit procedure and procedures already performed based on the entire population of cases,

- grouping variations by selected criteria, for example, shared case attributes,
- using statistical algorithms to cluster similar variations and
- using automated techniques to explore the root-cause of variations.

Addressing risks by analyzing sub-populations of transactions

Many researchers use the variation analysis as a starting point for identifying unusual process executions that may indicate a risk of material misstatement and require further investigation by the auditor. However, to meet the objective of many of the procedures performed, information on the process variations is not necessary to begin with. Instead of an isolated investigation of the individual process variations, the procedures may be performed jointly for the entire population of cases. In the following, different scenarios are discussed where the auditor may determine to resign from the manual variation analysis or put it behind the analyses based on the population of cases.

In the example described by the AICPA, the auditor obtains an understanding of each individual process variation and identifies instances where segregation of duties is not maintained.[665] By performing inquiries and further procedures, the auditor identifies a risk of material misstatement related to cutoff, as some of the affected cases relate to transactions recorded near the end of the audit period. As a result, the extent of cutoff procedures planned near period end is increased.[666] Instead of identifying cases where segregation of duties is not maintained as part of individually analyzing each process variation, the segregation of incompatible duties may be audited based on the entire population of cases.[667] For example, if the auditor expects that invoices are not approved by the same user creating the purchase order, he or she may filter the application to the cases where segregation of duties between these activities is not maintained. In a next step, the auditor may perform further analyses to identify any additional control activities that may have been performed to mitigate the related risk of material misstatement, for example, by identifying cases for which an additional invoice approval has been performed by a different authorized individual. Only in the last step, the individual process variations related to the remaining cases subject to the related

[665] Cf. AICPA (2017), para. A.72 and A.66.
[666] Cf. ibid., para. A.72.
[667] Cf. Chapter 2.3.3.

risk of material misstatement may be flagged for further investigation, including determining if the related transactions are material to the financial statements.

Similarly, by analyzing throughput times in the process, the auditor may determine that after the approval of the purchase order, goods are received within the usual cycle time in the process. By analyzing the purchase order amount related to cases where no goods or invoices have been received and no payments have been made yet,[668] the auditor may further determine that no risk of material misstatement is related to these transactions. Consequently, the auditor may use the process graph to exclude all transactions from the analysis that are not processed beyond the level of the purchase order, i.e., that only include activities related to the creation, change, approval or rejection of a purchase order, reducing the remaining number of process variations to review and evaluate.

As process variations represent the distinct ways the process is executed, the focus of the variation analysis usually is determining whether the sequence of activities complies with the auditor's understanding of the process, i.e., whether the cases start with the expected initiating activity and whether the order of activities or any missing activities give rise to a material misstatement. Based again on the population of cases, the ordering of activities within the process may be analyzed as follows:

- The analysis of the initiating and recording activities within the process[669] may be used to identify cases that do not start with the expected initiating activity.
- Missing activities in individual process traces may be identified by filtering the overall process graph for cases that do not pass through a specific key activity that is expected to be performed in the process. If a mitigating activity has been identified, for example, because the missing activity is redundant to another activity in the process, the auditor may further filter the data to cases where the mitigating activity has not been performed as well.
- In identifying inherent risks at the assertion level and related control activities addressing these risks, the auditor obtains an understanding of sequences of activities that are non-conforming with the entities policies and procedures ensuring the appropriate recording of transactions within the process. For example, in purchase to pay, the auditor may expect that (a) purchase orders and any

[668] Cf. Chapter 4.2.5.
[669] Cf. Chapter 3.3.1.

subsequent changes of the purchase orders are approved before the purchase order is released to the vendor in order to prevent fictitious transactions from being recorded, (b) no goods receipt, invoice or payment activities are performed for purchase orders that have been rejected and (c) payment blocks applied for cases with quantity or price disputes identified in the three-way-match need to be resolved by change events before any payment-related activity is performed, preventing, or detecting and correcting, the recording of transactions with incorrect details from the underlying documents. The "Activity sequence" dashboard[670] may be used to identify cases in the population with a pattern of activities that does not correspond to the auditor's understanding of the processing of transactions in accordance with applicable GAAP.

As such, the auditor may perform procedures related to the (non-)occurrence of individual activities and the sequence of activities within the process based on the total population of cases rather than duplicating the same procedures by individually analyzing each variation.

The examples demonstrate how the auditor may adjust the analysis to the identified risks of material misstatement and the objective of the audit procedure. Starting from the unfiltered process graph visualizing all variations in the process, subgroups of cases that require further investigation are identified. Filtering these cases enables the auditor to limit the analysis of individual process variations to those cases identified as giving rise to a risk of material misstatement. As the analysis is considering the entire population of cases, and the cases may be linked to the related journal entries posted to the financial accounts, the auditor may further reduce the evaluation of variations to those process traces that are material to the financial statements, significantly reducing the number of variations to investigate in detail.

Reducing complexity by altering the definition of a variation

A functionality for modifying the definition of a process variation may support the auditor in considering the evidence already obtained through other audit procedures when investigating the individual process variations.

The definition of a variation may be adjusted, for example, by:

- excluding certain activities from the construction of process variations,

[670] Cf. Appendix III.

- determining if direct repetitions of individual activities should be considered and
- determining if variations are created based on individual activities or based on the activity types.

In determining the activities to consider in the event log, some audit teams tend to include as much information as possible in the process graph.[671] As a result, the analyzer may include activities that are not relevant for performing the planned audit procedures but are redundant to other information provided in the analyzer or provide additional information identified as relevant for the entity but not for the auditor.[672] Consequently, the auditor may determine to not consider these activities the variation building algorithm. On the other hand, the auditor may also decide to exclude key control activities already covered by other audit procedures within the analyzer. For example, the auditor may identify the payment block activities related to the three-way-match as control activities relevant to the audit. However, if the appropriate use of payment blocks for disputes identified in the three-way-match has been confirmed, the auditor may not be interested in analyzing process variations resulting from a different ordering of payment block related activities in the process and exclude these activities from the construction of process variations right from the start.

The auditor may further decide to not consider repetitions of certain activities, i.e., cases where the same activity is performed twice without any other activity being executed in between. For example, the two process traces (1) “purchase order created, purchase order changed, purchase order approved” and (2) “purchase order created, purchase order changed, purchase order changed, purchase order approved” would result in two different process variations. If the auditor confirmed that changes made to purchase orders are approved by an authorized individual before the purchase order is released and segregation of incompatible duties is maintained, he or she will not be interested in process variations solely resulting from different numbers of consecutive purchase order changes. Similarly, whether consecutive goods receipts entered for a case represent duplicate entries of the same delivery or individual entries for distinct partial deliveries related to the purchase order line item is usually identified as part of the three-way-match[673] between purchase order, goods receipt and invoice. These

[671] Cf. Chapter 4.1.3.
[672] Refer to the analysis of the set of activities in Table 7.
[673] Cf. Chapter 3.2.1.

examples demonstrate how the risks of material misstatement related to repetitions of specific activities may be audited jointly for the entire population of cases instead of investigating each variation individually. At the same time, by disregarding consecutive repetitions of the same activities in constructing the process variations, the number of variations may be reduced significantly.

Lastly, the auditor may determine to summarize activities based on their activity type before creating the process variations. This may especially be relevant for datasets including a large number of activities. For example, the activities "purchase order approved (hierarchy 1)" and "purchase order approved (hierarchy 2)" refer to the approval of a purchase order on different authorization levels and may be assigned to the activity type "purchase order approval". Provided the auditor confirmed that purchase orders are approved by an authorized individual, he or she may decide to exclude the information on the hierarchy level in determining if a process execution confirms with his or her understanding of the entity's business.

Considering the results of the audit procedures already performed in the definition of a process variation may reduce process complexity by significantly decreasing the number of resulting process variations.

Grouping variations by selected criteria

When asked about missing process mining functionalities, one audit team fed back that when analyzing variations, their primary focus is on the activities that are occurring or not occurring in a specific process trace. Provided the sequence of activities relevant to the audit has been evaluated based on the entire population of cases as discussed earlier in this chapter, a functionality for grouping all variations that contain the same set of activities into a variation group (while disregarding the order of activities) may significantly reduce the number of variations. Instead of displaying the individual process traces, the process graph of a variation group displays all individual variations within the group for jointly investigation.

Instead of disregarding repetitions of specific activities in the variation definition, the grouping functionality may further be used to group the variations by disregarding these repetitions, i.e., "loops" within the process. For example, in this scenario, all cases containing a purchase order change or two or more subsequent purchase order changes will be summarized in a variation group, provided the remaining sequence of activities is the same.

Variations may be grouped by many additional criteria, in particular considering information obtained from other procedures performed within the process mining analyzer. For example, based on the identified risks of material misstatement, variations may be grouped according to the group of users involved in processing related cases, by the materials or material types that are purchased or by a specific vendor. The complexity of the variation analysis may be further reduced by combining the grouping functionality with other approaches for the qualitative evaluation of variations, for example, when not considering specific activities in the definition of a variation.

Every pilot team faced at least a small number of cases including a comparatively large number of events. Due to the multitude of events involved, each of these cases usually has a unique process trace, which may result in a large number of variations only including a single case. For example, the purchase to pay process of a large consumer goods company based in Switzerland includes 1.954 process variations with a total number of 15.192 cases. While the first five variations already cover 50 percent of the cases, there are 1.341 process variations (69 percent of all variations) only including a single case. An appropriate instrument to assess the risks of material misstatement that may be related to these process paths may reduce the number of variations to investigate significantly.

Infrequent paths of the process may represent overly complex and non-routine transactions, for example, cases subject to accounting practices including a large number of corrections and reversals. The information on the total transaction volume related to variations supports the auditor in identifying those outliers involving a risk of material misstatement for further investigation.

Frequently, cases with a large number of events include long-standing cases spanning over multiple years. Long-standing cases in purchase to pay may occur, for example, if frame agreements exist with specific suppliers and the related transactions use the same purchase order reference.[674] As the cases span over a large period of time, they usually have a unique sequence of activities. However, besides using the same purchase order reference and thus the same case identifier, transactions posted within the frame agreement are usually subject to the same process than other transactions. As such, although representing a unique process trace by themselves, long-standing cases

[674] For the consumer goods company, 829 of the variations with only a single case have been created in a period preceding the audit period and were subject to frame agreements with various suppliers.

frequently contain common patterns of the process, i.e., they loop over the most frequent variations in the process. As cases related to long-standing purchase orders may be identified by the timestamp of their first activity, the auditor may investigate these cases in a separate analysis, reducing the number of variations to examine in detail. For example, the auditor may summarize all long-standing cases (i.e., purchase order line items) by their purchase order document, the related vendor or the materials involved to understand the business reason behind the transactions and corroborate the processing with information obtained on existing frame agreements or similar contracts.

Grouping infrequent variations based on their characteristics for a jointly investigation may reduce the number of variations significantly. Considering the materiality of the related amounts posted to the financial statements helps the auditor in determining which variation groups to investigate in detail.

Using statistical algorithms to cluster similar variations

Statistical algorithms may be used to help the auditor in handling the complexity of a process graph by automatically grouping process variations that are similar to each other into a cluster. Clustering is an established technique in statistics used to reduce a problem's complexity. In the context of process mining, clustering techniques are referred to as "trace clustering", describing the grouping of process traces into homogeneous subsets that are easier to understand and evaluate than the overall process model.[675]

A prerequisite for applying clustering algorithms is to determine the similarity of the items to be clustered using a distance measure. In the context of process mining, an item is an individual process trace that is followed by at least one case in the event log. The similarity of cases may be determined, for example, based on the edges in the process that represent the transition from one activity to another, as many edges lead to a complex structure of the overall process. Additional information that may be incorporated include the case attributes of different process traces, the event attributes and the duration of the process execution.[676] After determining the influence factors, the similarity of cases may be calculated using distance measures. There are many

[675] Cf. NEUBAUER, THAIS R./FANTINATO, MARCELO F./PERES, SARAJANE M. (2019), p. 45; SONG, MINSEOK/GÜNTHER, C. W./VAN DER AALST, WIL M. P. (2008), p. 109.

[676] Cf. LEONARDI, GIORGIO et al. (2018), p. 15; SONG, MINSEOK/GÜNTHER, CHRISTIAN W./VAN DER AALST, WIL M. P. (2008), p. 114.

different distance measures available in literature which may be used individually or in combination.[677] In a final step, clustering algorithms may be used to cluster process traces with similar behavior (i.e., with a small distance) into a homogeneous subgroup. An overview of clustering algorithms is provided by BACHER/PÖGE/WENZIG.[678]

An early prototype for the clustering of similar process traces has been implemented in the audit firm's process mining application in parallel to the implementation described in Chapter 3. Appendix IV provides an overview of the clustering analysis and describes the key functionalities provided. The underlying dataset includes 1.554 process variations that are grouped into 175 clusters. As the clusters summarize variations whose cases are handled similar to each other, the auditor may evaluate multiple variations at the same time and may achieve a higher coverage of the overall number of cases (or monetary amount, respectively) faster than by analyzing the individual variations. In case a cluster is still very complex, the auditor can filter this cluster and switch back to the variation analysis to analyze any significant variations individually.

As clustering algorithms typically involve a trade-off between the number and the complexity of resulting clusters, the auditor is required to determine either the number of clusters or the desired complexity before clustering the data. However, this information is usually not known, as figuring out the number of homogeneous subgroups of process traces is one of the reasons to perform the analysis in the first place. Challenges have been further identified regarding the interpretation of the clusters from a business perspective. Instead of analyzing the resulting sub-graphs of the process, many auditors went back to analyzing the individual process variations or restricted the analysis of the clusters to the most frequent process traces within each cluster. These results conform with observations made in related literature on trace clustering. Frequently, although statistically correctly grouping the most similar variations, the clusters derived from the data fail to support the understanding of the process from a business perspective.[679] Recent literature discusses the concept of an "interactive" trace clustering including the involvement of expert knowledge in determining factors

[677] Cf. NEUBAUER, THAIS R./FANTINATO, MARCELO/PERES, SARAJANE M. (2019), p. 47; ADRIANSYAH, ARYA/VAN DONGEN, BOUDEWIJN. F./VAN DER AALST, WIL M. P. (2011a), p. 57; SONG, MINSEOK/GÜNTHER, CHRISTIAN W./VAN DER AALST, WIL M. P. (2008), p. 115.

[678] Cf. BACHER, JOHANN/PÖGE, ANDREAS/WENZIG, KNUT (2010), pp. 18ff.

[679] Cf. NEUBAUER, THAIS R./FANTINATO, MARCELO F./PERES, SARAJANE M. (2019), p. 46.

influencing the similarity of process traces to facilitate later interpretation.[680] However, this approach again involves detailed prior knowledge of the process before actually using process mining to understand the process in the first place. If the auditor already knew the criteria differentiating subgroups of process variations to begin with, it is questionable if a clustering functionality would be needed anymore. Instead, the auditor could simply group the variations by the identified criteria for a jointly analysis as described earlier in the chapter.

Reducing the complexity of the process by automatically grouping traces with a similar behavior into a sub-graph may be a valuable instrument to facilitate both risk assessment procedures and responding to the identified risks. However, provided the challenges identified in determining the similarity of process executions and interpreting the resulting clusters, further research is required for the concept of trace clustering to be superior to other means of evaluating the different sequences of activities.

Using automated techniques to explore the root-cause of variations

While the most frequent variation helps the auditor in understanding the critical path of the process, for the remainder of the variations the auditor is primarily interested in whether deviations from the routine process path give rise to a risk of material misstatement of the financial statements. The approaches to qualitatively evaluate process variations discussed in this chapter, together with the consideration of the variations' materiality to the financial statements, may support the auditor in handling the multitude of resulting process variations while focusing on related risks of material misstatement.

In executing a risk-based audit approach, the auditor is not particularly interested in each individual process variation but in their root-causes that may give rise to a risk of material misstatement. For example, purchase orders that are processed differently whenever a specific vendor and user is involved may give rise to a risk of material misstatement due to fraud. In exploring the root-cause of different process variations, the auditor would significantly benefit from an instrument for identifying the reasons for a case to use an alternative path through the process than other cases.

---

[680] Cf. NEUBAUER, THAIS R./FANTINATO, MARCELO F./PERES, SARAJANE M. (2019), p. 46.

Approaches to analyze process variations proposed in literature but not applied to the context of auditing so far include the model delta analysis, conformance checking techniques, sequence classification and the automated log delta analysis.

Model delta analysis

The model delta analysis uses two distinct event logs of positive and negative cases. A case is classified as positive if it meets specific performance objectives, for example, the treatments of a patient within a specific timeframe. The technique builds a process model separately for the positive and negative cases and supports a visual comparison of the two process graphs to identify distinguishing patterns that may explain the differences in performance.[681] When transferred to the context of auditing, negative cases would correspond to cases with a sequence of activities that gives rise to a risk of material misstatement. However, identifying if these process traces exist is subject to the variation analysis, i.e., to classify cases or variations as positive or negative, they need to be analyzed individually. Further, the manual comparison of the resulting process graphs does not scale up to complex processes and is prone to errors.[682] Consequently, the model delta analysis does not help in an auditing context.

Conformance checking techniques

Conformance checking techniques[683] aim to explain the conformance of an individual process trace with a normative process model.[684] However, one of the main benefits of using process mining in the audit is the potential of the solution to replace the traditional means of deriving a flowchart of the designed process model through inquiry, inspection and observation. In addition, the degree of conformance of a process trace does not provide information whether the discrepancy identified may give rise to a risk of material misstatement. As such, the suitability of conformance checking techniques to facilitate evaluating process variations in an audit of financial statements is limited.

Sequence classification

Sequence classification methods build a classifier (for example, a decision tree) that may be used to determine with a certain accuracy if a process trace belongs to the positive or the negative subgroup of cases.[685] The methods may be categorized into

[681] Cf. van Beest, Nick R. T. P. et al. (2015), p. 386.
[682] Cf. ibid.
[683] Cf. Chapter 2.1.
[684] Cf. van der Aalst, Wil M. P. (2016), p. 33.
[685] Cf. van Beest, Nick R. T. P. et al. (2015), p. 388.

(1) techniques for activity-based feature encoding, (2) frequent sequence mining techniques and (3) discriminative sequence mining techniques.[686] Techniques for activity-based feature encoding transform each process trace into a vector containing information on the frequency of each activity that is part of the event log. A classifier is applied on the vectors to distinguish positive and negative cases.[687] Frequent sequence mining techniques extract sequences of activities from the set of positive cases and negative cases separately. The frequencies of the patterns identified are used to construct a classifier to distinguish positive and negative cases.[688] Discriminative sequence mining techniques are comparable to frequent sequence mining techniques but consider the discriminative power of a pattern, i.e., a sequence of activities is selected only if it is frequent for one subgroup of cases but does not occur in the other.[689]

A detailed evaluation of existing sequence classification methods is provided by NGUYEN et al.[690] The authors conclude that the accuracy obtained with sequence classification methods is rather limited and the obtained sets of rules where overly complex.[691]

Automated log-delta analysis

The automated log delta analysis aims to explain differences in the behavior of cases within two different event logs or two different process variations.[692] The technique produces a list of statements in natural language describing present or frequent behavior in one event log or variation that is absent or infrequent in the other.[693] For example, such difference statements may include: "In event log 1, activity A is always executed but in event log 2, activity A is sometimes skipped"[694], or "In variant 1, activity B occurs after activity C, while in variant 2, it does not occur".[695] VAN BEEST et al. demonstrate that the automated log delta analysis requires significantly less statements to explain the differences between two event logs than the sequence classification methods and furthermore points to the exact point in the process where a

[686] Cf. VAN BEEST, NICK R. T. P. et al. (2015), p. 388.
[687] Cf. ibid.; NGUYEN, HOANG HUY et al. (2014), p. 5.
[688] Cf. NGUYEN, HOANG HUY et al. (2014), p. 5.
[689] Cf. ibid.; VAN BEEST, NICK R. T. P. et al. (2015), p. 388.
[690] Cf. NGUYEN, HOANG HUY et al. (2014), pp. 2ff.
[691] Cf. ibid., p. 16; DUMAS, MARLON et al. (2018), p. 459; VAN BEEST, NICK R. T. P. et al. (2015), p. 388.
[692] Cf. VAN BEEST, NICK R. T. P. et al. (2015), p. 386.
[693] Cf. ibid.
[694] Cf. DUMAS, MARLON et al. (2018), p. 459.
[695] Cf. VAN BEEST, NICK R. T. P. et al. (2015), p. 401. The authors further provide an explanatory list of statements derived from cases of a patient treatment process, cf. ibid.

behavioral difference occurs.[696] From the list of statements it may be determined which pattern or combination of patterns is capable to explain the behavioral differences between the subsets of cases by manually inspecting a sample of cases.[697] However, again, the usefulness for exploring the root-cause of variations in the context of auditing is limited. Even if applied to individual process variations without pre-classifying the event log into positive and negative cases, the auditor is required to review an extensive list of statements describing the behavior within the different subgroups of cases and this list increases with the number of variations.

In summary, the existing semi-automated or automated approaches of evaluating process variations are not suitable in the context of auditing. Further theoretical and empirical research is required with regard to an explorative approach that is capable to support the auditor in effectively identifying business rules, control activities or other characteristics of the process that determine the necessary conditions for a case to follow a certain process path. Based on the root-causes identified, the auditor may determine if they, individually or in combination, give rise to a risk of material misstatement without manually analyzing the resulting process variations. Until then, the proposed approach to evaluate process variations both qualitatively and quantitatively may be applied.

4.4 Lacking efficiency of process mining when applied in multiple periods

4.4.1 Practical challenges in comparing data with process mining over time

As many companies decide to use the end of the calendar year as their balance sheet date, including entities belonging to a group with tight reporting deadlines after financial statement close, auditors usually perform as many procedures as feasible before this date, i.e., based on interim data. In practice, this especially includes the procedures to understand the entity and its environment, to identify and assess risks of material misstatement and to evaluate the design and operating effectiveness of controls. At period end, the procedures performed at an interim date are updated to obtain audit evidence on the time period between the interim date and the balance sheet date.[698] As a result, auditors frequently compare information at different points in time. Besides

[696] Cf. van Beest, Nick R. T. P. et al. (2015), pp. 400 and 403.
[697] Cf. Dumas, Marlon et al. (2018), p. 459.
[698] Cf. IFAC (2021), ISA 330, para. 12 and 22.

the comparison of period end and interim data, analytical procedures[699] usually involve the comparison of the entity's financial or non-financial information with the previous audit period.[700]

However, current process mining techniques do not support comparing the information of the audit period with a comparative period directly within a single process mining application. As the data extraction period can be flexibly defined, theoretically, two different process mining applications may be set up for the audit period and the comparative period to support roll-forward procedures and analytical procedures. In this scenario, the auditor needs to reperform the process mining procedures at period end and compare the results with the analyzer set-up at the interim date or for the previous audit period.[701] However, there are several practical challenges related to this approach preventing a meaningful comparison of data between different process mining applications.

Process mining solutions do not exclusively consider the data from the audit period under review. The length of the period and the extent of data points included in the analyzer are depending on the definition of the data extraction strategy. For cases that have an event in the audit period under review, the process mining solution introduced in Chapter 3 considers events from up to four years before the balance sheet date and depending on the data extraction date, the analyzer may also include events that have been performed after the balance sheet date.[702] Similar configurations may be observed in the process mining research, for example, when the investigation only includes those process instances that lead to an invoice in the observation period.[703] Consequently, the definition of the data extraction period directly impacts the metrics and analyses within the process mining analyzer.

Company	Process	Case(s)	Company code	Local currency	Users
ABC Company	PTP	Invoice - With PO Invoice - Without PO	CH80	CHF	85
Purchase order amount	Invoice amount (CY)	Cases	Activities	Events	Variations
110,955,392	83,991,852.51	15,192	23	123,676	1,954

Figure 80: Key indicators of the process provided in the process mining application

---

699 Cf. IFAC (2021), ISA 520, para. 1; ELDER, RANDAL J. et al. (2020), pp. 167f.; WIESE, MICHAEL (2013), pp. 21f. Refer to Chapter 2.3 for the definition and purpose of analytical procedures.
700 Cf. IFAC (2021), ISA 520, para. A1.
701 Cf. Chapter 2.3.4.
702 Details on the data extraction strategy are provided in Chapter 4.1.6.
703 Cf. CHIU, TIFFANY (2018), p. 18.

For example, the total number of events and the number of activities summarized in Figure 80 are based on all events that are included in the event log, regardless of whether they have been performed in the period under review or not. Similarly, as the users performing a particular activity are stored as event attribute, any user related metrics are considering the entirety of events in the dataset. The total purchase order amount is based on the total value of the purchase orders relating to cases that have been processed throughout the audit period. As some of these purchase order documents might have been created prior to the audit period, both the "cases" and the "purchase order amount" metric are not limited to the audit period but considering the entire data extraction period. To enable reconciliation to the financial data, however, the invoice amount metric related to the process instances only considers the invoice amount that has been recorded in the audit period.[704] Consequently, there is a discrepancy in the calculation of different metrics and key indicators related to the process instances. These examples demonstrate that when designing analyses and audit procedures with process mining, careful consideration needs to be devoted to the definition of metrics and the presentation of information. The discrepancies between the data extraction period and the period under review especially pose challenges when comparing the data for different periods across multiple process mining analyzers. Cases that are processed in the audit period but have been initiated in a previous period are included in both analyzers, preventing a direct comparison of most of the information and metrics provided.

Despite the challenges in separating individual data points between different time periods, manually comparing information between two different process mining analyzers is a tedious task. For example, changes in the users involved in the process and in their individual responsibilities need to be identified by a manual comparison of related information between two different applications. Figure 81 shows the users approving a purchase order in the audit period on the left side and in the previous period on the right side.

[704] In Figure 80, the limitation of the invoice amount metric to the audit period is indicated by the label "invoice amount (CY)".

User ID	Events	%
GNO	3,082	39.01%
JCT	2,264	28.66%
HVN	1,345	17.03%
OFV	1,011	12.8%
OJI	132	1.67%
PAQ	61	0.77%
CPF	5	0.06%
Total	7,900	100%

User ID	Events	%
OJI	3,628	55.41%
GNO	1,521	23.23%
HVN	903	13.79%
OFV	281	4.29%
PAQ	214	3.27%
Total	6,547	100%

Figure 81: Manual comparison of users approving a purchase order in the current (left) and previous audit period (right)

Comparing both tables shows that two of the users in the previous period ("JCT" and "CPF") are no longer approving purchase orders in the audit period, that there are no new users involved in the purchase order approval process and that the involvement of user "OJI" significantly increased[705] in the audit period. However, as the data sources are not integrated, comparing the information and documenting the results of the comparison takes a considerable amount of time that, in the example provided, increases with the number of users and activities in the process. To overcome this limitation of a manual comparison, some audit teams exported the data from both process mining analyzers into Microsoft Excel and performed the actual comparison of the data outside of process mining. Still, both approaches significantly reduce audit efficiency.

Further, current process mining solutions do not support comparing the number and volume of cases following the same process trace at different points in time. Frequently, the numbering of variations is based on the number of cases following the process trace, starting with the most frequent variation 1. By this, individual variation IDs are not comparable between different process mining applications. Without investigating the individual process traces, the auditor may not determine, for example, if (a) the sequence of activities within a particular process variation has already been audited in a prior period, if (b) the process variations with the highest number of cases or the highest invoice amount volumes did change from one period to another or if (c) variations related to a risk of material misstatement in the prior period or unusual variations included in the management letter do still occur in the audit period under

[705] As the process instance is a purchase order line item while a purchase order is approved on the level of the entire document, it cannot be determined with absolute assurance from the analysis in Figure 81 if the user "OJI" approved a larger number of purchase orders or if the related purchase orders in the audit period contain a larger number of line items, cf. Chapter 4.1.4.

review. Hence, comparing variations across periods may increase the efficiency of process mining as an audit instrument significantly.[706]

Due to the challenges in comparing data, many pilot teams used the process mining analyzer to support obtaining an understanding of the process and identifying and addressing risks of material misstatement at the interim date but updated the procedures at the balance-sheet date without using process mining. Overall, the practical evaluation confirms that the efficiency of using process mining in subsequent audit periods could be increased significantly if the analyzer supported a meaningful comparison of data between different periods. This includes both the comparison of the data from the audit period with the data from the previous period and the comparison of the period end data with the data already investigated at an interim date throughout the period.

4.4.2 Concept and integration of a comparison functionality for process data

To ensure that all process instances that are relevant to the audit period under review are included in the process mining analyzer, the data extraction period and the audit period inherently differ.[707] Consequently, not all information provided in the process mining analyzer is limited to the current audit period. Hence, to compare the data between multiple periods, the definition of counting and allocating specific data points to a particular period needs to be adjusted. Due to the different nature of the information, a different approach is required for different types of data. In the following, a concept to compare different types of data across multiple periods is developed. Considerations include the allocation and comparison of:

- financial information,
- case and event information,
- event attributes, such as the user performing the event,
- interactions between users,
- vendor information and
- process variations.

To facilitate the comparison, the data from both the audit period and the comparative period may be integrated into the same process mining application. Providing a comparison functionality in the global filter pane of the analyzer that enables to investigate

[706] Cf. Chapter 4.3. The relevance of comparing individual process traces between periods increases even more if process mining is used to test internal controls, cf. Chapter 4.5.

[707] Cf. Chapter 4.1.6.

the data from the audit period, the comparative period or both periods together may significantly increase the efficiency of the analyses.

Comparing financial information

The concept of comparing financial data in process mining does not differ from other automated tools and techniques using general ledger or subledger data. The transactions are allocated to the audit period or the comparative period based on the effective date of the related journal entries. In an aggregated view summarizing the data for the entire audit period, the information of the comparative period may be added as separate column to the analysis. Identifying significant changes related to financial data is usually supported by two columns summarizing the absolute and relative changes between both periods. Table 25 illustrates the comparison of the general ledger accounts' transaction movements between the audit period and the prior period.[708]

GL account type	GL account class	Current period (2018)	Prior period (2017)	Difference	% Difference
Assets	1000-Liquid Assets	-19,222	10,168	-29,389	-289%
	1100-Trade Receivable	-2,147,509	2,091,718	-4,239,227	-203%
	1200-Inventories	-4,036,888	-4,051,560	14,672	0%
	1300-Other Receivables	3,281,854	2,984,508	297,346	10%
	1400-Accrued Income / Prepaid Expense	-9,841	57,748	-67,589	-117%
	1600-Total Property, Plant And Equiptmen	-84,174	-58,179	-25,995	45%
	1700-Intangible Assets	82,729	-1,284	84,013	-6,542%
	1800-Deferred Tax Assets	0	-996,679	996,679	-100%
Equity	3100-Reserves	3,000,000	400,000	2,600,000	650%
	3200-Retained Earnings	-944,130	2,596,678	-3,540,808	-136%
Expenses	4000-Gross Margin I-Total (Exp)	70,032,709	74,482,990	-4,450,281	-6%
	4100-Selling Expenses	280,633	-205,141	485,774	-237%
	4200-Logistics & Distribution	1,685,884	1,704,911	-19,027	-1%

Table 25: Comparison of the general ledger account movements between two different periods

Similarly, if the data is disaggregated on a monthly basis, the analysis may be extended by the respective months from the comparative period.

Comparing case and event information

When investigating a single period, a purchase order line item is counted when it has an event in the audit period. To ensure completeness of the related process trace in the analyzer, all events related to this process instance are counted, regardless of whether they fall in the audit period under review. When the metrics related to a case are stored

[708] The analysis of the general ledger account movements without an integrated comparison functionality is introduced in Figure 27 of Chapter 3.2.1.

as case attribute only, the appropriate allocation of recording-relevant process information is solely made on the level of the case metrics. For example, as the invoice amounts in the case table are summarizing the total invoice amount related to these cases, for calculating the invoice amount relevant to the current audit period, a new metric "invoice amount (CY)" is introduced, that only considers the invoice amount posted in the current period.

To enable a comparison between different audit periods, these definitions need to be adjusted. To ensure completeness of the process model, a case is still allocated to each period is has an event in. As cases may have events in both the audit and the comparative period the concept of "active cases" is introduced to indicate that cases may overlap between both periods. Events, however, need to be allocated to the period they have been performed in. The integration of journal entry posting events[709] and subledger data[710] enables to store the metrics related to a case as an event attribute. Hence, a precise allocation of the related purchase order, goods receipt, invoice and payment amounts may be achieved automatically by separating the events by the period they have been performed in. Using the adjusted definition of cases and events relevant to a particular period, Appendix V summarizes the implementation of the comparison functionality for the analyses of the "Overview" dashboard of the analyzer.

Comparing event attributes, including user information

Event attributes such as the user performing the event are allocated to the period the event has been performed in. By this, process mining facilitates comparing events and related attributes between different periods. Instead of manually comparing changes in the user involvement between two separate applications as illustrated in Figure 81, the analysis may be integrated in a single process mining application. Figure 82 shows the users approving a purchase order in the audit period and the comparative period.

[709] Cf. Chapter 4.2.2.
[710] Cf. Chapter 4.2.3.

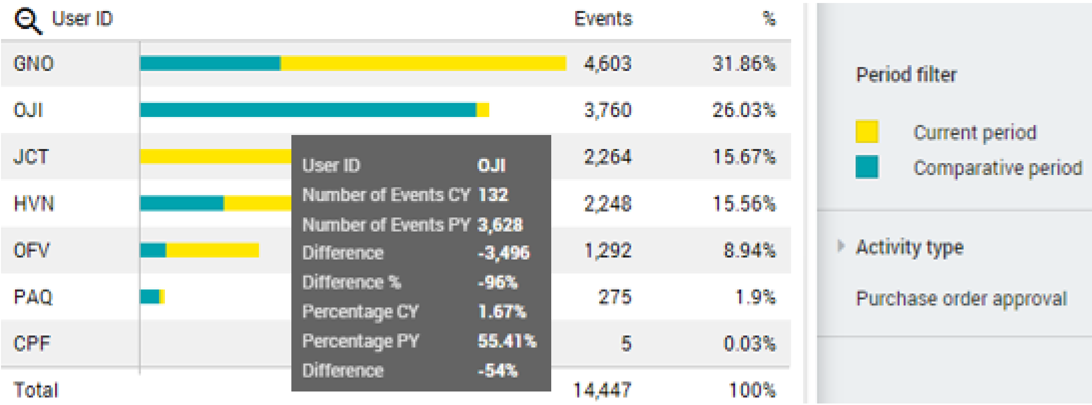


Figure 82: Integrated comparison of the users approving a purchase order in the audit period and the comparative period

Users approving a purchase order in the current period are colored yellow while the respective information from the comparative period is colored turquoise. When hovering a specific user, a tooltip summarizes the absolute and relative changes in the number of events performed, supporting analytical procedures.

Interactions between users

Using the user information in the event log, process mining supports social network analyses focusing on the collaboration between resources involved in the process.[711] A social network includes nodes representing the resources (such as users, roles or departments) and edges indicating the relationships between the users. A relationship between two users exists if a case is handed over from one user to another, for example, if the purchase order is created by user A and approved by user B. The weight, and thus, the thickness of the edges is determined by the number of interactions between the related users. Interactions may begin in one period and end in the next. That is, while the purchase order might be created and automatically handed over for approval in the audit period, the actual approval might be performed after the balance sheet date. Figure 83 illustrates the processing of five different cases during the period for which the data has been extracted.

[711] Cf. VAN DER AALST, WIL M. P. (2016), p. 33. To facilitate the initial adoption of process mining in the audit, in the implementation described in this thesis, the users' roles and responsibilities and appropriate segregation of incompatible duties are analyzed without the means of a social network analysis.

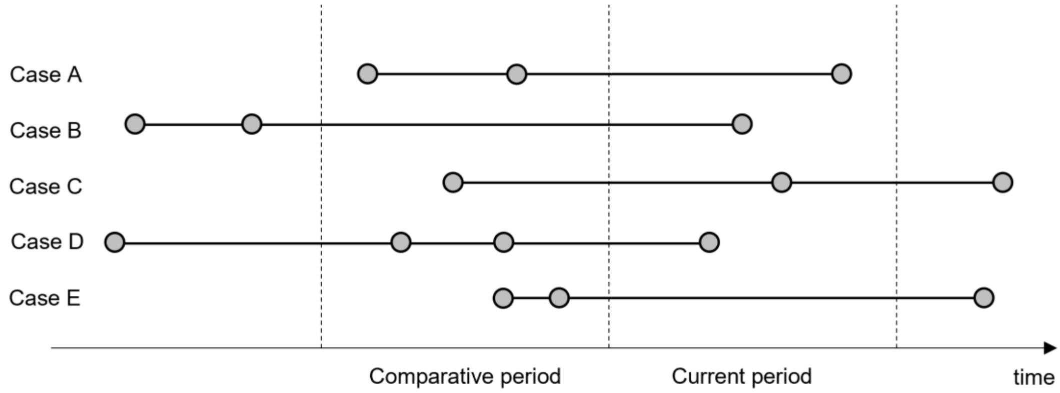


Figure 83: Collaboration between users processing transactions in the data extraction period

A gray circle represents an event performed by a particular user. Following the allocation of events and event attributes to different periods, an interaction between two events is considered as completed in the period the second event has been performed in. In the example presented, cases A, B, C and D have an interaction in the current period. Case E does not, as there is no event in the current audit period. Similarly, cases A and E have one interaction in the comparative period while case D has two interactions in the comparative period.

Comparing vendor information

When analyzing general ledger or subledger data, the auditor is used to identify changes (or the absence of changes) related to individual accounts using columns summarizing the data of both periods and the absolute and relative changes, respectively. As such, a multi-period comparison of the vendor related analyses described in Chapter 4.2.4 may be integrated similar to the comparison of the general ledger data illustrated in Table 25.

Frequently, automated tools and techniques support comparing information between multiple periods by providing visual presentations of the data. Figure 84 demonstrates how the data of the comparative period may be integrated in the analysis of the vendor structure introduced in Figure 78.

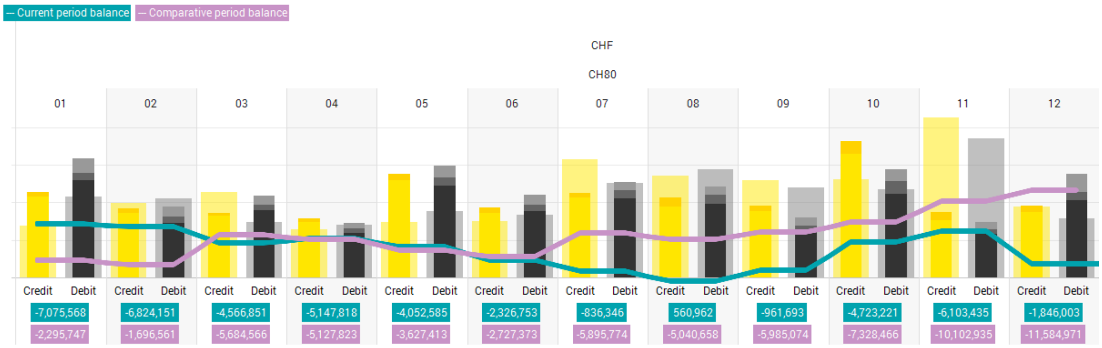


Figure 84: Comparison of the development of vendor balances in the audit period and the previous period

Similar to Figure 78, the columns display the transaction volumes debiting or crediting the vendor accounts in the audit period. The comparative information of the previous period is added as transparent columns in the background to support the auditor in identifying significant changes in the monthly development throughout both periods. Information on the monthly ending balances is provided in turquoise for the audit period and in purple for the comparative period.

When the financial data is combined with process information, for example, when displaying the number of cases and events by vendor, the considerations for allocating cases and events to the different periods apply. As the vendor is an event attribute, it is allocated to the period the event has been performed in.

Comparing process variations

Variations are typically chronologically numbered based on the number of cases they contain.[712] As the distribution of cases over the variations changes over time, the variation IDs change as well. Variation ID 15 may describe the process trace "purchase order created – purchase order approved – goods receipt posted" in the first audit period and the process trace "purchase order created – purchase order rejected – goods receipt posted" in the second audit period. With current process mining solutions, the auditor is not able to identify the variation ID used for the trace "purchase order created – purchase order approved – goods receipt posted" in the second audit period without clicking through the variations or manually filtering the process graph.

[712] Cf. Figure 34.

Figure 85 introduces a new concept for recording the trace of a process variation over time.

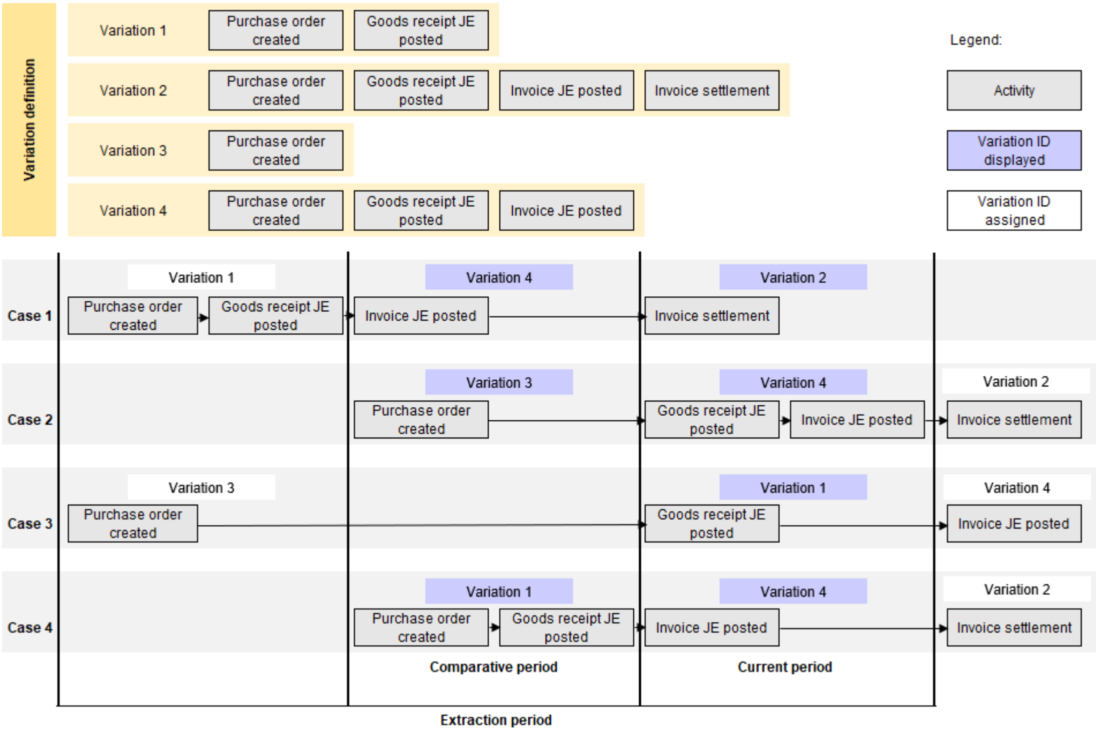


Figure 85: Definition of variations considering multiple periods

The numbering of variations is kept constant between periods, that is, the same sequence of activities has the same variation ID in each period. In the example provided, variation 1 describes the process trace "purchase order created – goods receipt JE posted". When assigning the variation ID to a specific process trace, only the events that are performed in the respective period or in prior periods are considered. In the example provided in Figure 85, case 1 will be included in variation 1 in the period before the comparative period. In the comparative period, the activity "invoice JE posted" is performed and thus, case 1 will be displayed under variation 4. When analyzing the data of the current audit period, the same case will be part of variation 2 as the invoice has been settled. The variations highlighted in purple show the variation IDs that are displayed in the analyzer, depending on the period selected.

This definition of variations implies that:

(1) The variation ID assigned to a case might change over time as the case is processed. Cases that have events in both the audit period and the comparative period will be assigned to a different variation ID in both periods.

(2) The same variation ID is used in all periods to describe the same sequence of activities.

The auditor is not particularly interested in the fluctuation of a specific case's variation but in the appropriateness of a specific process variation, i.e., a sequence of activities. As such, the proposed concept of defining variations may increase the efficiency of auditing process variations significantly over time. The auditor may still be interested in analyzing significant changes in the number or invoice amount volume of the cases processed in a specific variation. However, a variation ID that has already been audited in previous periods would not need to be investigated from a qualitative perspective anymore in the audit period and in subsequent periods.

Figure 86 shows the implementation of the concept by comparing the ten most frequent process traces in the audit period with the comparative period.

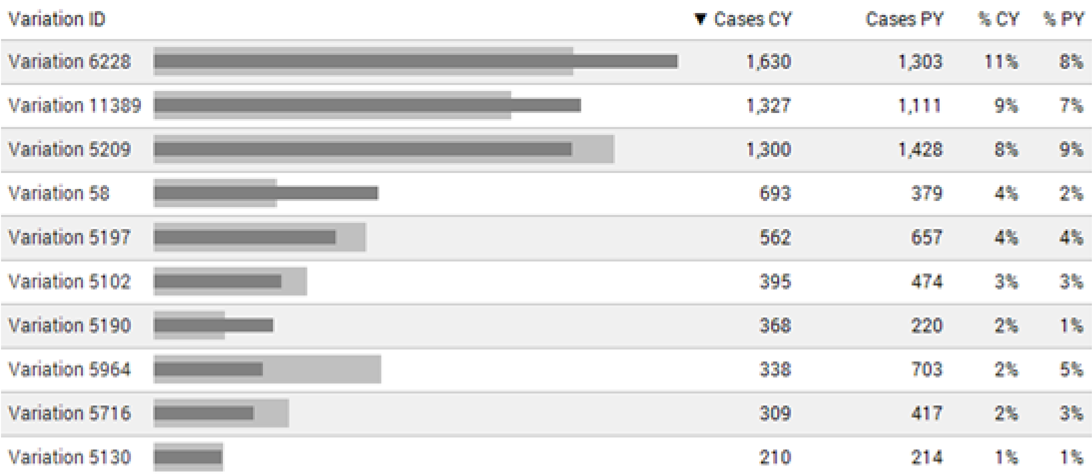


Figure 86: Comparing process traces between different periods

For each variation, information on the number of cases in the current and in the prior period is provided. A bar chart is used to visualize the number of cases in each period, where the comparative information of the prior period is displayed transparently behind the current period's information. The last two columns show the number of cases in a specific variation relative to the total number of cases in each period, respectively. In the example provided, there is a total number of 2.405 variations in the audit period and 2.263 variations in the prior period. Each variation ID describes the same process trace in the current and in the prior period.[713] By default, the variations are ordered descending by the number of cases in the audit period. Instead of analyzing the number

[713] The variation IDs in the dataset presented in Figure 86 range from 1 to 12709 while the total number of variations is 2.405 in the audit period and 2.263 in the comparative period. This may happen as the proposed concept involves assigning a new variation ID to each unique process trace, i.e., the variation IDs continuously increase with the number of new process variations that may be observed. In implementing the concept, a repository of unique process traces has been created considering multiple entities and observation periods, resulting in the significant discrepancy between the number of process variations and the maximum variation ID.

of cases, the metric may be adjusted to investigate the invoice amount processed in each variation and period. Assuming the variations have been investigated in the prior period from a qualitative perspective, the auditor may focus on analyzing significant changes in the number of cases or invoice volume processed in a specific variation. Sorting the data ascending by case number (or invoice amount, respectively) in the current or the prior period supports the auditor in identifying business practices that are new or do no longer occur in the current audit period.

Future implementations may provide advanced algorithms and analyses that support effectively and efficiently auditing these new process variations. The UiPath process mining solution used in the first feasibility assessment already included a dashboard for comparing different process variations.[714] However, due to the multitude of process variations resulting in practice, the analysis has not been used by the audit teams.[715] This may change with the integration of a constant variation ID for a given process trace, as the auditor may limit the analysis to variations that are new – either when compared to the variations of the previous audit period or to variations already evaluated at an interim date of the audit period. Process mining literature already includes a multitude of conformance checking techniques to align a process execution with the closest path of a model and quantify the conformance.[716] Besides the instruments for analyzing process variations discussed in Chapter 4.3.2, in the future, these algorithms may support the auditor in evaluating new variations by identifying the most similar process execution among those variations already investigated in previous audit periods.

The proposed concept of enhancing process mining with a functionality to compare data between two audit periods may not only facilitate analytical procedures but significantly increase the efficiency of using the technology in an audit of financial statements. Although the described implementation focuses on the comparison with the previous audit period, the concept may be adjusted to further account for the comparison of data between the interim date and period end. In this scenario, the adoption of process mining in an audit of financial statements may further benefit from an option to automatically reperform the analyses performed at the interim date for the

[714] Cf. Figure 33.
[715] Cf. Chapter 3.2.1.
[716] Cf. VOM BROCKE, JAN et al. (2021), p. 484; DE LEONI, MASSIMILIANO/VAN DER AALST, WIL M. P. (2013), p. 3; Chapter 2.2.

roll-forward period until period end, with a summary highlighting changes (or the absence of changes), for example, new process variations or open items that have been cleared throughout the roll-forward period.

4.5 Process mining for testing internal controls over financial reporting

4.5.1 Developments in process mining theory and audit practice related to internal control

Financial Market Integrity Strengthening Act reinforcing the importance of internal control

As a reaction to the Wirecard scandal in Germany and based on the recommendation of the Federal Ministry of Finance and the Federal Ministry of Justice and Consumer Protection,[717] the German Bundestag established the Financial Market Integrity Strengthening Act (Finanzmarktintegritätsstärkungsgesetz, FISG),[718] being effective since July 2021.[719] The FISG's objective is to restore and permanently strengthen the confidence in the German financial market by improving financial statement control.[720] To counteract financial statement fraud, the FISG focuses specifically on functions entrusted with the preparation and audit of financial statements, including the areas of corporate governance, external auditing and the enforcement procedures of the Federal Financial Supervisory Authority (Bundesanstalt für Finanzdienstleistungsaufsicht, BaFin).

Of particular interest in the light of this thesis are the new requirements for the boards of directors of entities listed on the stock exchange to establish an appropriate and effective system of internal control and a risk management system designed specifically to fit the individual business activity and risk environment situation of the entity.[721] Before the FISG was in effect it was the discretion of the board of directors to decide if and how internal controls and a risk management system are implemented, as there was no obligation to follow the respective recommendation[722] of the German

[717] Cf. BMF/BMJ (2020); BMJ (2021).
[718] Cf. BUNDESREGIERUNG (2021).
[719] Cf. ibid., article 27, para. 1 FISG. As the law has been established in short cause after the Wirecard scandal it includes various transitional provisions that became effective as of 1 January 2022, cf. ibid., article 27, para. 2 FISG.
[720] Cf. BMF/BMJ (2020).
[721] Cf. BUNDESREGIERUNG (2021), article 15, para. 1 FISG; § 91, section 3 AktG.
[722] Cf. REGIERUNGSKOMMISSION (2022), principle 4.

Corporate Governance Code (Deutscher Corporate Governance Kodex, DCGK). As the board of directors has to evidence that both the entity's business activity and the risk situation have been considered in designing and implementing the system of internal control and the risk management system,[723] the FISG leads to a more formalized monitoring of internal controls and risk management procedures. By this, the FISG increases the importance of an appropriate and effective system of internal control over financial reporting for both the entities and the external auditor.

The divergence of theoretical and empirical research on using process mining for tests of controls

Understanding business processes and auditing integrated internal controls is the predominant use case identified in scientific research for process mining in the field of auditing.[724] Yet, as demonstrated in Chapters 2.2 and 4.3.1, there is only limited empirical data available evidencing the appropriateness of process mining to support audit procedures. The case audit firm's empirical application described in Chapter 3.3 was limited to using the technology to support substantive audit procedures, i.e., information obtained regarding the entity's system of internal control was not used as audit evidence. Further, the implementation excluded those audits performed in accordance with PCOAB auditing standards where the auditor expresses an opinion on the effectiveness of the company's internal control over financial reporting.[725] The audit firm's rationale for the limitation to substantive audit procedures included both the lack of a proper methodical integration of the technology into the ISA, which has been provided in Chapter 2.3 of this thesis, and the lack of empirical studies confirming the theoretical consensus that process mining may be reliably used to perform tests of controls in an audit of financial statements.

However, the feedback received as part of the implementation includes occasions where the audit team states that they have been able to identify and evaluate controls with process mining. Table 26 summarizes both control activities and control deviations identified by the audit teams of six different entities using process mining.

---

[723] Cf. REGIERUNGSKOMMISSION (2022), article 15, para. 1 FISG; § 91, section 3 AktG.
[724] Cf. Chapter 2.2.
[725] Cf. PCAOB AS 2201.01 and 2201.03.

Control activities or control deviations identified	Entity 1	Entity 2	Entity 3	Entity 4	Entity 5	Entity 6
Purchase order approval						
Purchase order approval	●	●	●	●	●	●
Approval on the appropriate authorization level	●	○	●	○	●	●
Approval after the purchase order has been changed	●	●	○	○	○	●
Three-way-match (considering SAP tolerances)						
Quantity (Purchase order vs. Goods received)	●	●	●	●	●	●
Quantity (Goods received vs. Invoice received)	●	●	●	●	●	●
Price (Purchase order vs. Goods received)	●	●	●	●	●	●
Price (Goods received vs. Invoice received)	●	●	●	●	●	●
Timely follow-up of exceptions (disputes) identified	●	●	○	○	●	●
Open cases with goods receipt but no invoice (accruals), coverage of GR/IR account	○	●	●	●	●	●
Activity sequence						
Automatic recording of bank statement on payment run	●	○	○	○	●	○
Retrospective purchase order creation	○	○	○	●	○	○
Segregation of duties						
Create purchase order & approve purchase order	●	●	●	●	●	●
Change purchase order & approve purchase order	○	●	●	○	●	○
Change purchase order quantity & post goods receipt	●	○	○	●	○	○
Create purchase order & post goods receipt	●	○	●	●		
Post goods receipt & post invoice	○	○	●	○	●	●
Remove payment block & post invoice	○	○	●	●	○	○
User analysis						
Users performing approval activities (purchase orders and/or invoices)	●	●	●	●	●	●
Users initiating payment run	●	○	●	○	●	●
Users releasing blocked invoices for payment	○	○	●	○	○	○
Master data						
Changes of master data	●	●	●	●	●	●
Changes of master data on weekend	●	○	●	●	●	●
Changes of bank account information	●	●	●	●	●	●

Table 26: Control activities and control deviations identified by the audit teams

All controls analyzed by the audit teams relate to the purchase to pay process. A black circle indicates that the control has been identified by the team, while controls with an empty circle have not been identified.[726] Six major categories of controls have been identified, including:

- the approval of the purchase order document by an authorized person,
- the price and/or quantity match between purchase order, goods receipt and invoice and the follow-up procedures performed for disputes identified,
- the sequence and timing of activities, including automated activities,
- the segregation of incompatible duties, i.e., activities that are expected to be performed by different users,
- the analysis of roles and responsibilities of the users performing key (control) activities and
- the master data maintenance, including changes made to bank information or at unusual times.

The analysis shows that the key IT dependent manual controls and application controls identified by the audit teams for the purchase to pay process are similar. Despite the controls related to the timing of activities, each audit team identified at least two control activities within each category, indicating that process mining supports identifying relevant internal controls in audit practice. However, the analysis is limited to six audit engagements only as after piloting, testing controls has been excluded from the scope of the implementation, and the teams did not cover the identification and evaluation of internal controls extensively in their feedback. Further empirical research may confirm common controls frequently implemented in purchase to pay or other business processes (which might depend, for example, on the industry or company size) and evaluate the audit evidence obtained related to the effectiveness of controls. However, over the course of the implementation project, many audit teams indicated that they would find it more intuitive if they would be able to use the technology for tests of controls rather than for substantive procedures only.[727]

[726] The information on control activities analyzed has been obtained both by reviewing the analyses performed using the process mining application and as part of the feedback conversations.
[727] Cf. Chapter 3.2.2 and Chapter 3.3.2.

Professional inspection findings related to tests of controls

An insufficient understanding of business processes relevant to financial reporting and inadequate tests of controls represent common findings in regulatory audit inspections. Quality issues identified as part of the annual inspection of the Big Four companies performed by the PCAOB frequently include occasions where:

- the audit team failed to identify and test relevant controls,
- the impact of control deficiencies on the financial statements has not been evaluated,
- the level of a control's precision has not been considered,
- the procedures performed to understand the relevant business process were limited to inquiry only,
- the extent of procedures performed to test controls, i.e., the sample sizes used, have been too small and
- the audit team failed to test the completeness of information produced by the entity as required by ISA 500.[728]

The introduction of the FISG, the divergence of theoretical and empirical process mining research and the recurring professional inspection findings give rise to further discuss the suitability of process mining for testing controls. Based on the methodical integration conducted in Chapter 2.3, the remainder of this Chapter specifically focuses on the technology's potential to improve the quality of traditional tests of controls and to address common findings of professional inspections.

4.5.2 Considering the holistic business process and financial reporting in testing internal controls

Shortcoming of the risk-based audit approach in today's business environment

The concept of process mining reveals several weaknesses of the audit risk model[729]. The traditional sample-based audit procedures performed to assess control risk do not support a robust assessment with regard to the reliability of the system of internal control – and this is not their purpose. They are designed to reduce the auditor's detection risk and by this the extent of remaining substantive procedures required to keep the audit risk at an acceptable low level.[730] However, this fundamental concept of the audit

[728] Cf. PCAOB (2018), pp. 7ff.; PCAOB (2016), p. 10.
[729] Cf. AICPA (1983), SAS No. 47.
[730] Cf. IFAC (2021), ISA 315 (Revised 2019), para. 34; Chapter 2.3.1.

risk model has its origin in the 1980s[731] and is based on the former state of information technology.[732] With the increased complexity of information technology supporting today's business processes,[733] the effort of understanding these processes and testing integrated internal controls increased significantly, leading to occasional doubts if auditing internal controls is still cost-effective as presumed in the audit risk model.[734]

Provided the importance of internal controls over financial reporting for the entity, its stakeholders and the external auditor – that has recently been reinforced by the FISG – it suggests itself to evaluate if process mining may help to overcome the shortcomings of the risk-based audit approach underlying today's audits of financial statements.

Evaluating the combined effectiveness of internal controls

In the audit risk model,[735] the multiplicative relation of IR and CR presumes that both components are stochastically independent.[736] This assumption does not necessarily hold true in practice.[737] Yet, controls are tested using samples that are considered independent from each other to a large extent. As such, when testing controls based on independent samples it is not possible to conclude on the effectiveness of controls within an individual process execution. The combination of multiple control weaknesses per case would only become transparent if the case is an element of all samples used to test relevant controls. While this might hold true when testing a control identified to mitigate another control's specific exception, no generalization of the joint effectiveness of controls for the entire population of transactions is possible.

As of today, it is not feasible for the auditor to identify all controls performed for an individual sampling item without manual walkthrough procedures, i.e., without following the transaction's processing along the critical path of the business process. Process mining enables to meet this objective for the entire population of cases simply by filtering the data.[738] For example, the process graph can be filtered for all cases for which the "purchase order approval" activity has not been performed in order to (a) identify if the approval should have been performed considering the control design and

[731] Cf. AICPA (1983), SAS No. 47.
[732] Cf. MOCHTY, LUDWIG (2015), p. 47.
[733] Cf. Chapter 4.1.2.
[734] Cf. MOCHTY, LUDWIG (2015), p. 47.
[735] Cf. Chapter 2.3.1.
[736] Cf. WIESE, MICHAEL (2013), p. 18.
[737] Cf. ibid.
[738] Cf. Chapter 2.3.4.

the characteristics of cases and to (b) determine if a mitigating control activity, like the approval of the related invoice, has been performed. By this, process mining supports to evaluate the combined effectiveness of relevant controls along the critical path of each individual transaction in the population in an effective and efficient manner.

Establishing the link between tests of controls and the financial statements

The traditional audit procedures related to internal controls over financial reporting are not directly linked to the financial statements but focus on the process-related aspects of transactions. When testing the operating effectiveness of a control based on a representative sample, the audit objective is to determine if the control effectively prevents (or detects and corrects) material misstatements. Frequently, the bidirectional connection between internal controls and the entity's bookkeeping remains unclear,[739] both in the auditor's walkthrough procedures to evaluate the design effectiveness of the control and when testing its operating effectiveness. Consequently, in performing conventional sample-based tests of controls, the auditor cannot precisely determine the total number and volume of transactions that actually was subject to the internal control and determine how these transactions have been recorded on the financial accounts.

As demonstrated in Chapter 4.2, the current process mining may be enhanced with information on the monetary amount (for example, the invoice amount) of a case at a specific point in time and the amount may be reconciled to the amount recorded in the entity's accounting system. By this, process mining provides the opportunity to exactly identify not only the transactions that have (or have not) been subject to a specific control activity, but also determine the monetary impact of these transactions on the financial statements. Due to the holistic analysis from the initiation of a transaction over its recording and processing up to the reporting in the financial statements, for the first time, process mining may enable a connection of controls with the (potential or actual) monetary impact of a control deviation on the financial statements. Consequently, the magnitude of (potential) misstatements resulting from a control deficiency may be precisely determined. As a result, control risk can be measured and a large extent of the sample-based audit of internal controls becomes obsolete.[740] Connecting the event log with the journal entry data further supports identifying and prioritizing

---

[739] Cf. MOCHTY, LUDWIG (2015), p. 35.
[740] Also see MOCHTY, LUDWIG (2015), p. 46.

relevant controls by determining which controls need to be tested in which order to ensure control objectives (derived from management's assertions[741]) are met.[742]

Removing sampling risk and the auditor's expectation about deviation rates

If an appropriately designed test of control is performed without any exception identified, the auditor may conclude that the control is operating effectively to prevent, or detect and correct, a material misstatement.[743] However, a test of control with process mining based on the entire population of transaction might show control exceptions that a traditional sample-based test of the same control would not show.[744] In this scenario, using the sample-based approach for testing the control would lead to the erroneous conclusion that the control is more effective than it actually is. Vice versa, if a control deviation is identified in a sample, the auditor needs to evaluate if it is systematic and represents a control deficiency or if the deviation is an anomaly that is not representative for the population.[745] With process mining, the representativeness of the deviation for the population may be determined from the data, supporting a more robust assessment of control deficiencies and removing the limitations inherent to sampling techniques, such as the sampling risk as defined in ISA 530.[746]

In addition, determining the extent of a test of control traditionally requires the auditor to determine the expected deviation rate of the control, as the sample size increases with the number of expected control deviations.[747] The comparison of the expected deviation rate with the detected rate may further indicate if the auditor can rely on the control to reduce the risk of material misstatement at the assertion level to the level assessed by the auditor.[748] As the auditor only tests the operating effectiveness of controls that are designed effectively,[749] usually no control deviations are expected. In considering the population of transactions instead of a sample, process mining shows if a control is sufficient to reduce the risk of material misstatement at the assertion level to the level assessed by the auditor, eliminating the need for a presumption about the expected deviation rate to design the sample and determine the sample size.

---

[741] Cf. Chapter 2.3.1.
[742] Cf. Mochty, Ludwig (2015), p. 32.
[743] Cf. IFAC (2021), ISA 330, para. 4(b); Chapter 2.3.4.
[744] Cf. Chapter 3.2.2.
[745] Cf. Chapter 2.3.4.
[746] Cf. ibid.; IFAC (2021), ISA 530, para. 5(c).
[747] Cf. IFAC (2021), ISA 330, para. 46.
[748] Cf. ibid., para. A41.
[749] Cf. IFAC (2021), ISA 315 (Revised 2019), para. A179.

If the audit opinion is based on the entire population of transactions instead of a sample, in the future, the lines between audit procedures for risk assessment (especially related to the entity's system of internal control) and substantive audit procedures may blur. This is also confirmed by the practical application, for example, the three-way-match analysis. The three-way-match between the purchase order, goods receipt and invoice is widely established as an important control mechanism in purchasing. However, at the same time, the related documents serve as audit evidence for the related transactions and are traditionally obtained when performing substantive procedures. As with process mining both tests of controls and substantive procedures may be performed based on the population instead of independent samples, the auditor may determine to follow a rely on controls approach for the transactions passing an effective set of control activities and only test the remainder of transactions substantively.

Addressing professional inspection findings related to tests of controls

Table 27 summarizes how the described configuration of process mining may support addressing common inspection findings identified by the PCAOB.[750]

Inspection finding	Solution approach[751]
The audit team failed to identify and test relevant controls.	The analysis of the number of volume of transactions passing a control activity helps to evaluate the relevance of a control to financial reporting. Relevant controls included in the process mining analyzer and are linked to the journal entries they initiate.
The impact of control deficiencies on the financial statements has not been evaluated.	The impact of control deficiencies is represented by the amount related to the case (purchase order, invoice or payment amount). Relevant amounts may be reconciled to related amounts recorded in the financial statements.
The level of a control's precision has not been considered.	The precision of controls is determined by the data model.
The procedures performed to understand the relevant business process were limited to inquiry only.	The procedures performed with process mining inherently go beyond inquiry only.
The extent of procedures performed to test controls, i.e., the sample sizes used, have been too small.	With process mining, control testing is based on the entire population of transactions instead of a sample only.
The audit team failed to test the completeness of information produced by the entity as required by ISA 500.	Reconciliation to the financial data supports assessing the completeness of the data.

Table 27: Potential of process mining to address common findings of professional inspections

Process mining may support reacting to the increased importance of the entity's system of internal control with an increased quality of control testing, including a precise determination of control risk based on the population of transactions. By this, the technology has the potential to revolutionize the audit of an entity's system of internal control. However, a prerequisite for the wider adoption of the technology in audit practice and further empirical exploration of its application in testing internal controls is that the weaknesses identified for the current process mining are addressed. As described earlier in this sub-chapter, this includes connecting the process and financial data[752] to support a robust assessment of control risk based on the entire population of transactions. Further limitations especially arise from the prerequisites of defining the

[750] Cf. Chapter 4.5.1; PCAOB (2016), p. 10; PCAOB (2018), pp. 7ff. The solution approaches assume that the challenges with regard to assembling relevant and reliable input data (cf. Chapter 4.1) have been resolved.
[751] Refer to Chapters 2.3.3 and 2.3.4 for a detailed integration of process mining into the auditor's control testing procedures required by the ISA.
[752] Cf. Chapter 4.2.

data model and preparing the process mining application.[753] The current process mining only supports evaluating controls that are covered by the data, i.e., control activities that have already been identified before actually applying process mining and performing the analyses. If internal controls need to be identified, described and technically defined as part of the data model specification, process mining does not replace the traditional audit procedures to understand the process, including the walkthrough procedures to identify relevant control activities and determine if they are implemented. Consequently, as of today, the potential benefit of using process mining to test internal controls only applies to recurring audits and is limited to the assessment of the effectiveness of controls whose definition and technical implementation has been determined beforehand.

[753] Cf. Chapter 4.1.

5 Concluding assessment of the appropriateness of process mining as an audit instrument

While the use of data analytics is widely established in the audit profession, the audit of internal processes and integrated controls continues to be a manual task. Almost unchanged from a methodical perspective since decades, understanding and evaluating an entity's process includes reading process narratives, inquiring process owners, observing procedures and behavior and testing controls or process outcomes (i.e., individual transactions) based on samples. Process mining promises to transfer the idea of "data analytics" to the audit of internal processes and integrated controls. Related research confirms the suitability of the technology as an audit instrument but is, however, predominantly based on theoretical considerations only.

This thesis contributed to the integration of process mining into the field of auditing by supplementing a detailed theoretical discussion of the technology's ability to meet relevant ISA requirements with an empirical evaluation of its implementation in the audit of various different entities. Based on the empirical findings, key points of criticism and acceptance challenges have been highlighted and solution approaches have been developed for those findings being of significant importance from an audit methodology and application perspective. By this, the thesis contributed to a broader dissemination of process mining in the audit profession.

A summarizing concluding assessment of the appropriateness of process mining as an audit instrument is provided by answering the research questions raised.

At (1): Which requirements have to be addressed by a data-based process analysis in an audit of financial statements?

Although process mining researchers and audit practitioners continuously promote process mining as an audit instrument, a systematic integration of the technology into relevant auditing standards has been missing so far. Chapter 2.3 of this thesis makes up for this discrepancy. In answering the first research question, the first interim conclusion of this thesis[754] summarizes applicable ISA requirements relevant for a data-driven process analysis and the key mechanisms substantiating the theoretical suitability of process mining to meet these requirements.[755]

[754] Cf. Chapter 2.3.5.
[755] Cf. Table 5.

Process mining may be used as risk assessment procedure to support identifying and assessing risks of material misstatement. By reconstructing the process model from the data extracted from the entity's information system and visualizing how and by whom the process instances are processed through the system, process mining inherently supports understanding business processes relevant to financial reporting. The activities in the event log and the process understanding obtained support the auditor in identifying those activities that represent control activities relevant to financial reporting, including controls over both routine and non-routine transactions. Instead of performing inquiries, observations and manual walkthrough procedures on an individual transaction only, process mining supports investigating the implementation and design effectiveness of control activities along the critical path for all cases included in the data.

When addressing identified risks, instead of drawing a sample to test the operating effectiveness of controls, tests of controls with process mining may be performed based on the population of transactions. The case and event attributes help to define a control exception, identify the root-cause of control deviations and ultimately assess control risk. For transactions that pass a compliant set of control activities, the auditor may follow a rely on controls approach at a high acceptable detection risk. For the subset of non-compliant transactions, a substantive audit approach is followed at a lower acceptable detection risk. Tests of details are supported by supplementing the transactions related to selected cases with supporting evidence, i.e., the document information extracted from the ERP system.

As demonstrated in this thesis, process mining's areas of application in the field of auditing might be further extended by enhancing the technology with financial information. As the audit objectives are linked to transactions and accounts, the auditor would especially benefit from determining the transaction volume (not) covered by the process data at hand. Determining the impact of a process execution on the financial statements would facilitate evaluating the magnitude of control exceptions identified from testing internal controls and the materiality of misstatements identified from substantive procedures.

In summary, the theoretical analysis of ISA requirements applicable to a data-driven process analysis supports the positive statements made in scientific research about the appropriateness of process mining as an audit instrument.

At (2): Which findings and challenges can be identified based on the empirical evaluation of the implementation of process mining in the audit practice?

Provided the theoretical consensus that process mining adds value to auditing,[756] the second research question aims to address the lack of empirical validation of available process mining research. In Chapter 3 of this thesis, process mining's practical suitability in the field of auditing is empirically evaluated based on the results from the deployment of 157 process mining applications to audit teams across 20 different countries. Contributions to the predominantly theoretical research on process mining are made by identifying practical challenges and highlighting key findings based on empirical feedback obtained from the audit teams of one of the Big Four audit firms and as part of the coaching and quality assurance procedures performed by the author of this thesis.

The practical evaluation confirms the potential of the solution to support audit procedures both in the phase of identifying and assessing risks and in designing and executing audit procedures in response to assessed risks. Audit teams acknowledge that on the one hand, the level of detail and the overall quality of the process understanding increased significantly when compared to traditional inquiry, observation and manual walkthrough procedures. On the other hand, the conclusions reached are more precise than evidence obtained based on a sample of transactions only.

However, several challenges are identified as part of the different waves of the practical application. The findings demonstrate that the implementation of process mining in auditing needs to go hand in hand with significant training and education efforts. While related acceptance challenges are partly comparable to those already observed when introducing other automated tools and techniques, process mining introduces a level of complexity that may not be compared to performing data analytics on populations of financial data. While today's auditor is familiar with analyzing financial data using data analytics, the analysis of process data requires extensive IT knowledge to avoid systematic errors impacting the audit evidence that may be obtained from the procedures performed. Piloting results such as the challenges in analyzing resulting process variations further show that training efforts need to be devoted to the interpretation of audit results obtained based on an investigation of the entire population of

[756] Cf. Chapter 2.2.

transactions rather than a sample. This is surprising provided the broad rejection of sampling techniques observable in the audit practice in other respects.

While several findings have been addressed over the course of the implementation project by modifying and expanding the process mining analyzer, there is a multitude of methodical and technical questions and challenges remaining that need to be answered and resolved, respectively, to establish process mining as an audit technique. The detailed findings obtained in response to the second research question of this thesis are listed and evaluated in Chapters 3.1.2, 3.2.2 and 3.3.2, respectively, and particularly relate to challenges with regard to (1) the ETL process, (2) the integration of the procedures performed with process mining into the audit approach, (3) the tool handling and tool performance and (4) the analyses and functionalities provided in the analyzer.

This thesis reveals five challenges regarding the application of process mining in the audit that are of significant importance from an audit methodology and application perspective:

(1) *The missing explorative character of process mining:* The "log files" automatically and chronologically recording data relevant to the process that are propagated as easily accessible and reliable input data for process mining are not existing in today's ERP systems. Resulting challenges in scoping and assembling the event log include understanding the entity's IT landscape to identify and locate relevant data, determining the process instance, relevant activities and the data extraction period, and technical constraints with regard to analyzing the timing and duration of events. The previous knowledge and process understanding required to set-up an analyzer that promises to facilitate understanding this very process is perceived as blocker for adoption by many auditors. The relevance of the technology for the audit profession is further reduced by inherent limitations regarding the completeness and accuracy of the data.

(2) *The missing link between process and financial data:* If the process data may not be reconciled to the auditee's financial statements, the auditor may not determine if the population of cases underlying the recording of transactions on the general ledger accounts related to the process is complete. Piloting further showed the audit teams' interpretation difficulties when analyzing the

development of the invoice volume related to a specific vendor without information on the related liabilities recorded on the vendor account. A missing integration of process and financial data may lead to over-auditing of classes of transactions or accounts with a limited risk of material misstatement and under-auditing of significant classes of transactions or accounts.

(3) *Numerous resulting process variations that cannot be audited in their entirety:* In contrast to many examples provided in related process mining literature, the extensive number of process variations resulting in practice may not be handed over to the auditor for "further investigation". Piloting showed uncertainty about the degree to which variations need to be analyzed and the evaluation criteria applicable to this analysis.

(4) *Lacking efficiency of process mining due to missing comparison and roll-forward functionalities:* The comparison of different process mining applications is not only a tedious task but frequently ineffective due to the characteristics of case and event data that prevent a meaningful comparison of data between different applications.

(5) *Using process mining to remove limitations inherent to today's tests of controls rather than for performing substantive procedures only:* Despite the case audit firm's decision to limit the implementation to those audit teams following a substantive audit approach, the results of the implementation confirmed that process mining may support identifying and testing internal controls. In fact, audit teams stated it would be more appropriate to use the solution to test the design and operating effectiveness of controls rather than to perform substantive procedures only.

At (3): Provided the methodical requirements on the one hand and the practical challenges on the other hand, which modifications of process mining are necessary to address the key points of criticism?

The five key points of criticism related to the practical applicability of the current process mining in an audit of financial statements are systematically discussed in Chapter 4 of this thesis. In answering the third research question, the discussion includes the development of solution approaches and the proposal of modifications for future implementations of process mining to address each challenge identified.

Chapter 4.1 is devoted to the circumstance that process mining does not exploratively "mine" the IT systems for relevant events but rather descriptively discovers the process model from a given event log. A solution approach may involve ERP providers to introduce an "event ledger", logging the accounting relevant process execution similar to the recording of financial data in the general ledger. In the future, advanced algorithms may further support exploratively extracting event data related to the audit period's financial accounting documents (i.e., the journal entries in the general ledger) from the entity's IT systems. Many of the challenges identified throughout this thesis relate back to the use of relational databases in today's process mining applications. Consequently, the proposed solution approach further includes the consideration of graph databases to establish required links between data sources and maintain the process mining data model in an effective manner. As an entity's bookkeeping may already be represented as a network as of today,[757] storing process mining input data as a graph would further facilitate the integration of process and financial data, enabling the investigation of the complete flow of transactions as required by ISA 315.

As today's process mining is based on relational databases and an approach to effectively identify and extract relevant data from the source systems is missing, the remainder of the discussion in Chapter 4 focuses on modifications of the process mining based on relational databases to address the key points of criticism.

In addressing the missing link between process and financial data, Chapter 4.2 presents how process mining may be modified to (a) consider the entity's general ledger and (b) subledger data and their reconciliation to the process data and by this (c) support determining the impact of a process instance on the financial statements. For the first time, the modifications enable to transfer the concept of materiality to process mining and fully benefit from the data-driven analysis of the entire critical path of a transaction. Considering the general ledger data expands process mining's field of application to the assessment of the significance of a class of transactions, the deconstruction of the process into significant classes of transactions that have not been previously identified, the identification of significant transaction volumes that have been expected to relate to the analyzed process but are not covered by the extracted data and, at the same time, supports challenging the completeness of the event log. The integration of the subledger data addresses the audit teams' interpretation difficulties when analyzing

[757] Cf. MOCHTY, LUDWIG (1985), pp. 2ff.

vendor information without considering related liabilities recorded on the vendor accounts. Determining the monetary financial impact of process instances throughout the process execution helps to transfer familiar substantive procedures commonly performed by the auditor to process mining, addressing adoption challenges especially with regard to procedures performed on subledger data. On the other hand, it was demonstrated how the combined consideration of process and financial data may, for the first time, enable an analytical approach to address the completeness of purchases during the audit period and liabilities at period end.

Chapter 4.3 evaluates the shortcomings of theoretical research in evaluating process variations and analyzes the suitability of available techniques to handle the extensive number of process variations resulting in practice. As a large number of variations does not necessarily lead to a high risk of material misstatement, instruments are identified and evaluated that may help the auditor to determine the process executions to investigate in detail. While the connection of the process to the financial data helps in determining the materiality of individual variations, qualitative factors are needed both to respond to completeness risks and to identify those process variations that may, individually or in combination, lead to a material misstatement of the financial statements. The evaluation demonstrates how especially (a) identifying and addressing risks by analyzing the population of transactions, (b) altering the definition of a variation based on the objective of the audit procedure and (c) grouping variations by selected criteria while considering the results from the procedures already performed based on the entire population may reduce the remaining process paths to investigate in detail significantly.

Chapter 4.4 addresses the issue that current process mining techniques do not support comparing data from different periods, impeding the comparison of the audit period's data with an interim date or the previous period. After discussing the practical challenges preventing an effective comparison of data between different process mining applications, a concept for integrating a comparison functionality within an individual process mining application is developed by the author of this thesis. The overall concept summarized in Appendix V considers that different types of data require different approaches to allow for a meaningful comparison between different observation periods. The largest benefit not only for the auditor but across all disciplines considering the implementation of process mining is derived from the comparison of process variations. The proposed modification enables to use the same variation ID in all periods to describe the same sequence of activities. While the auditor may still analyze significant changes in the number or volume of cases in a specific variation, a variation ID that has already been audited in previous periods would not need to be investigated from a qualitative perspective anymore in the audit period and in subsequent periods. Besides facilitating effective analytical procedures, the proposed enhancement of process mining with a functionality to compare data between different periods significantly increases the efficiency of using the technology in an audit of financial statements.

Although the audit firm's final implementation of process mining did not focus on using the technology to test internal controls, Chapter 4.5 identifies that key IT dependent manual controls and application controls identified by the audit teams for the purchase to pay process are the same for different engagements, and many audit teams fed back that they have been able to identify and test relevant controls with process mining. It was demonstrated how process mining may help to overcome the shortcomings of the risk-based audit approach underlying today's audits of financial statements and address common control testing related findings from professional inspections. Testing controls based on the population of transactions in particular removes the assumptions and limitations inherent to a sample-based approach for testing internal controls over financial reporting. The proposed integration of the process data and the financial data further helps the auditor to determine the impact of a control exception (or a misstatement identified through substantive procedures) on the financial statements. When identifying control exceptions, the auditor does not extend the sample size but assesses whether the monetary amount of the affected transactions is material to the financial statements. However, even if the auditor concludes that the financial accounts related to the process are free from material misstatements, discussions within the audit profession will be needed with regard to how the auditor may provide an opinion of the effectiveness of internal controls if a known and significant deviation rate is evidenced by the data. As the lines between control testing and substantive procedures may blur with process mining and error rates determined with process mining may not be interpreted as probability in the sense of the audit risk model, a new concept for interpreting error rates identified based on the entire population of transactions instead of a sample only will be required for concluding on the effectiveness of internal controls.

In summary it can be stated that current process mining techniques may be modified to support an effective audit in line with applicable auditing standards. From the modifications made, in particular the integration of financial data may significantly increase the technology's areas of application in a financial statement audit while helping to avoid over- and under-auditing of classes of transactions and accounts related to the process. The proposed concepts for evaluating process variations and comparing process data between different periods help to increase the efficiency of the audit procedures performed.

A prerequisite for a broad adoption of the technology within the audit profession is that the weaknesses of today's process mining solutions and the challenges identified are resolved. This thesis revealed the huge complexity of the process models resulting in practice and the analyses and extensive documentation required to appropriately identify and address related risks of material misstatement. The empirical evaluation further confirmed that a considerable extent of training, coaching and quality assurance procedures needs to accompany the adoption of the technology.

Limitations and acceptance challenges especially arise from the costly prerequisites of determining the input data, defining the data model and preparing the process mining application. If the auditor requires detailed knowledge of relevant IT systems and the business process to assemble the data in the first place, the value added by process mining decreases significantly. Consequently, as of now, today's process mining does not replace the traditional audit procedures to understand the process, including walkthrough procedures to identify relevant control activities that need to be included in the data model. While the reconciliation of the process data to the transaction volume covered on related financial accounts helps to assess the completeness of the data model, inherent limitations regarding the completeness of data further require traditional audit procedures to identify, assess and address risks of material misstatement for those parts of the process that are not covered by the process mining analyzer.

6 Outlook

While this thesis demonstrated how today's process mining may be modified and used within an audit of financial statements, it will still take many years - and additional theoretical and empirical research - to broadly implement a data-driven audit in all steps of the risk-based audit approach. This thesis brought up several areas where further research may help to accelerate the practical adoption of process mining.

The largest benefit may be derived from establishing a best practice in preparing the process mining input data. Today's process of building the event log for process mining is a manual process that is prone to errors. The numerous decisions that need to be made as part of this process include, for example, the choice of the process instance and its granularity and the determination of the activities to consider in the event log. These decisions systematically limit the application scope of later analyses.

Further research is required in effectively identifying and assembling relevant data from ERP systems and related preprocessing applications. Besides the consideration for ERP providers to develop an event ledger chronologically recording the execution of activities in a business process,[758] a potential alternative scenario for future research may include enhancing process mining by an actual explorative "mining" of relevant IT systems. In the future, advanced algorithms might be used to train a system to learn from existing data models in order to automatically recognize and extract data having the nature of an event. When extracting the data strings that appear to be an event from the IT systems of the entity, the auditor may receive a list of activities that have been identified, including their meta data, related documents and information on their source systems. Based on this list of activities, the auditor may determine the activities relevant to financial reporting.[759] While this scenario requires further research, theoretically, it may remove the limitations with regard to the choice of a process instance.[760] The document that is followed through the process may be determined flexibly, depending on the definition of related activities and the objective of the analysis. For example, if the auditor analyzes the approval of purchase orders (or invoices), the dynamic process instance would be the entire purchase order document (or invoice document, respectively).

[758] Cf. Chapter 4.1.8.
[759] Cf. Chapter 4.1.3.
[760] Cf. Chapter 4.1.4.

If the relevant process data can be reliably extracted from the source systems, Chapter 4.1.8 suggests additional research to explore how to effectively combine process and financial data in a graph database, automatically establish links between data including the same references and support the auditor in easily accessing this data to design and perform audit procedures. The effective extraction and integration of process and financial data would not only reduce the number of manual steps and extensive involvement of IT specialists in data preparation but expand process mining's areas of application in auditing significantly.

For example, procedures designed to identify and address risks of material misstatement due to fraud usually involve the so-called "journal entry testing" that is traditionally performed based on general ledger data only. However, the information content of a few columns of general ledger data is limited and frequently leads to "false positive" results. For example, in practice, journal entries usually lack a meaningful description of their business purpose, as the related document is attached in the ERP system. However, traditional automated tools and techniques based on general ledger data do not extract related documents from the system. The information on the documents related to a transaction and the activities performed in advance to the recording of a journal entry may significantly enhance both effectiveness and efficiency of fraud detection procedures in the future.

With a holistic analysis of an entity's business processes relevant to financial reporting, the time consuming, restrictive and costly development of an individual process mining application for each individual business process may become obsolete in the future. In exploring the benefits that enhancing the journal entry network with event log data may provide to the auditor,[761] MOCHTY highlights the analysis of relationships and dependencies between business processes.[762] Research topics further arise with regard to the utilization of resources, resource planning as well as collaboration or collusion between resources involved in processing transactions. Process mining in the "BPM in the large"[763] context may further enable to expand the application of process mining to those COSO components[764] of internal control operating at

---

761 Cf. MOCHTY, LUDWIG (2015), p. 46.
762 Cf. ibid., p. 32; Chapter 4.1.8.
763 Cf. Chapter 4.1.8.
764 The COSO committee develops guidance that helps organizations to design, implement, execute and monitor a system of internal control over financial reporting, cf. COSO (1994). The individual COSO components of internal control have been incorporated into the ISA, cf. Chapter 2.3.3; IFAC (2021), ISA 315 (Revised 2019), para. 12(m).

the level of the entity rather than on the transaction or process level.[765] Related research requestions that arise include analyzing to what extent process mining may support identifying and auditing so-called "entity level controls" and general IT controls that do not focus on a particular business process only.

While the largest potential for process mining in the field auf auditing may be derived from further theoretical and empirical research with regard to the data preparation discussed in Chapter 4.1, this thesis brought up several additional areas that might benefit from further exploration.

Besides the proposed concept for comparing process data at different points in time,[766] the audit profession may benefit from the ability to automatically reperform analyses at a future point in time. For example, the analyses performed at the interim date may be automatically reperformed within the process mining analyzer for the roll-forward period until period end. Documentation may be facilitated by a summary highlighting changes (or the absence of changes), for example, new process variations or open items that have been cleared throughout the roll-forward period.

In exploring the root-cause of different process variations, the auditor would significantly benefit from an instrument for identifying the reasons for a case to use an alternative path through the process than other cases. This may involve the use of machine learning algorithms to reliably identify decision points in the process, i.e., points with exclusive branches into different process execution due to the characteristics of the underlying cases. In this regard, further research may answer the question if it is possible to exploratively identify business rules or control activities from the process data that determine the necessary conditions for a case to follow a certain path of the process.[767] Provided the challenges identified in Chapter 4.3.2 with regard to clustering similar process executions while still being able to interpret resulting clusters from a business perspective, research may further explore if in the future, the use of advanced trace clustering techniques may be superior to other means of evaluating process variations. The concept for the comparison of process variations discussed in Chapter 4.4.2 may be further developed to support the auditor in evaluating new process variations occurring in the audit period under review. In the future, algorithms may

---

765 Cf. Figure 15.
766 Cf. Chapter 4.4.2.
767 Cf. ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012), p. 358.

help to identify the most similar process execution among those variations already audited in previous periods to facilitate evaluating the impact of the deviations.

Process mining in auditing is still in its infancy. As of today, the practical adoption in the audit profession is primarily limited to the Big Four audit firms. Within these audit firms, the application of process mining is limited to handpicked companies meeting specific criteria that allow for the implementation of a particular process mining solution. This process mining solution is limited to the analysis of an isolated business process. An audit team may apply the process mining solution introduced in this thesis to an entity's purchase to pay process. However, in order to evaluate all business processes relevant to financial reporting with process mining, an additional process mining solution and appropriate analyses would need to be designed for each remaining significant class of transaction, such as the order to cash cycle, the payroll cycle, or the inventory and warehousing cycle. Further, the audit teams need continuous support from a team of experts in order to obtain sufficient and appropriate audit evidence from applying process mining.

While this thesis demonstrated that an effective use of process mining in auditing requires both researchers and audit practice to jointly work on resolving current weaknesses and challenges identified, finally, additional work also results for standard setters. Current auditing standards have been written at a point in time where data was not readily available. With the revision of ISA 315 and the AICPA's guidance to data analytics[768] the regulators started to acknowledge the circumstance that audits may be supported by data analytics. This dissertation demonstrated that from a regulatory perspective the ISA do not contradict an application of process mining, however, the introduction of terms like "automated tools and techniques" neither supports this application. This is especially resulting from the defective operationalization of the ISA – a task the audit profession should continuously work on to remove acceptance challenges and accelerate the practical adoption of digital solutions in auditing.

[768] Cf. AICPA (2017).

List of references

ACCORSI, RAFAEL/ULLRICH, MEIKE/VAN DER AALST, WIL M. P. (2012): Process Mining. In: Informatik Spektrum, no. 35, pp. 354-359.

ACFE (2008): Fraud Examiners Manual. International edition, Austin, Texas.

ADRIANSYAH, ARYA/SIDOROVA, NATALIA/VAN DONGEN, BOUDEWIJN F. (2011): Cost-based fitness in conformance checking. In: 11th International Conference on Application of Concurrency to System Design, pp. 57-66.

ADRIANSYAH, ARYA/VAN DONGEN, BOUDEWIJN. F./VAN DER AALST, WIL M. P. (2011a): Conformance checking using cost-based fitness analysis. In: 15th IEEE International Enterprise Distributed Object Computing Conference, pp. 55-64.

ADRIANSYAH, ARYA/VAN DONGEN, BOUDEWIJN F./VAN DER AALST, WIL M. P. (2011b): Towards Robust Conformance Checking. In: ZUR MUEHLEN M./SU J. (editors): Business Process Management Workshops. Lecture Notes in Business Information Processing, pp. 122-133.

AGRAWAL, RAKESH/GUNOPULOS, DIMITRIOS/LEYMANN, FRANK (1998): Mining process models from workflow logs. In: Proceedings of the 6th International Conference on Extending Database Technology: Advances in Database Technology, pp. 469-483.

AGUIRRE, SANTIAGO/PARRA, CARLOS/SEPÚLVEDA, MARCOS (2017): Methodological proposal for process mining projects. In: International Journal of Business Process Integration and Management, volume 8, no. 2, pp. 102-113.

AICPA (1983): Audit risk and materiality in conducting an audit. Auditing Standards Boards, Statements on Auditing Standards 47.

AICPA (2017): Guide to Audit Data Analytics. Wiley, Hoboken, New Jersey.

AILENEI, IRINA/ROZINAT, ANNE/ECKERT, ALBERT/VAN DER AALST, WIL M. P. (2011): Definition and validation of process mining use cases. In: International conference on business process management. Springer, Heidelberg, Germany, pp. 75-86.

ALLES, MICHAEL/JANS, MIEKE/VASARHELYI, MIKLOS A. (2011): Process mining: a new research methodology for AIS. Annual Conference 2011, SSRN Electronic Journal. URL https://ssrn.com/abstract=1746926.

ARIAS, MICHAEL/SAAVEDRA, RODRIGO/MARQUES, MARIA R./MUNOZ-GAMA, JORGE/SEPÚLVEDA, MARCOS (2018): Human resource allocation in business process management and process mining: A systematic mapping study. In: Management Decision, volume 56, issue 2, pp. 376-405.

AUGUSTO, ADRIANO/CONFORTI, RAFFAELE/DUMAS, MARLON/LA ROSA, MARCELLO (2017): Split miner: discovering accurate and simple business process models from event logs. In: 2017 IEEE International Conference on Data Mining. New Orleans, LA, pp. 1-10.

AUGUSTO, ADRIANO/CONFORTI, RAFFAELE/DUMAS, MARLON/LA ROSA, MARCELLO/MAGGI, FABRIZIO M./MARRELLA, ANDREA/MECELLA, MASSIMO/SOO, ALLAR (2019a): Automated discovery of process models from event logs: review and benchmark. In: IEEE Transactions on Knowledge and Data Engineering, volume 31, issue 4, pp. 686-705.

AUGUSTO, ADRIANO/CONFORTI, RAFFAELE/DUMAS, MARLON/LA ROSA, MARCELLO/POLYVYANYY, ARTEM (2019b): Split miner: Automated discovery of accurate and simple business process models from event logs. In: Knowledge and Information Systems, volume 59, issue 2, pp. 251-284.

BACHER, JOHANN/PÖGE, ANDREAS/WENZIG, KNUT (2010): Clusteranalyse: Anwendungsorientierte Einführung in Klassifkationsverfahren. 3. Auflage, De Gruyter, Oldenburg, Germany.

BECKER, MICHAEL/BUCHKREMER, RÜDIGER (2019): A practical process mining approach for compliance management. In: Journal of Financial Regulation and Compliance, volume 27, no. 4, pp. 464-478.

BIERSTAKER, JAMES/JANVRIN, DIANE/LOWE, D. JORDAN (2014): What factors influence auditors' use of computer-assisted audit techniques. In: Advances in Accounting, volume 30, issue 1, pp. 67-74.

BMF/BMJ (2020): Entwurf eines Gesetzes zur Stärkung der Finanzmarktintegrität (Finanzmarktintegritätsstärkungsgesetz – FISG [WWW Document]. URL https://www.bundesfinanzministerium.de/Content/DE/Gesetzestexte/Gesetze_Gesetzesvorhaben/Abteilungen/Abteilung_VII/19_Legislaturperiode/2021-06-10-FISG/1-Referentenentwurf.pdf?__blob=publicationFile&v=3 (last accessed: 1 January 2023).

BMJ (2021): Gesetzgebungsverfahren: Gesetz zur Stärkung der Finanzmarktintegrität [WWW Document]. URL https://www.bmj.de/SharedDocs/Gesetzgebungsverfahren/DE/Finanzmarktintegritaet.html (last accessed: 1 January 2023).

BODENMANN, ANDREAS/MAAS, ANNALOUISE/SCHWAN, MORITZ (2018): Die Zukunft des digitalisierten Rechnungswesens. In: Schweizerische Zeitschrift für Wirtschafts- und Finanzmarktrecht, volume 1, pp. 33-47.

BPM CENTER (2023): A collaborative virtual research center in the area of BPM [WWW Document]. URL http://bpmcenter.org/reports (last accessed: 19 February 2023).

BRAUN, ROBERT L./DAVIS, HAROLD E. (2003): Computer-assisted audit tools and techniques: analysis and perspectives. In: Managerial Auditing Journal, volume 18, no. 9, pp. 725-731.

BRYNES, PAUL/CRISTE, TOM/STEWART, TREVOR/VASARHELYI, MIKLOS A. (2017): Reimagining Auditing in a Wired World. AICPA Whitepaper [WWW Document]. URL www.aicpa.org/interestareas/frc/assuranceadvisoryservices/downloadabledocuments/whitepaper_blue_sky_scenario-pinkbook.pdf (last accessed: 9 May 2022).

BUIJS, JOOS C. A. M./LA ROSA, MARCELLO/REIJERS, HAJO A./VAN DONGEN, BOUDEWIJN F./VAN DER AALST, WIL M. P. (2013): Improving Business Process Models Using Observed Behavior. In: CUDRE-MAUROUX, P./CERAVOLO, P./GAŠEVIĆ, D. (editors) (2013): Data-Driven Process Discovery and Analysis. Lecture Notes in Business Information Processing, no. 162, pp. 44-59.

BUIJS JOOS C./VAN DONGEN, BOUDEWIJN F./VAN DER AALST, WIL M. P. (2014): Quality dimensions in process discovery: the importance of fitness, precision, generalization and simplicity. In: International Journal of Cooperative Information Systems, volume 23, no. 1, p. 1440001.

BUNDESREGIERUNG (2021): Gesetz zur Stärkung der Finanzmarktintegrität (Finanzmarktintegritätsstärkungsgesetz – FISG). In: Bundesgesetzblatt Jahrgang 2021, Teil 1, Nr. 30, pp. 1534-1567 [WWW Document]. URL https://www.bundesfinanzministerium.de/Content/DE/Gesetzestexte/Gesetze_Gesetzesvorhaben/Abteilungen/Abteilung_VII/19_Legislaturperiode/2021-06-10-FISG/3-Verkuendetes-Gesetz.pdf?__blob=publicationFile&v=2 (last accessed: 31 December 2022).

BURATTIN, ANDREA (2015): Process Mining Techniques in Business Environments: Theoretical Aspects, Algorithms, Techniques and Open Challenges in Process Mining, Band 207. Springer International Publishing, Heidelberg, Germany.

CHIU, TIFFANY (2018): Exploring new audit evidence: The application of process mining in auditing, PhD thesis.

CHIU, TIFFANY/BROWN-LIBURD, HELEN/VASARHELYI, MIKLOS A. (2019): Performing Tests of Internal Controls using Process Mining: What could go wrong? In: The CPA Journal [WWW Document]. URL: https://www.cpajournal.com/2019/07/10/performing-tests-of-internal-controls-using-process-mining (last accessed: 22 March 2023).

CICA (2003): Electronic Audit Evidence: Research Report, Toronto, Ontario.

CLAES, JAN/GEERT, POELS (2012): Process Mining and the ProM Framework: An Exploratory Survey. In: ROSA, MARCELLO/SOFFER, PNINA (editors): Business Process Management Workshops. Springer, Tallinn, Estonia, pp. 187-198.

COOK, JONATHAN E./WOLF, ALEXANDER L. (1998a): Discovering models of software processes from event-based data. In: ACM Transactions on Software Engineering and Methodology, volume 7, issue 3, pp. 215-249.

COOK, JONATHAN E./WOLF, ALEXANDER L. (1998b): Event-based detection of concurrency. In: Proceedings of the Sixth International Symposium on the Foundations of Software Engineering, pp. 35-45.

COOK, JONATHAN E./WOLF, ALEXANDER L. (1999): Software Process Validation: Quantitatively Measuring the Correspondence of a Process to a Model. In: ACM Transactions on Software Engineering and Methodology, volume 8, no. 2, pp. 147-176.

COSO (1994): Internal Control - Integrated Framework, 2nd edition, New Jersey: Committee of Sponsoring Organizations of the Treadway Commission.

DE CNUDDLE, SOFIE/CLAES, JAN/POELS, GEERT (2014): Improving the quality of the heuristics miner in ProM 6.2. In: Expert Systems with Applications, volume 41, no. 17, pp. 7678-7690.

DE GIACOMO, GIUSEPPE/LEMBO, DOMENICO/LENZERINI, MAURIZIO/POGGI, ANTONELLA/ROSATI, RICCARDO (2018): Using Ontologies for Semantic Data Integration. In: FLESCA, SERGIO/GRECO, SERGIO/MASCIARI, ELIO/SACCÀ, DOMENICO (editors): A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years: Studies in Big Data. Springer, pp. 187-202.

DE LEONI, MASSIMILIANO/VAN DER AALST, WIL M. P. (2013): Aligning event logs and process models for multi-perspective conformance checking: an approach based on integer linear programming. In: DANIEL, F./WANG, J./WEBER B. (editors): Business process management. Springer, Heidelberg, pp. 113-129.

DE MEDEIROS, ALVES A. K. (2006): Genetic Process Mining. Eindhoven University of Technology, Eindhoven, The Netherlands.

DE MEDEIROS, ALVES A. K./VAN DONGEN, B./VAN DER AALST, WIL M. P./WEIJTERS, A. J. M. M. (2004): Process mining: Extending the α-algorithm to mine short loops, Beta Working Paper Series.

DE MEDEIROS, ALVES A. K./VAN DER AALST, WIL M. P. (2007): Process Mining towards Semantics. In: DILLON, THARAM S./CHANG, ELIZABETH/MEERSMAN, ROBERT/SYCARA, KATIA (editors): Advances in Web Semantics I: Ontologies, Web Services and Applied Semantic Web. Springer, Berlin, Heidelberg, Germany, pp. 35-80.

DE MEDEIROS, ALVES A. K./WEIJTERS, A. J. M. M./VAN DER AALST, WIL M. P. (2007): Genetic process mining: an experimental evaluation. In: Data Mining and Knowledge Discovery, volume 14, no. 2, pp. 245-304.

DE WEERDT, JOCHEN/VANDEN BROUCKE, SEPPE/VANTHIENEN, JAN/BAESENS, BART (2013): Active trace clustering for improved process discovery. In: IEEE Transactions on Knowledge and Data Engineering, volume 25, no. 12, pp. 2708-2720.

DEBRECENY, ROGER/LEE, SOOK-LENG/NEO, WILLY/SHULING TOH, JOCELYN (2005): Employing generalized audit software in the financial services sector: Challenges and opportunities. In: Managerial Auditing Journal, volume 20, no. 6, pp. 605-618.

DELOITTE (2020): Process Mining in der Wirtschaftsprüfung [WWW Document]. URL https://www2.deloitte.com/de/de/pages/audit/articles/process-mining-wirtschaftspruefung.html (last accessed: 16 November 2020).

DOS SANTOS GARCIA, CLEITON/MEINCHEIM, ALEX/RIBEIRO FARIA JUNIOR, ELIO/ROSANO DALLAGASSA, MARCELO/VECINO SATO, DENISE M./RIBEIRO CARVALHO,

DEBORAH/PORTELA SANTOS, EDUARDO A./SCALABRIN, EDSON E. (2019): Process mining techniques and applications – A systematic mapping study. In: Expert Systems with Applications, volume 133, pp. 260-295.

DREHER, SIMON/REIMANN, PETER/GRÖGER, CHRISTOPH (2020): Application Fields and Research Gaps of Process Mining in Manufacturing Companies. In: REUSSNER, R. H./KOZIOLEK, A./HEINRICH, R. (editors): Informatik 2020. Gesellschaft für Informatik, Bonn, Germany, pp. 621-634.

DUMAS, MARLON./LA ROSA, MARCELLO/MENDLING, JAN/REIJERS, HAJO A. (2018): Fundamentals of business process management, 2nd edition. Springer, Heidelberg, Germany.

EDGINGTON, THERESA M./RAGHU, T. SANTANAM/VINZE, AJAY S. (2010): Using process mining to identify coordination patterns in IT service management. In: Decision Support Systems, volume 49, pp. 175–186.

ELDER, RANDAL J./BEASLEY, MARK S./HOGAN, CHRIS E./ARENS, ALVIN A. (2020): Auditing and Assurance Services: International Perspectives. 17th Edition, Pearson Education Limited.

EY (2019): Die Digitalisierung der Abschlussprüfung [WWW Document]. URL www.ey.com/Publication/vwLUAssets/ey-reporting-magazin-2016/$FILE/ey-reporting-magazin-2016.pdf (last accessed: 21 November 2020).

FERNÁNDEZ-LLATAS, CARLOS/LIZONDO, AROA/MONTON, EDUARDO/BENEDÍ, JOSE-MIGUEL/TRAVER, VICENTE (2015): Process mining methodology for health process tracking using real-time indoor location systems. In: Sensors, volume 15, issue 12, pp. 29821-29840.

FINANCIAL REPORTING COUNCIL (2017): Audit Quality Thematic Review: The Use of Data Analytics in the Audit of Financial Statements.

GAALOUL, WALID/GODART, CLAUDE (2005): Mining workflow recovery from event based logs. In: VAN DER AALST, WIL M. P./BENATALLAH, BOUALEM/CASATI, FABIO/CURBERA, FRANCISCO (editors): Business Process Management. Springer, Heidelberg, Germany, pp. 169-185.

GÄRTNER, MICHAEL (1994): Analytische Prüfungshandlungen im Rahmen der Jahresabschlussprüfung: ein Grundsatz ordnungsmäßiger Abschlussprüfung. Hitzeroth, Marburg, Germany.

GEHRKE, NICK/MÜLLER-WICKOP, NIELS (2010): Basic principles of financial process mining: a journey through financial data in accounting information systems. In: Proceedings of the 16th Americas Conference on Information Systems. AMCIS, Lima, Peru.

GEHRKE, NICK/WERNER, MICHAEL (2013): Process Mining. In: Die Zeitschrift für den Wirtschaftsstudenten, volume 7, pp. 1-16.

GOEDERTIER, STIJN/MARTENS, DAVID/VANTHIENEN, JAN/BAESENS, BART (2009): Robust process discovery with artificial negative events. In: Journal of Machine Learning Research, volume 10, pp. 1305-1340.

GRECO, GIANLUIGI/GUZZO, ANTONELLA/PONTIERI, LUIGI/SACCA, DOMENICO (2006): Discovering expressive process models by Clustering Log Traces. In: IEEE Transactions on Knowledge and Data Engineering, volume 18, no. 8, pp. 1010-1027.

GRISOLD, THOMAS/MENDLING, JAN/OTTO, MARKUS/VOM BROCKE, JAN (2021): Adoption, use and management of process mining in practice. In: Business Process Management, volume 27, no. 2, pp. 369-387.

GROMOV, VLADIMIR (2014): Diagnostics in compliance checking. Master thesis, University of Eindhoven, Eindhoven, The Netherlands.

GUO, QINLONG/WEN, LIJIE/WANG, JIANMIN/YAN, ZHIQIANG/PHILIP, S. YU (2015): Mining invisible tasks in non-free-choice constructs. In: MOTAHARI-NEZHAD, HAMID REZA/RECKER, JAN/WEIDLICH, MATTHIAS (editors): Business Process Management. Springer, Innsbruck, Austria, pp. 109-125.

GÜNTHER, CHRISTIAN W./VAN DER AALST, WIL M. P. (2007): Fuzzy mining - Adaptive process simplification based on multi-perspective metrics. In: Business Process Management, Proceedings, pp. 329-343.

HARTKE, LARS/HOHNHORST, GEORG/SATTLER, GERNOT (2010): SAP-Handbuch Sicherheit und Prüfung: Praxisorientierter Revisionsleitfaden für SAP-Systeme. 4th Edition, IDW Verlag, Düsseldorf, Germany.

HERBST, JOACHIM (2001): Ein induktiver Ansatz zur Akquisition und Adaption von Workflow-Modellen, PhD thesis, University of Ulm, Germany.

HERBST, JOACHIM/KARAGIANNIS, DIMITRIS (1998): Integrating Machine Learning and Workflow Management to Support Acquisition and Adaption of Workflow Models. In: Proceedings of the Ninth International Workshop on Database and Expert Systems Applications, pp. 745-752.

HERBST, JOACHIM/KARAGIANNIS, DIMITRIS (1999): An Inductive Approach to the Acquisition and Adaptation of Workflow Models. In: Proceedings of the IJCAI Workshop on Intelligend Workflow and Process Management: The New Frontier for AI in Business, pp. 52-57.

HOBECK RICHARD/KLINKMÜLLER CHRISTOPHER/BANDARA, H. M. N. DILUM/WEBER, INGO/VAN DER AALST, WIL M. P. (2021): Process mining on blockchain data: a case study of Augur. In: POLYVYANYY, ARTEM/WYNN, MOE THANDAR/VAN LOOY, AMY/REICHERT, MANFRED (editors): Business Process Management. Springer, Cham, Switzerland, pp. 306-323.

HUANG, ZAN/KUMAR, AKHIL (2012): A study of quality and accuracy trade-offs in process mining. In: INFORMS Journal on Computing, volume 24, no. 2, pp. 311-327.

IBM CLOUDANT (2015): Why NoSQL? Your database options in the new non-relational world [WWW Document]. URL http://book.itep.ru/depository/nosql/AST-0146405_Why_NoSQL_IBM_Cloudant.pdf (last accessed: 14 May 2022).

IEEE TASK FORCE ON PROCESS MINING (2012): Process Mining Manifesto. In: DANIEL, FLORIAN/BARKAOUI, KAMEL/DUSTDAR, SCHAHRAM (editors): Business Process Management Workshops, Springer, Berlin, Heidelberg, Germany, pp. 169-194.

IFAC (2021): International Standard on Auditing 200 (ISA 200): Overall Objectives of the Independent Auditor and the Conduct of an Audit in Accordance with International Standards on Auditing. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 50-67.

IFAC (2021): International Standard on Auditing 210 (ISA 210): Agreeing the Terms of Audit Engagements. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 68-81.

IFAC (2021): International Standard on Auditing 220 (ISA 220): Quality Control for an Audit of Financial Statements. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 82-92.

IFAC (2021): International Standard on Auditing 230 (ISA 230): Audit Documentation. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 93-101.

IFAC (2021): International Standard on Auditing 240 (ISA 240): The Auditor's Responsibilities relating to Fraud in an Audit of Financial Statement. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 102-130.

IFAC (2021): International Standard on Auditing 300 (ISA 300): Planning an Audit of Financial Statements. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 165-172.

IFAC (2021): International Standard on Auditing 315 (ISA 315) (Revised 2019): Identifying and Assessing the Risks of Material Misstatement. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 173-246.

IFAC (2021): International Standard on Auditing 320 (ISA 320): Materiality in Planning and Performing an Audit. In: IAASB (editor): Handbook of International

Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 247-252.

IFAC (2021): International Standard on Auditing 330 (ISA 330): The Auditor's Responses to Assessed Risks. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 253-267.

IFAC (2021): International Standard on Auditing 450 (ISA 450): Evaluation of Misstatements Identified during the Audit. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 281-288.

IFAC (2021): International Standard on Auditing 500 (ISA 500): Audit Evidence. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 289-301.

IFAC (2021): International Standard on Auditing 505 (ISA 505): External Confirmations. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 309-315.

IFAC (2021): International Standard on Auditing 520 (ISA 520): Analytical Procedures. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 325-329.

IFAC (2021): International Standard on Auditing 530 (ISA 530): Audit Sampling. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 330-339.

IFAC (2021): International Standard on Auditing 550 (ISA 550): Related Parties. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 382-397.

IFAC (2021): International Standard on Auditing 560 (ISA 560): Subsequent Events. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 398-404.

IFAC (2021): International Standard on Auditing 700 (ISA 700) (Revised): Forming an Opinion and Reporting on Financial Statements. In: IAASB (editor): Handbook of International Quality Control, Auditing, Review, Other Assurance, and Related Services Pronouncements, 2021 Edition, Volume I, New York, pp. 489-522.

JANS, MIEKE (2009): Internal fraud risk reduction by data mining and process mining: framework and case study. PhD thesis, Diepenbeek, Hasselt University.

JANS, MIEKE (2012): Process mining in auditing: from current limitations to future challenges. In: DANIEL, F./BARKAOUI, K./DUSTDAR, S. (editors): Business Process Management Workshops. Springer, Berlin, Heidelberg, pp. 394–397.

JANS, MIEKE (2019): Auditor choices during event log building for process mining. In: Journal of Emerging Technologies in Accounting, volume 16, no. 2, pp. 59-67.

JANS, MIEKE/ALLES, MICHAEL/VASARHELYI, MIKLOS A. (2010): Process mining of event logs in auditing: Opportunities and challenges. SSRN Electronic Journal. URL https://ssrn.com/abstract=2488737, pp. 1-32.

JANS, MIEKE/LYBAERT, NADINE/VANHOOF, KOEN/VAN DER WERF, JAN MARTIJN E. M. (2010): Business process mining for internal fraud risk reduction: results of a case study. In: Proceedings of the International Research Symposium on Accounting Information Systems, Paris.

JANS, MIEKE/DEPAIRE, BENOÎT/VANHOOF, KOEN (2011): Does Process Mining Add to Internal Auditing? An Experience Report. In: HALPIN, TERRY/NURCAN, SELMIN/KROGSTIE, JOHN/SOFFER, PNINA/PROPER, ERIK/SCHMIDT, RAINER/BIDER, ILIA (editors): Enterprise, Business-Process and Information Systems Modeling. Springer, Berlin, Heidelberg, Germany, pp. 31-45.

JANS, MIEKE/ALLES, MICHAEL/VASARHELYI, MIKLOS A. (2012): Process mining of event logs in internal auditing: a case study [WWW Document]. URL https://www.researchgate.net/publication/268344444_Process_mining_of_event_logs_in_internal_auditing_a_case_study (last accessed: 1 May 2021).

JANS, MIEKE/ALLES, MICHAEL/VASARHELYI, MIKLOS A. (2013): The case for process mining in auditing: Sources of value added and areas of application. In: International Journal of Accounting Information Systems, issue 1, pp. 1–20.

JANS, MIEKE/ALLES, MICHAEL/VASARHELYI, MIKLOS A. (2014): A field study on the use of process mining of event logs as an analytical procedure in auditing. In: The Accounting Review, volume 89, no. 5, pp. 1751–1773.

KEBEDE, MUSIE (2015): Comparative Evaluation of Process Mining. Master thesis, University of Tartu, Estonia.

KERREMANS, MARC (2019): Market guide for process mining. Gartner Report.

KLINKMÜLLER, CHRISTOPHER/PONOMAREV, ALEXANDER/TRAN, AN BINH/WEBER, INGO/VAN DER AALST, WIL M. P. (2019): Mining blockchain processes: extracting process mining data from blockchain applications. In: International conference on business process management. Springer, Cham, Switzerland, pp. 71–86.

KPMG (2018): Process Mining - Find and Present Process Alternatives to Real Life Processes to Make Them more Efficient and Controls more Effective [WWW Document]. URL www.home.kpmg.com/be/en/home/insights/2017/09/process-mining.html (last accessed: 2 February 2020).

KRANAKIS, EVANGELOS (2013): Advances in Network Analysis and its Applications. Springer.

LEEMANS, SANDER J. J./FAHLAND, DIRK/VAN DER AALST, WIL M. P. (2013): Discovering block-structured process models from event logs containing infrequent behaviour. In: Business Process Management Workshops, 2013 International Workshops, Beijing, China, pp. 66-78.

LEONARDI, GIORGIO/STRIANI, MANUAL/CAVALLINI, ANNA/MONTANI, STEFANIA/QUAGLINI, SILVANA (2018): Leveraging Semantic Labels for Multi-level Abstraction in Medical Process Mining and Trace Comparison. In: Journal of Biomedical Informatics, volume 83, pp. 10-24.

LEVY, HOWARD B. (2021): A Proper Risk-Based Approach to the Search for Unrecorded Liabilities. In: The CPA Journal, November 2017 issue [WWW Document]. URL https://www.cpajournal.com/2017/11/21/proper-risk-based-approach-search-unrecorded-liabilities/ (last accessed: 3 June 2022).

LY, LINH T./RINDERLE-MA, STEFANIE/KNUPLESCH, DAVID/DADAM, PETER (2011): Monitoring Business Process Compliance Using Compliance Rule Graphs. In: MEERSMAN, ROBERT/DILLON, THARAM/HERRERO, PILAR/KUMAR, AKHIL/REICHERT, MANFRED/QING, LI/OOI, BENG.CHIN/DAMIANI, ERNESTO/SCHMIDT, DOUGLAS C./WHITE, JULES/HAUSWIRTH, MANFRED/HITZLER, PASCAL/MOHANIA, MUKESH (editors): On the Move to Meaningful Internet Systems. Springer, Crete, Greece, pp. 82-99.

MAGGI, FABRIZIO M./BOSE, R. P. J. C./VAN DER AALST, WIL M. P. (2012): Efficient Discovery of Understandable Declarative Process Models from Event Logs. In: Advanced Information Systems Engineering – Proceedings of the 24th International Conference, Gdansk, Poland, pp. 270-285.

MAGGI, FABRIZIO M./MARRELLA, ANDREA/MILANI, FREDRIK/SOO, ALLAR/KASELA, SILVA (2018): A user evaluation of automated process discovery algorithms [WWW Document]. URL https://www.researchgate.net/publication/338765603_A_User_Evaluation_of_Automated_Process_Discovery_Algorithms (last accessed: 17 June 2021).

MAITA, ANA R. C./MARTINS, LUCAS C./LÓPEZ PAZ, CARLOS R./RAFFERTY, LAURA/HUNG, PATRICK C. K./PERES, SARAJANE M./FANTINATO, MARCELO (2017): A systematic mapping study of process mining. In: Enterprise Information Systems, volume 12, no. 5, pp. 505-549.

MANNHARDT FELIX/DE LEONI, H. A./REIJERS, HAJO A./VAN DER AALST, WIL M. P. (2017): Data-Driven Process Discovery-Revealing Conditional Infrequent Behavior from Event Logs. In: International Conference on Advanced Information Systems Engineering, Springer, pp. 545-560.

MANNHARDT FELIX/KOSCHMIDER AGNES/BARACALDO NATHALIE/WEIDLICH MATTHIAS/MICHAEL, JUDITH (2019): Privacy-preserving process mining. In: Business Information System Engineering, volume 61, issue 5, pp. 595-614.

MANS, RONNY S./VAN DER AALST, WIL M. P./VANWERSCH, ROB J. B. (2015): Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes. Springer.

MARTIN, NIELS/FISCHER, DOMINIK A./KERPEDZHIEV, GEORGI D./GOEL, KANIKA/LEEMANS, SANDER J. J./RÖGLINGER, MAXIMILIAN/VAN DER AALST, WIL M. P./DUMAS, MARLON/LA ROSA, MARCELLO/WYNN, MOE T. (2021): Opportunities and challenges for process mining in organisations – results of a Delphi study. In: Business Information System Engineering, volume 63, issue 5, pp. 511-527.

MARUSTER, LAURA/WEIJTERS, A. J. M. M./VAN DER AALST, WIL M. P./VAN DEN BOSCH, ANTAL (2002): Process mining: Discovering direct successors in process logs. In: LANGE, STEFFEN/SATOH, KEN/SMITH, CARL H. (editors): Discovery Science. Springer, Berlin, Germany, pp. 364-373.

MENDLING, JAN/WEBER, INGO/VAN DER AALST, WIL M. P./VOM BROCKE, JAN/CABANILLAS, CRISTINA/DANIEL, FLORIAN/DEBOIS, SOREN/DI CICCIO, CLAUDIO/DUMAS, MARLON/DUSTDAR, SCHAHRAM/GAL, AVIGDOR/GARCÍA-BANUELOS, LUCIANO/GOVERNATORI, GUIDO/HULL, RICHARD/LA ROSA, MARCELLO/LEOPOLD, HENRIK/LEYMANN, FRANK/RECKER, JAN/REICHERT, MANFRED/REIJERS, HAJO A./RINDERLE-MA, STEFANIE/SOLTI, ANDREAS/ROSEMANN, MICHAEL/SCHULTE, STEFAN/SINGH, MUNINDAR P./SLAATS, TIJS/STAPLES, MARK/WEBER, BARBARA/WEIDLICH, MATTHIAS/WESKE, MATHIAS/XU, XIWEI/ZHU, LIMING (2018): Blockchains for business process management – challenges and opportunities. In: ACM Transactions on Management Information Systems, volume 9, no. 1, pp. 1-16.

MICHAEL, JUDITH/KOSCHMIDER, AGNES/MANNHARDT, FELIX/BARACALDO, NATHALIE/RUMPE, BERNHARD (2019): User-centered and privacy-driven process mining system design for IoT. In: International conference on advanced information systems engineering. Springer, Cham, Switzerland, pp. 194-206.

MOCHTY, LUDWIG (1985): Zur Automatisierung der Buchprüfung – Ein Bayes'sches Suchspiel-theoretisches Konzept. PhD Thesis, Technical University of Vienna, Austria.

MOCHTY, LUDWIG (2015): Journal Entry Network- und Geschäftsprozess-Analyse im Verbund: Ein dualer Revisionsansatz. In: Deggendorfer Forum zur digitalen Datenanalyse e. V. (editor): Prozessoptimierung mit digitaler Datenanalyse: Ansätze und Methoden, Erich Schmidt Verlag, Berlin, Germany, pp. 27-47.

MOCHTY, LUDWIG/WIESE, MICHAEL (2012): Die Netzwerkstruktur der Buchhaltung als Grundlage des risikoorientierten Prüfungsansatzes. In: Betriebswirtschaftliche Forschung und Praxis, volume 64, no. 5, pp. 479-507.

MOCHTY, LUDWIG/WIESE, MICHAEL/MAAS, ANNALOUISE (2022): Die doppelte Buchhaltung als graphentheoretisches Netzwerk. In: Der Wirtschaftstreuhänder - Fachjournal, WP Spezial, issue 5, pp. 372-378.

MÜHLBERGER, ROMAN/BACHHOFNER, STEFAN/DI CICCIO, CLAUDIO/GARCÍA-BANUELOS, LUCIANO/LÓPEZ-PINTADO, ORLENYS (2019) Extracting event logs for process mining from data stored on the blockchain. In: International conference on business process management. Springer, Cham, Switzerland, pp. 690-703.

MÜLLER-WICKOP, NIELS/SCHULTZ, MARTIN/GEHRKE, NICK/NÜTTGENS, MARKUS (2011): Towards automated financial process auditing: aggregation and visualization of process models. In: NÜTTGENS, MARKUS/THOMAS, OLIVER/WEBER, BARBARA (editors): Enterprise Modelling and Information Systems Architectures. Gesellschaft für Informatik e.V., Bonn, Germany, pp. 121-134.

MÜLLER-WICKOP, NIELS/SCHULTZ, MARTIN (2013): ERP Event Log Preprocessing: Timestamps vs. Accounting Logic. In: Design Science at the Intersection of Physical and Virtual Design, volume 7939, pp. 105-119.

NEUBAUER, THAIS R./FANTINATO, MARCELO/PERES, SARAJANE M. (2019): Interactive trace clustering. Conference paper, XV Simpósio Brasileiro de Sistemas de Informação, pp. 45-48.

NGUYEN, HOANG HUY/DUMAS, MARLON/LA ROSA, MARCELLO/MAGGI, FABRIZIO/SURIADI, SURIADI (2014): Mining business process deviance: a quest for accuracy. In: LIU, L./DILLON, T./PASTOR, O./SELLIS, T./PANETTO, H./MEERSMANN, R. et al. (editors): On the move to meaningful internet systems, volume 8841, Springer, Heidelberg, Germany, pp. 436-445.

OKOYE, KINGSLEY (2020): Applications and Developments in Semantic Process Mining. IGI Global, Hershey.

PCAOB (2004): Auditing Standard No. 2: An Audit of Internal Control Over Financial Reporting Performed in Conjunction with an Audit of Financial Statements. In: PCAOB (editor): PCAOB Release No. 2004-001. Washington, D. C, pp. 27-115.

PCAOB (2016): Report on 2015 Inspection of Ernst & Young LLP. In: PCAOB (editor): PCAOB Release No. 104-2016-142. Washington, D. C, pp. 1–66.

PCAOB (2018): Report on 2016 Inspection of Ernst & Young LLP. In: PCAOB (editor): PCAOB Release No. 104-2018-018. Washington, D.C., pp. 1–70.

PIKA, ANASTASIIA/WYNN, MOE T./BUDIONO, STEPHANUS/TER HOFSTEDE, ARTHUR H. M./VAN DER AALST, WIL M. P./REIJERS, HAJO A. (2020): Privacy-preserving process mining in healthcare. In: International Journal of Environmental Research and Public Health, volume 17, no. 5, pp. 1612-1640.

PUFAHL, LUISE/OHLSSON, BRIDIE/WEBER, INGO/HARPER, GARRETT/WESTON, EMMA (2021): Enabling financing in agricultural supply chains through blockchain. In: VOM BROCKE JAN/MENDLING, JAN (editors): Business process management cases, volume 2. Springer, Heidelberg, Germany, pp. 41-56.

PWC (2019): Process Intelligence - Optimising your Processes [WWW Document]. URL www.pwc.com/m1/en/events/iia2018/process-intelligence.pdf (last accessed: 11 May 2022).

RAMEZANI ELHAM/FAHLAND, DIRK/VAN DER AALST, WIL M. P. (2012): Where did I misbehave? Diagnostic information in compliance checking. In: BARROS, ALISTAIR/GAL, AVIGDOR/KINDLER, EKKART (editors): Business Process Management. Springer, Berlin, Heidelberg, Germany, pp. 262-278.

REBUGE, ÁLVARO/FERREIRA, DIOGO R. (2012): Business process analysis in healthcare environments: a methodology based on process mining. In: Information Systems, volume 37, issue 2, pp. 99-116.

REGIERUNGSKOMMISSION (2022): German Corporate Governance Code (as amended on 28 April 2022, convenience translation) [WWW Document]. URL https://dcgk.de//files/dcgk/usercontent/en/download/code/220627_German_Corporate_Governance_Code_2022.pdf (last accessed: 5 January 2023).

ROZINAT, ANNE (2010): Process Mining: Conformance and Extension. PhD thesis, Technical University of Eindhoven, Eindhoven, The Netherlands.

ROZINAT, ANNE/VAN DER AALST, WIL M. P. (2006): Conformance testing: Measuring the fit and appropriateness of event logs and process models. In: BUSSLER, CHRISTOPH J./HALLER, ARMIN (editors): Business Process Management Workshops, Nancy, France, pp. 163-176.

ROZINAT, ANNE/VAN DER AALST, WIL M. P. (2008): Conformance checking of processes based on monitoring real behavior. In: Information Systems, volume 33, issue 1, pp. 64-95.

ROZINAT, ANNE/DE MEDEIROS, ALVES K. A./GÜNTHER, CHRISTIAN W./WEIJTERS, A. J. M. M./VAN DER AALST, WIL M. P. (2008): The need for a process mining evaluation framework in research and practice. In: HOFSTEDE, ARTHUR/BENATALLAH, BOUALEM/PAIK, HYE-YOUNG (editors): Business Process Management Workshops. Springer, Brisbane, Australia, pp. 84-89.

SAP (2022): Shopping cart. In: SAP Support Wiki [WWW Document]. URL https://wiki.scn.sap.com/wiki/display/SRM/Shopping+cart (last accessed: 2 February 2023).

SCHIMM, GUIDO (2002): Process Miner: A Tool for Mining Process Schemes from Event-based Data. In: FLESCA, SERGIO/GRECO, SERGIO/IANNI, GIOVAMBATTISTA/LEONE, NICOLA (editors): Logics in Artificial Intelligence, Springer, Cosenza, Italy, pp. 525-528.

SCHIMM, GUIDO (2003): Mining most specific workflow models from event-based data. In: VAN DER AALST, WIL M. P./WESKE, MATHIAS (editors): Business Process Management, Springer, Eindhoven, The Netherlands, pp. 25-40.

SONG, MINSEOK/GÜNTHER, C. W./VAN DER AALST, WIL M. P. (2008): Trace clustering in process mining. In: ARDAGNA, D./MECELLA, M./YANG, J. (editors): Business Process Management Workshops (BPM 2008). Lecture Notes in Business Information Processing, volume 17. Springer, Berlin, Heidelberg, Germany, pp. 109-120.

TIWARI, ASHUTOSH/TURNER, CHRIS/MAJEED, BASIM (2008): A review of business process mining: state-of-the-art and future trends. In: Business Process Management Journal, volume 14, issue 1, pp. 5-22.

UIPATH (2021): UiPath Process Mining Guide [WWW Document]. URL https://docs.uipath.com/process-mining/docs/overview-dashboard-role-analyst (last accessed: 8 July 2022).

VAN BEEST, NICK R. T. P./DUMAS, MARLON/GARCÍA-BANUELOS, LUCIANO/LA ROSA, MARCELLO (2015): Log Delta Analysis: Interpretable Differencing of Business Process Event Logs. In: MOTAHARI-NEZHAD, HAMID/RECKER, JAN/WEIDLICH, MATTHIAS (editors): Business Process Management, Springer, Innsbruck, Austria, pp. 386-405.

VAN DER AALST, WIL M. P. (2005): Business alignment: using process mining as a tool for delta analysis and conformance testing. In: Requirements Engineering Journal, volume 10, no. 3, pp. 198-211.

VAN DER AALST, WIL M. P. (2010): Process discovery: capturing the invisible. In: Computational Intelligence Magazine, volume 5, issue 1, pp. 28-41.

VAN DER AALST, WIL M. P. (2016): Process Mining: Data Science in Action, 2nd edition. Springer, Berlin, Germany.

VAN DER AALST, WIL M. P./VAN DONGEN, BOUDEWIJN F./HERBST, JOACHIM/MARUSTER, LAURA/SCHIMM, GUIDO/WEIJTERS, A. J. M. M. (2003): Workflow Mining: A survey of issues and approaches. In: Data & Knowledge Engineering, volume 47, issue 2, pp. 237-267.

VAN DER AALST, WIL M. P./WEIJTERS, A. J. M. M. (2004): Process Mining: A Research Agenda. In: Computers in Industry, volume 53, pp. 231-244.

VAN DER AALST, WIL M. P./WEIJTERS, A. J. M. M./MARUSTER, LAURA (2004): Workflow mining: Discovering process models from event logs. In: IEEE Transactions on Knowledge and Data Engineering, volume 16, no. 9, pp. 1128-1142.

VAN DER AALST, WIL M. P./DE BEER, HUUB T./VAN DONGEN, BOUDEWIJN F. (2005): Process mining and verification of properties: an approach based on temporal logic. In: MEERSMAN, ROBERT/TARI, ZAHIR (editors): On the Mode to Meaningful Internet Systems. Springer, Heidelberg, Germany, pp. 130-147.

VAN DER AALST, WIL M. P./DE MEDEIROS, ALVES A. K. (2005): Process Mining and Security: Detecting Anomalous Process Executions and Checking Process Conformance. In: Electronic Notes in Theoretical Computer Science, volume 121, pp. 3-21.

VAN DER AALST, WIL M. P./DE MEDEIROS, ALVES A. K./WEIJTERS, A. J. M. M. (2005): Genetic process mining. In: CIARDO, GIANFRANCO/DARONDEAU, PHILIPPE (editors): Applications and Theory of Petri Nets 2005. Springer, Miami, USA, pp. 48-69.

VAN DER AALST, WIL M. P./VAN HEE, KEES M./VAN DER WERF, JAN M./VERDONK, MARC (2010): Auditing 2.0: using process mining to support tomorrow's auditor. In: Computer, volume 43, no. 3., pp. 90-93.

VAN DER WERF, JAN MARTIJN E. M./VAN DONGEN, BOUDEWIJN F./HURKENS, COR A. J./SEREBRENIK, ALEXANDER (2009): Process discovery using integer linear programming. In: Applications and Theory of Petri Nets, volume 94, no. 3-4, pp. 387-412.

VAN DONGEN, BOUDEWIJN F./DE MEDEIROS, ALVES A. K./VERBEEK, HENRICUS M. W./WEIJTERS, A. J. M. M./VAN DER AALST, WIL M. P. (2005): The ProM Framework: a new era in process mining tool support. In: 26th International Conference on Application and Theory of Petri Nets. Springer, Heidelberg, Germany, pp. 444-454.

VAN ZELST, SEBASTIAN J./VAN DONGEN, BOUDEWIJN F./VAN DER AALST, WIL M. P./VERBEEK, HENRICUS M. W. (2017): Discovering workflow nets using integer linear programming. In: Computing, volume 100, issue 5, pp. 529-556.

VOM BROCKE, JAN/JANS, MIEKE/MENDLING, JAN/REIJERS, HAJO A. (2021): A five-level framework for research on process mining. In: Business and Information Systems Engineering, volume 63, issue 5, p. 483-490.

WEIJTERS, A. J. J. M./VAN DER AALST, WIL M. P. (2001): Rediscovering Workflow Models from Event-based Data. In: Proceedings of the 11th Dutch-Belgian Conference on Machine Learning, pp. 93-100.

WEIJTERS, A. J. J. M./VAN DER AALST, WIL M. P. (2003): Rediscovering Workflow Models from Event-based Data using little thumb. In: Integrated Computer-Aided Engineering, volume 10, no. 2, pp. 151-162.

WEIJTERS, A. J. J. M./VAN DER AALST, WIL M. P./DE MEDEIROS, ALVES K. A. (2006): Process mining with the heuristics miner-algorithm. Technische Universiteit Eindhoven, volume 166, pp. 1-34.

WEIJTERS, A. J. M. M./RIBEIRO, JOEL T. S. (2011): Flexible Heuristics Miner. In: IEEE Symposium on Computational Intelligence and Data Mining, pp. 310-317.

WEN, LIJIE/VAN DER AALST, WIL M. P./WANG, JIANMIN/SUN, JIAGUANG (2007): Mining process models with non-free choice constructs. In: Data Mining and Knowledge Discovery, volume 15, no. 2, pp. 145-180.

WERNER, MICHAEL (2016): Process Model Representation Layers for Financial Audits. In: 49th Hawaii International Conference on System Sciences, Koloa, USA pp. 5338-5347.

WERNER, MICHAEL (2017): Financial process mining: accounting data structure dependent control flow inference. In: International Journal of Accounting Information Systems, volume 25, pp. 57-80.

WERNER, MICHAEL/GEHRKE, NICK/NÜTTGENS, MARKUS (2012): Business process mining and reconstruction for financial audits. In: Proceedings of the 45th Hawaii International Conference on System Science, Maui, pp. 5350-5359.

WERNER, MICHAEL/GEHRKE, NICK (2019): Identifying the absence of effective internal controls: An alternative approach for internal control audits. In: Journal of Information Systems, volume 33, issue 2, pp. 205-222.

WERNER, MICHAEL/WIESE, MICHAEL/MAAS, ANNALOUISE (2021): Embedding process mining into financial statement audits. In: International Journal of Accounting Information Systems, volume 41, pp. 1-15.

WIESE, MICHAEL (2013): Aussagebezogene Abschlussprüfung, Düsseldorf: IDW-Verlag.

WYNN, MOE THANDAR/OUYANG, CHUN/TER HOFSTEDE, ARTHUR H. M./FIDGE, COLIN J. (2009): Workflow Support for Product Recall Coordination. BPM Center Report, Queensland University of Technology: Business Process Management Group.

YANG, WAN-SHIOU/HWANG, SAN-YIH (2006): A process-mining framework for the detection of healthcare fraud and abuse. In: Expert Systems with Applications, volume 31, issue 1, pp. 56-68.

ZHANG, SHAO-HUA/GU, NING/LIAN, JIE-XIN/LI, SAIHAN (2003): Workflow process mining based on machine learning. In: Proceedings of the 2nd International Conference on machine Learning and Cybernetics, IEEE Computer Society, volume 4, Washington D. C., pp. 2319-2323.

Appendix I: Glossary

Activity: Activities are the different types of procedures performed in the process. For example, the approval of a purchase order represents an activity. This single activity might have been performed once for 90 cases, resulting in 90 events.

Activity type: Activities may be grouped into activity types. For example, the activity type "purchase order modification" summarizes all activities that deal with changes made to purchase orders, regardless of the specific field that has been changed.

Case: A case (also referred to as "process instance") summarizes one execution of the process and represents the document that is followed through the process. In the implementation described in this thesis, the case is a purchase order line item (for SAP MM transactions) or an invoice line item (for SAP FI transactions). Thus, all different events that are performed for an individual line item make up one case.

Case type: Case types are used to differentiate cases into different categories. For example, cases that represent a purchase order line item may be assigned to the case type "case with purchase order", whereas cases representing an invoice line item are summarized as "case without purchase order". By this, procedures may be designed explicitly for specific case types only.

Edge: An edge in the process graph is the path that connects two activities. The more cases went through the same process path, the thicker the edge in the process graph.

Effective date: The effective date or posting date is the date on which the transaction becomes effective to the financial statements and is a mandatory field required by the bookkeeping system. When entering the transaction, the effective date may be set to the same date as the entry date of the transaction or, alternatively, may be future- or back-dated.

Entry date: The entry date is automatically applied by the system and is the date on which the transaction is entered in the system.

Event: An event describes an activity in the process that is performed by a specific user for a certain case at a particular point in time. For example, the execution of the activity "purchase order approval" by user FMB for purchase order PO-100 at the 2nd of February 2022.

Event log: An event log is a chronological record of activities performed in an information system that is derived from the log files recorded by the system.[769] In general, an event log is related to a separated process, for example the purchase to pay process of an entity. Information contained in the event log includes, for example, the date, time and user of the activity performed and a reference to the related case.

Event time: The event time is the time at which the event takes place. All event times are directly extracted from the ERP system.[770]

Loop: Loops in the process graph represent activities that are being repeated without any other activity being executed in between.

Process instance: The terms "process instance" and "cases" are used synonymously. Refer to the definition of a "case".

Process graph: A process is a graphical and analytical representation of all events that are performed for a defined set of cases, grouped into activities and ordered by the timestamp of events. The main objective of the process mining algorithms is reconstructing how the process occurred.

Process mining: Process mining is a technology that extracts process related data from an information system in order to reconstruct the process flow as it actually occurred.[771]

Variation: Process instances that follow the same activities in exactly the same sequence are summarized in a process variation, i.e., one specific process path ("trace"). Thus, a process variation includes all cases for which the process has been executed in exactly the same way.

[769] Cf. JANS, MIEKE/ALLES, MICHAEL G./VASARHELYI, MIKLOS A. (2010), p. 3.
[770] Cf. Chapter 4.1.5.
[771] Cf. VAN DER AALST, WIL M. P. (2016), p. 25.

Appendix II: Overview dashboard of the audit firm's process mining application

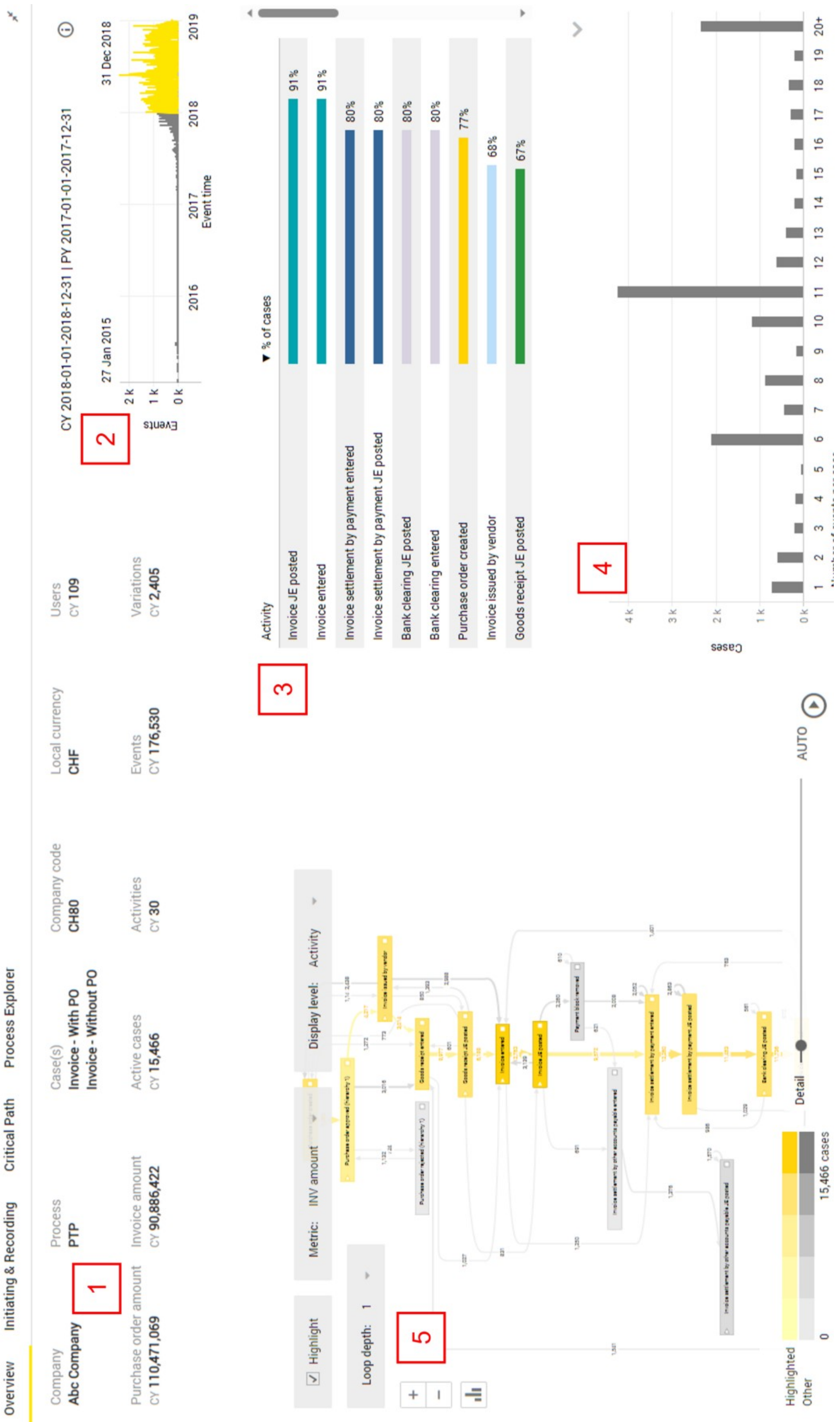


Figure 87: Overview dashboard of the process mining application

Key indicators (1): From the key indicators the auditor can obtain an initial understanding about the nature and complexity of the process. Besides general information, such as the company's name and its local currency, the case information indicates that both invoices related to a purchase order and direct invoices without a purchase order are processed. Using the key indicators, the auditor may determine if the number of cases, the total amounts for purchase orders and invoices, the number of activities and events and the number of users involved in the process are in line with the understanding of the entity's business and organizational structure.

Event timeline (2): The figure at the top-right side of the dashboard summarizes the distribution of events over the observation period. It may be used to determine on a high level if activities have been performed throughout the entire period or if there are gaps or peaks at certain points in time requiring further investigation. The events related to the audit period are colored yellow. Gray coloring is used for any events performed before or after the audit period that result due to the different lengths of the data extraction strategy period and the audit period.

Activity chart (3): From the chart of activities, the auditor obtains an understanding of the activities performed within the process and their frequencies, i.e., the percentage of cases for which an activity has been performed at least once. The activities are colored according to their activity type. For example, all activities dealing with goods receipt processing are colored green. The activities provide an initial overview of the process composition and quality. While the chart of activities especially supports the initial process understanding, a reasonableness test of the case coverage may be performed as part of data validation.

Events per case (4): The distribution of events per case summarizes the number of events within each case and supports understanding and validating the process composition. In the example provided, cases related to a purchase order usually contain eleven events while the majority of direct purchases consists of six events.

Process graph (5): The process graph on the overview dashboard may be used to obtain an initial understanding of the process execution on a high level of aggregation, including the identification of the most common process path (colored yellow) and the order of activities within.

Appendix III: Structure and key analyses of the audit firm's process mining application in 2018

Menu page	Dashboard	Description of analyses
Financial reconcili-ation	Account coverage	Identify the transaction volume of financial accounts that is covered/ not covered through the process data.
	General ledger activity	Analyze the general ledger data and identify the root cause of non-covered transaction volume.
	Booking pattern	Analyze the journal entries initiated by recording activities and identify related debited/credited accounts.
Under-stand process	Overview	Obtain an overview of the dataset, including KPIs, activities and the distributions of cases, events and variations.
	Open/Close	The dashboard supports analyzing the first and the last activity of cases in the dataset.
	Process explorer	Analyze the process graph visualizing the sequence of activities as re-constructed from the event log. The dashboard may be used as the starting point for root-cause analyses on any items selected for further investigation.
	Critical path	Understand and evaluate process variations including their number and volume of cases. A process graph visualizes the process trace for the variation currently selected.
Controls evalua-tion	Compare	Compare the sequence of activities in two different process variations. Customize the dashboard to compare, for example, the processing of cases for different material groups or vendors only.
	Throughput time	Analyze the duration of activities and the time between the end of an activity and the start of the next activity.
	Activity sequence	Investigate cases including a particular sequence of activities, for example, cases where a purchase order has not been approved before the posting of the vendor invoice.
	Segregation of duties	Identify the cases, activities and users for which segregation of incom-patible duties is not maintained.
	Three-way-match	Reperform the price and quantity match between the purchase order, goods receipt and invoice documents.
User involve-ment	Activities	Analyze roles and responsibilities by investigating the activities per-formed by users or departments.
	Cases	Investigate which individuals are involved in which cases. Information is provided on the cases for which specific activities have been per-formed by a particular user.
Vendor profile	Vendor structure	The dashboard lists the vendors that are related to the process. The re-lated number of cases and the purchase order and invoice volume help to understand the vendor structure and changes therein.
	Vendor master data	The dashboard contains the vendor master data extracted from the sys-tem.
	Master data changes	The dashboard includes detailed information on changes made to mas-ter data fields.
Details	Cases and events	The dashboard lists the cases and events in the process, cf. Table 1 and Table 2.
	Document details	The dashboard contains the detailed header and line item information of all documents related to cases in the process.
	Journal en-tries	The dashboard includes all journal entries that are initiated by cases in the process.

Table 28: Structure of the audit firm's process mining application in 2018

Appendix IV: Implementation of the trace clustering prototype on the critical path dashboard

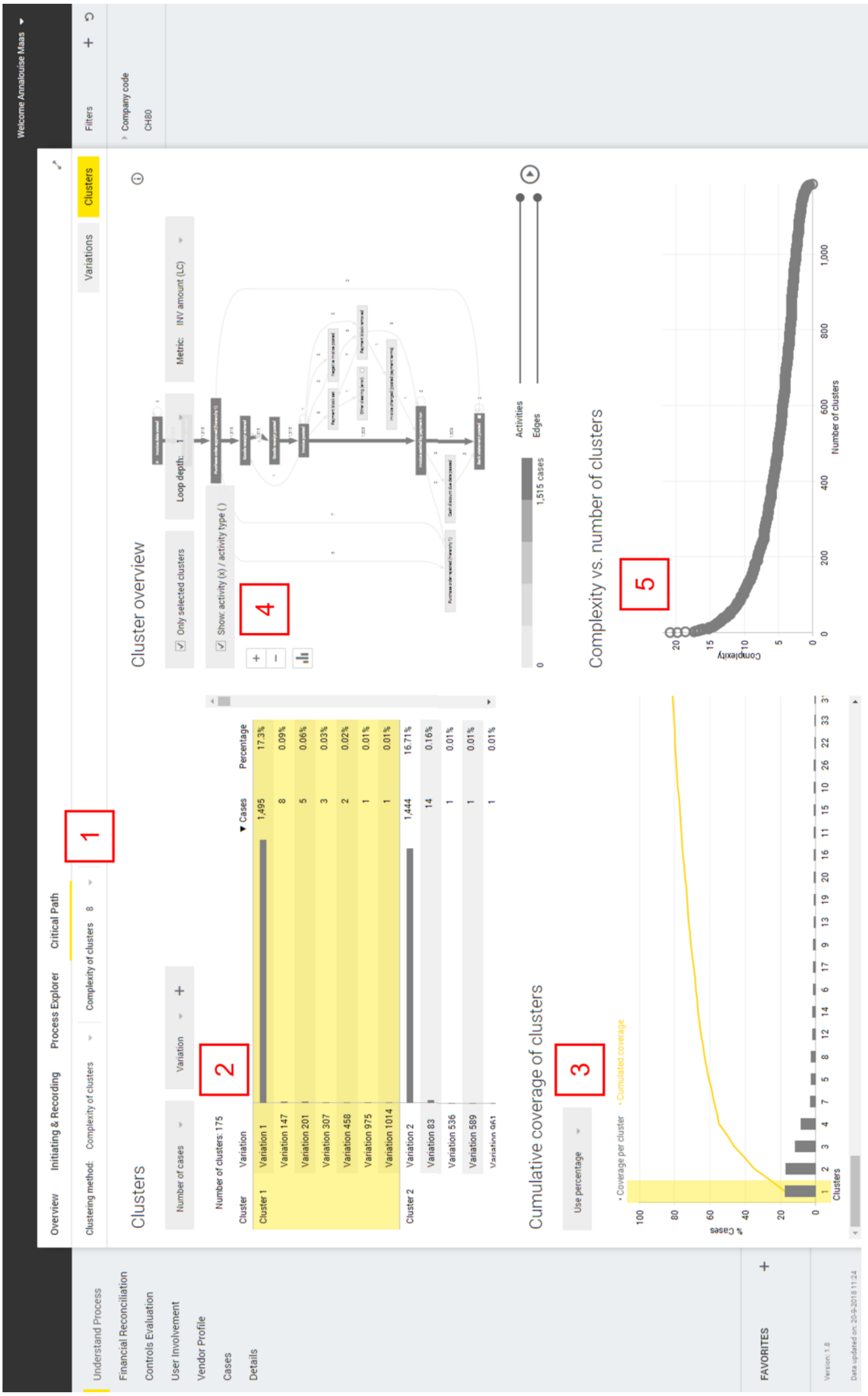


Figure 88: Implementation of the trace clustering prototype on the critical path dashboard

Clustering process variations may help the auditor in dealing with the extensive number of process variations that are usually resulting in practice. Instead of individually investigating each process variations separately, the auditor may consider auditing groups of variations that have a similar behavior. The clustering prototype enables to group variations with similar patterns into a cluster with the aim of reducing the number of process paths to analyze.

Definition of the clustering method: (1): The options at the top of the dashboard can be used to adjust the clustering method. With these filter criteria, the auditor determines how similar variations are grouped into a cluster to reduce the process complexity. There are two different clustering methods available:

- **Number of clusters:** The auditor may determine a target number of clusters that is being used to group the variations. The behavior of the variations within each cluster is similar to each other. In case a low number of clusters is selected, more variations are grouped together into one cluster. Consequently, as the number of clusters decreases, the clusters' complexity (that is, the heterogeneity of process paths included in the cluster) increases.
- **Complexity of clusters:** The auditor may determine a maximum complexity value. A low complexity within the clusters will result in a higher number of clusters. A higher complexity will lead to less, but more complex clusters.

Cluster bar chart (2): The cluster bar chart may be used to validate the resulting number of clusters, their structure and the variations included in each cluster. The "cluster" column lists all clusters that have been generated. The "variation" column shows all variations within a cluster. The table can either show the number of cases or the monetary amount associated to each variation. By setting the metric to the invoice amount, the auditor may investigate the materiality of the resulting clusters. In case a cluster is still very complex, the auditor can filter this cluster and analyze the significant variations included individually.

Cumulative coverage of clusters (3): The cumulative case coverage of clusters is used to determine how many cases (as a percentage of all cases) are covered by each cluster. The line illustrates the cumulated case coverage. In the example, approximately 50 percent of all cases may be covered by investigating the first three clusters.

Cluster overview (4): The process graph on the dashboard displays the process paths of the cases included in the cluster currently selected in the cluster bar chart (2).

Complexity vs. number of clusters (5): The chart at the bottom right of the dashboard shows the relation between the complexity and the number of clusters. It is used to facilitate determining an appropriate clustering method (1) suitable for the dataset, as with rising complexity, the evaluation and interpretation of the clusters from a business perspective becomes more complex. The auditor can obtain an understanding of how the complexity within the clusters behaves with regard to the number of clusters in order to specify the number or complexity that should be used to cluster the variations.

Appendix V: Implementation of the comparison functionality on the overview dashboard

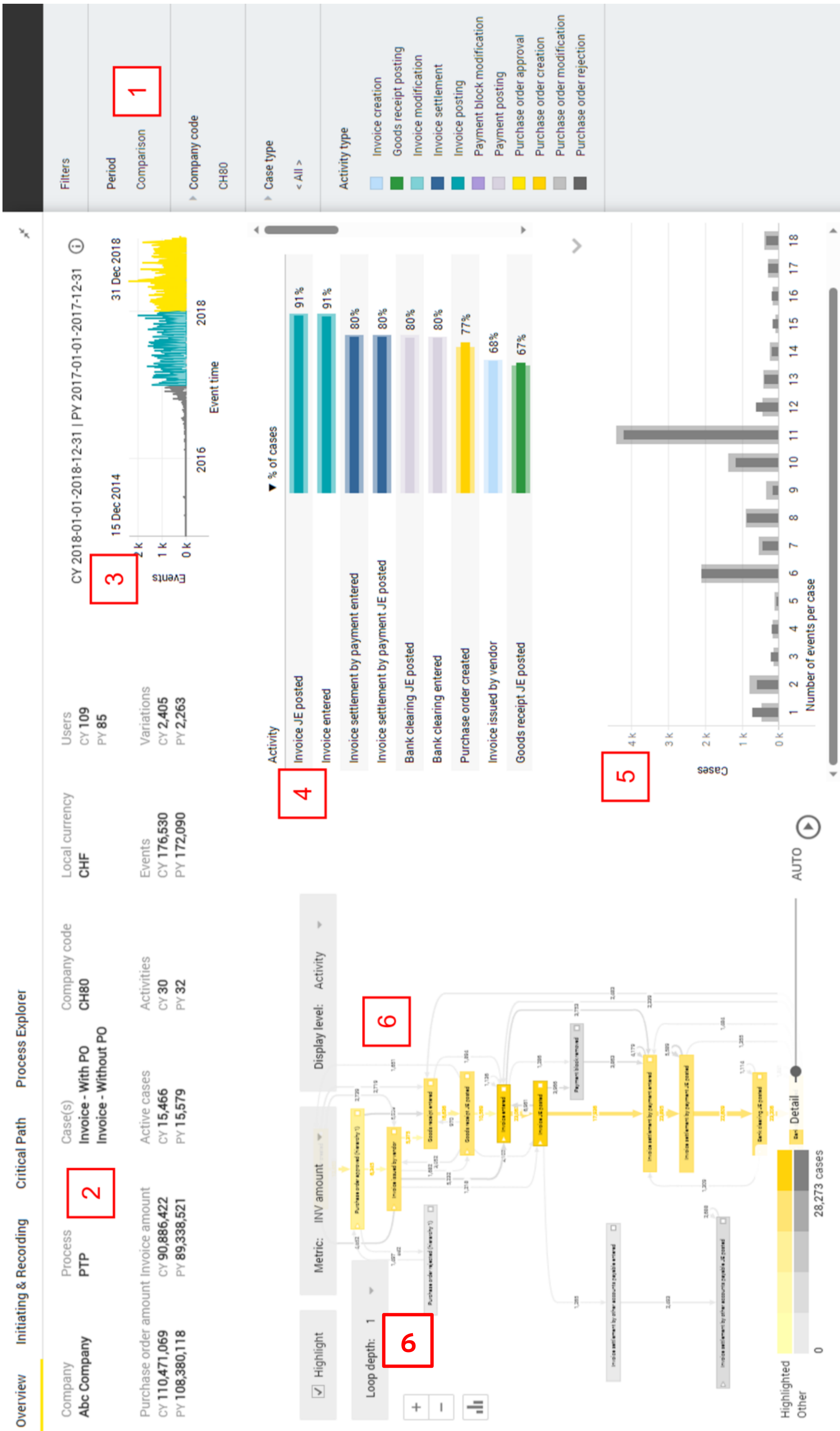


Figure 89: Implementation of the comparison functionality on the overview dashboard

Using the “Overview” dashboard of the audit firm’s process mining application as an example, Figure 89 illustrates how process mining may be enhanced with a comparison functionality. The general functionalities of the “Overview” dashboard are described in Appendix II.

Comparison functionality (1): The comparison feature can be activated using the “period” filter in the global filter pane on the right side of the analyzer. Here, the option “current period” can be selected to display the data from the current period, “comparative period” selects the data from the comparative period. If “comparison” is selected, all analyses in the application are filtered to display the data from the audit period in comparison to the data of the comparative period.

Key indicators (2): Comparative information may be added to the key indicators by adding respective labels (for example, “CY” for the current year, “PY” for the previous year) to the indicators that have changed between both periods. In the example provided, the total purchase order volume, the invoice amount and the number of cases processed in both periods did not change significantly. On the other hand, the auditor may determine to further investigate the two activities that are no longer performed in the audit period under review and consider the new users involved in processing transactions in the audit period when analyzing roles and responsibilities in the process.

Event timeline (3): In the comparative view, the event timeline is adjusted to display events related to the audit period in yellow and events related to the comparative period in turquoise. Gray coloring is used for any events performed before the comparative period or after the audit period that result due to the different lengths of the data extraction strategy period and the audit period.

Activity chart (4): Comparative information from the previous period is added to the activity chart using slightly thicker and transparent bars in the background. A tooltip displays the relative change of the cases for which the respective activity has been performed in the audit period compared to the previous period. Using the chart of activities, the auditor may identify significant fluctuations, new activities in the audit period as well as activities that have only been performed in the previous period.

Events per case (5): Similar to bar charts, in column charts, lighter-colored bars in the background of the audit period’s data display the data of the comparative period.

Process graph (6): If the “comparison” view is selected in the global filter pane (1), the process graph displays the total number of cases processed in both periods. By filtering the data to the individual periods (and by comparing the individual process variations as described in Chapter 4.4.2) the auditor may identify significant differences in the processing of cases. In the future, a highlighting functionality may be integrated for new process paths occurring in the audit period under review that have not been used in the comparative period.

Eidesstattliche Erklärung

Ich gebe folgende eidesstattliche Erklärung ab:

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig ohne unzulässige Hilfe Dritter verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und alle wörtlich oder inhaltlich übernommenen Stellen unter der Angabe der Quelle als solche gekennzeichnet habe.

Die Grundsätze für die Sicherung guter wissenschaftlicher Praxis an der Universität Duisburg-Essen sind beachtet worden.

Ich habe die Arbeit keiner anderen Stelle zu Prüfungszwecken vorgelegt.

Ort, Datum	Unterschrift