

## A Case Study on the Application of Process Mining in Combination with Journal Entry Tests for Financial Auditing

Sebastian Stephan  
 German Research Center for  
 Artificial Intelligence (DFKI),  
 Saarland University, Germany  
[Sebastian.Stephan@dfki.de](mailto:Sebastian.Stephan@dfki.de)

Johannes Lahann  
 German Research Center for  
 Artificial Intelligence (DFKI),  
 Saarland University, Germany  
[Johannes.Lahann@dfki.de](mailto:Johannes.Lahann@dfki.de)

Peter Fettke  
 German Research Center for  
 Artificial Intelligence (DFKI),  
 Saarland University, Germany  
[Peter.Fettke@dfki.de](mailto:Peter.Fettke@dfki.de)

### Abstract

*The increasing complexity of data and processes within companies makes it increasingly difficult for auditors to ensure that annual audits are free of material misstatement. To cope with this complexity, a variety of analytical procedures have been developed in the last years. However, most of the existing procedures focus on conspicuous statements in the general ledger, and thus not consider behavioral aspects. In this paper, we show how journal entry tests can be effectively combined with process mining to capture a more comprehensive view within a company's audit. Therefore, the paper gives a comprehensive description of the purchase-to-pay-process and its realization in current SAP software as well as the required mechanism to extract event logs from raw SAP database tables. The conducted analysis is based on a dataset provided by a German medium-sized audit firm. The results suggest that we can discover anomalies that are not traceable through traditional analysis.*

### 1. Introduction

The advancing digitization has enabled companies to achieve many efficiency gains. A multi-layered system landscape and infrastructure has been created, which allows many business processes to be accelerated and automated [1]. In addition, large quantities of data are generated, recorded and evaluated every day. This poses new challenges for financial auditing. In order to detect financial accounting irregularities, the auditor must understand the business processes and its internal controls of the company [2, 3]. Due to the increasing data and process complexity within a company, the audit process can not be executed manually anymore. According to the International Standard on Auditing (ISA) 330, tests of controls and substantive procedures are an integral part of the audit in order to reveal material misstatements [4]. Over the last decade, there have been many publications in accounting research literature that

have developed new analysis procedures or optimized existing approaches. In general, journal entry tests are applied exclusively on the data stored in the general ledger of an enterprise resource planning system (ERP) [5]. However, context and process information on how the financial statements are formed is usually not taken into account.

Process mining is a method to understand how business processes operate. It allows a fact-based analysis of the underlying process along different perspectives [6]. Using an event log as input, process mining focuses on the reconstruction of the time-logical sequence of the recorded events in general. As such, process mining appears to be a suitable technique for auditors within the risk-oriented audit to enhance the data-driven analysis focusing on statistical quantities with a process-oriented perspective. In future, this will be more crucial for auditors because ISA 315 states that the auditor should obtain an understanding of the "[...] information system, including related business processes, relevant to financial reporting"[7, p. 292]. However, the existing literature in this context is rather limited. In particular, the question of how journal entry tests and process mining can be efficiently combined for auditing purposes has not been addressed yet. This work tries to close the research gap by providing a case study based on the purchase-to-pay-process. The focus was set to this process since its creditor part has a considerable influence on the financial statements of a company. For example, payments describe a risky area because they cause an outflow of money, but are not always recorded in the ERP system directly. Accordingly, it can be of interest to filter out these cases and map them along the whole business process in order to understand how the ERP as a system, but also the process itself works. Based on a real-world SAP dataset provided by a German medium-sized audit firm, we combine traditional journal entry tests with process mining techniques in order to analyze the process behavior along identified conspicuous account postings. All necessary prerequisites in order to apply

journal entry tests and process mining to a SAP-based infrastructure will be explained.

Section 2 discusses literature of process mining and journal entry tests that have been applied in the financial auditing domain. Section 3 introduces the reader to required background information and describe the conducted case study. Section 4 focuses on the event log extraction from raw SAP database tables. In Section 5, process mining and journal entry tests are combined and discussed. Section 6 discusses the key findings and limitations of the presented work. Section 7 concludes with a summary on the major results and an outline of future work.

## 2. Related Work

Since this work is a practical case study based on several closely connected works from different areas, the related work is divided into three subsections.

### 2.1. Event Log Extraction from ERP-Systems

In order to apply process mining on ERP data, an event log has to be extracted. In recent years, several approaches have been developed involving different trade-offs. One of the first event log extraction approach focused mainly on SAP transaction data using data from the change document header *CDHDR* and position table *CDPOS* [8]. The data was then combined with ontological information using the tables with the transaction codes *TSTC* and their descriptions *TSTCT* to reconstruct activities. The authors developed the *EVS ModelBuilder* to automate the extraction of event logs from a SAP system. However, since this approach lacks with focus on one specific business process in detail, in practice it often leads to cluttered so-called *spaghetti processes*. In contrast, in [9] the authors tried to reconstruct the purchase-to-pay-process by starting from a reference process model and deriving all required SAP tables that are relevant for the process. One major challenge the authors have mentioned was the determination of the *caseid* that can not directly be interfered from the SAP tables. While both papers show concrete solutions how to create an event log, the benefit, which results from process mining, is only discussed in very general terms and only from a control-flow perspective. On the contrary, Jans et al. developed a general procedure for the extraction of event logs, which is not limited to SAP data, but can also be applied to all types of relational databases [10].

The focus of this paper is not to develop a new event log extraction method. Instead, it combines the above described approaches and utilizes them in a realistic environment, pointing out existing issues and potential

pitfalls while generating an appropriate event log.

### 2.2. Process Mining in Audit

Based on the extracted event log from an ERP system, multiple methods have been proposed utilizing process mining for financial auditing. In [6], the authors introduced a framework utilizing process mining based on historic data, existing process models and current data, which was also applied for auditing purposes. Another approach focuses on a process mining implementation featuring a comprehensive visualization that provides the user with the control-flow as well as specific audit relevant information [11, 12]. In [13], the same authors introduce a variation of their method that is specifically designed to deal with the high complexity of mined process models for audit. In [5], the authors elaborate how process mining can be combined effectively with the control and accounting mechanisms within ERP systems. In [14], the control-flow perspective and the data perspective of process mining is utilized together to gain new insights for audit analysis. In [15], an alternative process mining approach based on account postings is introduced. Here, the approach uses instead of the temporal order of events the accounting specific dependencies. The results suggest that this method allows to extract cleaner and more comprehensive process models. In [16], the potential of process mining in audits and current limitations and future challenges are stated. To back the claim, the authors elaborated the value added by process mining methods to auditing in a field study using procurement data from a global bank [17, 18]. In a further work, Jans and Hosseinpour develop an actionable framework to address the transaction verification level of continuous auditing incorporating data mining and process mining techniques as well as the expertise of an auditor [19].

In contrast to the existing methods, this paper does not apply process mining on its own. By combining process mining with journal entry tests, we focus on incorporating both the control-flow as well as the data-perspective within the risk-oriented audit to capture a more comprehensive picture of the control environment of a company.

### 2.3. Journal Entry Testing

In addition to the above described idea of applying process mining techniques to support financial audits, there is a big branch of computer assisted audit techniques (CAAT) focusing mainly on the analysis of journal entries to detect material misstatements and potential conspicuous transactions. Most existing

approaches that find application in current software products are based on handcrafted rules implementing knowledge of auditing regulations and best practices [20, 21]. Furthermore, there are a few approaches utilizing statistical methods such as Benford's law [22] and extreme value theory [23]. In recent years, implementations utilizing deep learning for journal entry tests led to promising results [24, 25].

Journal entry tests have the disadvantage that during the financial audit the analysis is only based on entries of the general ledger. Therefore, it is only able to detect certain kind of problems isolated. By combination with process mining, an integrated view with regard to the entity's environment, risks, operational processes and internal controls will be supported. In sum, the contribution of this research paper lies in the practical use and how process mining and journal entry tests can be combined on a SAP-ERP database.

### 3. Case Study: Context and Background

Audit firms are striving to close the gap between the increasing complexity of data and business processes and the use of innovative technologies to ensure economic efficiency [26]. For this reason, a real-world SAP database export was made available at the end of 2019 from a medium-sized German audit firm in which personal data attributes were anonymized. Since accounting-relevant business transactions are of special interest for an auditor, the focus of this case study was initially limited to the purchase-to-pay-process in order to research the potentials of journal entry tests in combination with process mining.

#### 3.1. Fundamentals of the Risk-Oriented Audit

The risk-oriented audit approach has to be distinguished from the full audit, as the focus of this approach relies on generating an understanding of the business, risks, operational processes and internal controls. To obtain a sufficient audit evidence, appropriate responses to the assessed risks have to be implemented [7, 4]. With its focus on implementing substantive procedures, different audit procedures were applied based on the audit risk. For example, along the identified risky areas, the auditor inquires about controls in place to mitigate those risks. Afterwards, those controls in place were tested and subsequently re-assessed. Finally, according to ISA 330, tests like substantive procedures were performed on the risky areas in order to determine the adequacy and applicability of internal controls as well as to identify conspicuous financial statements, respectively. The scope of the test may vary based on the risk rating [2, 3].

#### 3.2. Challenges of the Risk-Oriented Audit

In order to assess the risks of material misstatement, the auditor needs a comprehensive understanding of the company's business processes and internal controls [7]. However, the quality of identifying risky areas depends on the effectiveness of the audit procedures used. The risk-based audit approach is based on the expectation that business transactions and risks have undergone a number of internal controls before the produced data is included in the financial statements. Therefore, the auditor must know and understand the underlying processes end-to-end in order to be able to test financial statements and to understand their composition [2, 3]. However, due to increasing amounts of data, traditional approaches such as inquiry, observation or examinations to determine the audit risk, documentation and reporting become increasingly ineffective and inefficient. Another challenge is the fact that the quality of identified risks depend on the questions asked. In general, business transactions that are not explicitly questioned by the auditor might not be considered during audit. While journal entry tests cover various approaches of empirical revision methods, the result sets are generated independent of its corresponding processes and tests of controls [1, 27].

While there are several procedures assessing risk and related activities, process mining can be used for generating an understanding of the company and its environment, including its internal controls. Process mining can realize potentials according to the company's risk assessment, for example, checking control activities according to *segregation of duties* or *information processing*. Furthermore, activities can be monitored by different views on the process. Here, for example, suspicious communication paths or critical organizational areas can be highlighted [6]. Combined with journal entry tests, we are of the opinion that considering both, financial accounting data and process data supports the auditor in (1) gaining an understanding of the company and its environment, (2) identifying risky business areas with regard to accounting-related internal controls, (3) supporting the assessment of material misstatements of financial statements considering the account posting behaviour, and (4) checking the company's internal controls of financial reporting.

#### 4. Event Log Extraction

To apply process mining using a SAP database, transforming operational data to an appropriate event log is a necessary preliminary step. In this section,

we describe how an event log can be extracted from a SAP database. This will be described by the example of the purchase-to-pay-process. First, we define the scope of the process and briefly describe each process activity. Followed by a description of its typical implementation in the SAP system, we provide an overview and description of the involved SAP tables and used attributes.

#### 4.1. Purchase-to-Pay-Process

The purchase-to-pay-process usually consists of the following six core process steps [2, 21]: (1) *requisition*, (2) *purchase order*, (3) *goods receipt*, (4) *invoice receipt*, (5) *accounting document* and (6) *payment*. In (1), suppliers are selected and purchase prices and conditions are negotiated. Needs are triggered based on an inventory control in the warehouse or individual needs. Afterwards, data is entered into the system. In (2), purchase orders are created, either as individual purchase order or as release order from a master agreement stored in the system. Tendering processes, where a selection of suppliers or the creation of new contracts are happen, can be also necessary in this phase. In (3), goods receipts are recorded with reference to the purchase order. Deviations to the quantity and quality of the goods delivered are checked and reported. In (4), invoice receipts are checked and entered into the system. In (5), an accounting document is generated. During invoice verification, a formal and substantive examination of the invoice items with the delivered goods takes place. Subsequently, the accounting documents can be used to identify deviations regarding price or quantity of the items shown. Finally, the payment will be released. In (6), the payment process is triggered which includes the generation of the payment proposal list and creation of a payment medium. After the approval of the payment, the data is transferred to the financial institution and the involved accounts are cleared.

#### 4.2. The Purchase-to-Pay-Process in SAP

In the SAP system, transaction data is processed in form of documents consisting of a header and segment [21]. The document header contains general information such as the type of document, date or number. The document segment contains information that belong to the document item. Typical information include posting key, amount and currency. Header and segment information are stored in different tables. For financial accounting documents, the tables *BKPF* and *BSEG* are used, for purchasing documents *EKKO* and *EKPO*, respectively. The different tables are linked with each

other using key relations like the *document number*. In order to extract an event log, we map the previously described process activities to the database tables. For the purchase-to-pay-process, we identified ten activities and eight different tables that are needed. A description of the used tables and its relation to each process activity is given in Table 1.

Table	Description	Activity
EBAN	Contains information of a purchase requisition, whether manual or automatically created.	requisition
EKKO	Stores the header information that are valid for the entire purchasing document. In addition to the vendor data, the table includes information regarding currency and terms of payment	request, contract, order
EKPO	Stores the data associated with a purchase order that refers to the individual order items. Together with EKKO, this table describes a purchase order.	request, contract, order
EKBE	Stores the goods receipts and invoice receipts for each order item, and thus gives details like whether goods receipt is done or invoice is done.	goods receipt, invoice receipt
BKPF	Contains the general ledger document headers. It includes information valid for the entire accounting document.	accounting document
BSEG	In contrast to the subledgers, all posting lines of an accounting document are stored in this table, regardless of the account type. Together with BKPF, this forms the general ledger. This table is used to get detailed information of each accounting transaction.	accounting document
BSIK	Accounting subledger of vendors that contains all open vendor items.	open payment.
BSAK	Accounting subledger of vendors that contains all cleared vendor items. While all information due to automatic payment run are stored in table REGUH, BSAK includes all cleared vendor items.	payment

**Table 1. SAP tables accessed during the purchase-to-pay-process (based on [21])**

#### 4.3. From SAP Tables to Event Log

To extract the event log, for each activity the corresponding tables and interdependencies have to be identified. Subsequently required data attributes for the extraction of the event log can be derived and extracted.

Purchase requisitions are stored in *EBAN*. The header information is stored in *EKKO*. Position data is stored in *EKPO*. The material receipt can be figured out via table *EKBE*. Here, we use *VGABE* as filter criterion to differentiate between *goods receipt*, *invoice receipt* and other documents like *subsequent debit/credit*. The tables *BKPF* and *BSEG* are used to reconstruct the accounting document activity. Information regarding clearing of an account is stored in *BSIK* and *BSAK*. *BSIK* tracks all items that have not been cleared so far, whereas *BSAK* contains all cleared items. Therefore, this tables can be used to check if there are non-cleared invoices or if any segregation of duties conflicts exist in

the payment program [21]. In order to create an event log, we extend the preprocessing steps defined in [10] as followed:

**(1) Document Selection:** During our case study, we will focus on typical process start documents like *requisition*, *request*, *contract* and *order*. A back-tracing from an end document like an invoice will not be considered during our case study. If a requisition is selected, all events related to the requisition will be comprised in the process instance. This may include a request, an order as well as several goods receipt entries and invoices. This procedure applies analogously to request, contract and order.

**(2) Granularity Level Selection:** Based on the process analysis objective, we distinguish between two granularity levels of a document. Activities on parent level usually refer to document header information and thus refer to a complete document (e.g. requisition). Positions of the requisitions will be aggregated and not considered separately. On the other hand, activities on child level refer to line items (position data) in the document segment. For example, *goods receipt* or *invoice receipt* refers to a single line item of a document like an order. This differentiation and decision making is required for preprocessing due the fact of the relational data structure provided by the SAP system. For example, booking an invoice is done on header level of granularity (*BKPF*). However, the link to an order is provided in the document segment (*BSEG*). Here we have to decide on which level the activities, e.g. item level or aggregated, should be used in our event log. In this case study, we select the document on parent level. This means that we only look at cases that belong to the previously mentioned start documents as a whole. The granularity level for the activities itself is set to line items (child level).

**(3) Caseid Determination:** Every process instance must have a unique *caseid* [15]. In our case, the *caseid* is created by combining the three-digit client number (MANDT) with a ten-digit document number and a five-digit item number. The selection of the document number and item number depends on the document category. The document categories correspond to the selected start documents. In the SAP system, every business transaction has its own document number. To ensure that the individual documents of each table belong to the same case, the documents must be reconciled with the documents from it's previous activities. If this is the case, the document number and item number of the previous activity is assigned. This step is executed recursively along the document chain, until no previous activity exists anymore. For example, we have to check recursively for

each request, contract or order if they refer to an existing requisition. If this is the case, we will create the *caseid* using the attributes *MANDT-BANFN-BNFP*, otherwise *MANDT-EBELN-EBELP*. The *BANFN* and *BANFP* attributes describe the purchase requisition number and position, respectively. The *EBELN* and *EBELP* attributes describe the purchase document number and item, respectively. This has to be done in order to assign logically related transaction postings to the same case and was performed accordingly for all specified start documents. Afterwards, we extracted sub-event-logs for each process activity and merged them. A description of the extracted event log is provided in Table 2.

Description of the Event Log: Purchase-to-Pay		
#Events		309.357
#Trace		38.491
Activity		10
#Resource		296
#Variant		1.309
Mean Process Duration		37 days
Start	2016-11-02 00:00:00	
End	2019-02-06 16:59:32	

Activity	Frequency	Percentage
Requisition	37.147	12.00
Request	62	0.02
Contract	232	0.08
Order	38.707	12.51
Goods receipt	63.547	20.54
Invoice receipt	67.794	21.91
Invoice surcharge	980	0.32
Accounting document	68.513	22.15
Open payment	85	0.03
Payment	32.290	10.44
<b>Sum</b>	<b>309.357</b>	<b>100.00</b>

Table 2. Event Log Description

## 5. Process Mining and Journal Entry Tests

In the following, we will combine process mining and journal entry tests. The procedure is two-fold. First, we conduct a descriptive analysis and mine a process model. Afterwards, journal entry tests were applied and combined with process mining in order to evaluate the potential of combining both approaches.

### 5.1. Application of Process Discovery

First, we explore the most frequent process variants. As shown in Figure 1, the most frequent process variant is (1). This process appears 19.563 times in the event log. As we can see, each activity is performed once, except *accounting document* which is performed multiple times. Variant (2) does not contain any *accounting document* or *payment* activities. Due to its high frequency of 5.622 cases, the underlying order documents should be inspected more in detail. In this

case, it might be an indication of outgoing payments that are not recorded by the accounting department. Trace (3) can be observed 2.330 times. This may indicate poor data quality or inconvenient data delimitation. Looking at trace (4), first it seems to be normal, but in 1.161 cases the activity *goods receipt* is missing. This can indicate violations of internal guidelines and controls such as n-way-matching. Trace (5) with a frequency of 988 shows that there are purchase orders consisting of multiple *invoices receipt* activities.

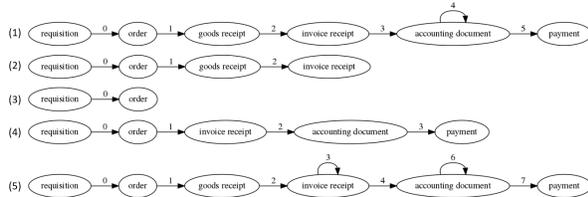


Figure 1. The five most frequent process variants

Afterwards, we mined various process models using the *Alpha Miner*, *Inductive Miner* and *Heuristics Miner*. We measured and compared the *fitness*, *precision* and *simplicity* of each process model in order to get a good visual representation of the purchase-to-pay-process. Compared to the *Alpha Miner*<sup>1</sup> and *Inductive Miner*<sup>2</sup>, the *Heuristic Miner* provides the best results according to *fitness* (91,43%), *precision* (81,18%) and *simplicity* (68,83%). The dependency threshold was set to 70% to keep the complexity of the mined model not that high. As representation, we used a heuristic net as shown in Figure 2. For process discovery, we used the PM4Py-library (Version: 1.2.11) [28].

In the following, dedicated journal entry tests based on [20, 21] are presented. Afterwards, these will be combined with process discovery techniques.

## 5.2. Timely Recording of Documents

With regard to Generally Accepted Accounting Principles (GAAP) every company is required to provide valuation information of its financial statements [29]. For this reason, an analysis of timely recording of documents is crucial, such as the recording of *goods receipts* or the *posting of invoices*. These activities are decisive for the company in order to comply with GAAP requirements. According to ISA 240 A.43, accounts that "[...] have not been reconciled on a timely basis or contain unreconciled differences [...]" [30, p. 180] are of interest for journal entry tests. In order to analyze timely recording of documents, tests are carried out with regard to the promptness as well as the period accuracy.

<sup>1</sup> fitness: 6,54%, precision: 34,43% & simplicity: 66,67%

<sup>2</sup> fitness: 100%, precision: 41,48% & simplicity: 72,66%

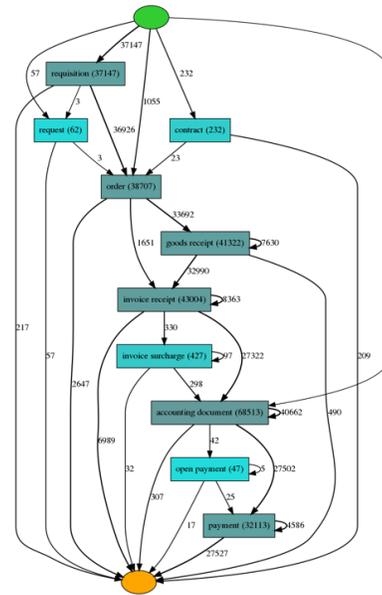


Figure 2. Heuristic net of the purchase-to-pay-process

The SAP system differentiate between three types of timestamps [20]. The *entry date* is automatically set by the system when the user enters a document. The *document date* is set by the user and describes the date of the document received. The *booking date* is used to assign the document to the accounting period concerned. Whereas the *document date* is compared to the *entry date* for promptness, the *booking date* is compared with the *document date* for period accuracy. To analyze timely recording of booking transactions, the *booking date* is compared to the *entry date*. In the following, we will focus on the timely difference between the document date and entry date which should correspond only to a few working days.

Table 3 gives an overview of the difference in days between entry dates and document dates. Here, we filtered by the data attribute *BLART*="WE". In sum, there are 210.865 goods receipts. 19 goods receipts are entered before its document date. These receipts can be classified as suspicious because usually goods receipt should be at least entered on the same date as the document date. This is true for nearly all goods receipts (96,4%). However, more than 1.500 goods receipt postings take longer than *5days*. Whether these time delays are process-related or related to delivery difficulties of the supplier cannot be conclusively assessed on the basis of these test results. In order to identify possible process-related anomalies, a process-oriented visualization of the provided results can lead to further interesting insights. For example,

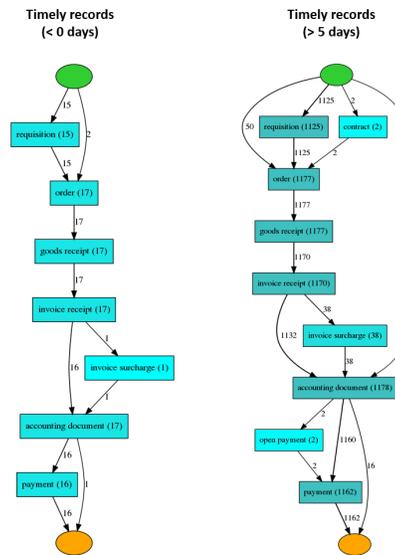
Difference in days	Number of goods receipts	Total value
< 0	19	804.663,28
0	203294	723.663.067,46
1-5	5869	180.527.950,58
6-10	948	73.968.242,34
> 10	735	21.205.912,66
<b>Total</b>	<b>210.865</b>	<b>1.000.169.836,31</b>

**Table 3. Time difference in days between document dates and entry dates**

is there a link between the process and the set of rules specified? Based on the identified conspicuous goods receipts, we mined two process models for goods receipts having a time delta of < 0 days and > 5 days respectively. As shown in Figure 3, a delta analysis of both mined process models revealed deviations with regard to the process behaviour. In contrast to the process model with timely records > 5 days, it looks like that only transactions relating to a purchase requisition cause such suspicious behaviour. On the other side, the process model with timely records > 5 days shows that there are processes having no requisition. Especially with regard to the orders and the contracts, a more detailed test of them may be relevant. This can result, for example, from a deficiency of internal controls. For example, documents that are not relate to a purchase requisition are not centrally parked in financial accounting department, but are received directly by the purchasing department and are therefore not forwarded promptly. Furthermore, a process trace that leads directly into the event *accounting document* can be identified. This may be due to the temporal delimitation of the data set, so that the underlying order is from 2017 and no anomaly can be assumed. However, it could also be a posting that is not based on a purchase order entered in the system. Such behavior would be a flaw in the internal control system. With regard to the timely entry of a business transaction, a test of detail of the underlying controls could be triggered by the auditor.

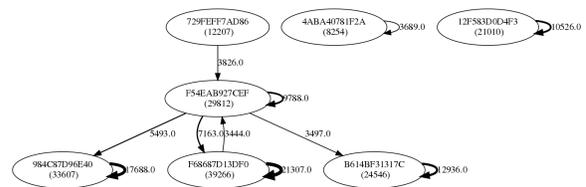
### 5.3. Manual Payments

In the SAP system, one can distinguish between two types of outgoing payments [21]. Automated payment runs usually take place at specific times. The due date of a payable is checked automatically, just as the payment is triggered. There is no manual intervention along the entire payment procedure. The system ensures that posted invoices are assigned to a vendor. Invoice verification is carried out automatically and there are no payment blocks. On the other side, all payments that do not take place within these automated payment runs are referred to as manual payments. For manual payments, the above mentioned conditions do not have to be met. In addition, manual payments do not appear



**Figure 3. Comparison of two control-flows of timely recorded documents**

as payments in the SAP system because they are usually executed by a third-party system like a the banking software. From a data-oriented view, no payment data is stored in the SAP system. This means that if the data from the third-party software will not be made available for analysis, the data has to be integrated or even compared manually. Due to this, they are considered to be particularly risky and should be checked since only their accounting effects are recorded in SAP system. Since the verification of the payment is the last step in the purchase-to-pay-process, transparency along the entire business process is of great importance due to the company's liquidity. Questions like *are there differences between manual payed invoices and invoices automatically payed?* may lead to interesting insights that trigger further tests.



**Figure 4. Social network of manual payments**

Because it is not limited to the control-flow perspective, process mining can be used to analyze the organizational structure as well. Here, a network analysis analyzing the handover of work was implemented. With regard to manual payments it may be interesting to identify those users who interact quite often with each other. Therefore, we

filtered on users who appear more than  $> 5.000$  times in the event log. As depicted in Figure 4, user *F54EAB927CEF* is strongly involved by users *984C87D96E40*, *F68687D13DF0* and *B614BF31317C*. Having a look at the control-flow of the activities related to the users, we see that the four users are involved in around 55% of all payments of the mined purchase-to-pay-process (32.290 times). While there are 17.976 manual payments, one open payment as well as 4.518 cases could be identified where activity *accounting document* does not relate to *payment*. Even if the latter may be caused by the data delimitation, the open payment might be interesting for the auditor. While such individual cases are difficult to detect by random sampling, journal entry tests combined with process mining can be seen as a versatile analytic tool to analyze large amounts of data from different perspectives. Figure 5 shows the corresponding control-flow.

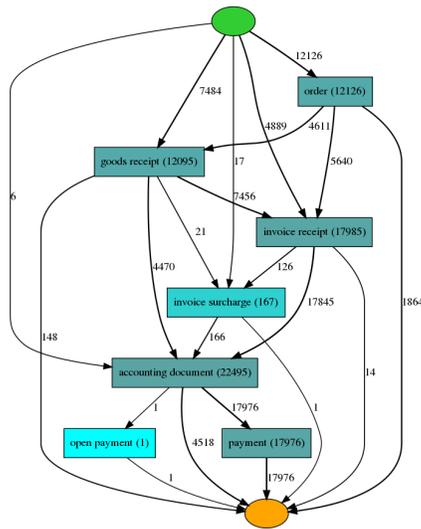


Figure 5. Heuristic net of manual payments

#### 5.4. Counter Account Analysis

The aim of the counter account analysis is to assign the counter accounts for a particular account and to identify conspicuous posting structures and variants [20]. Figure 6 shows an excerpt of the posting structure. While the journal entry test leads to interesting insights which account structures are booked very often, a process visualization of the underlying booking transaction can be used to make suspicious control-flow aspects or organizational structures visible. Thus, compared to the conventional opposite account analysis, not only quantitative aspects are taken into consideration. For example it can be checked if internal

controls are systematically circumvented when posting a transaction to a specific set of accounts. However, a description of the general ledger accounts using table *SKAT* in which the account descriptions are stored, was not possible because they were anonymized.

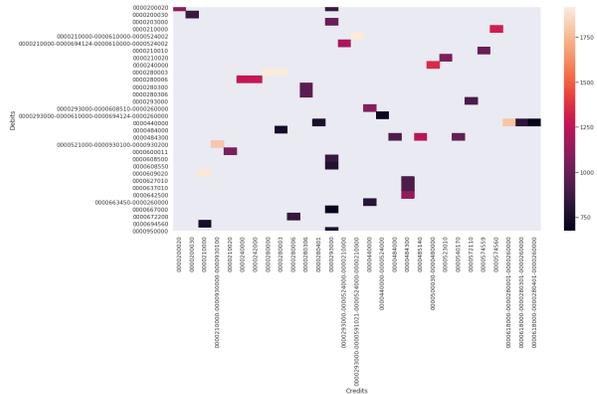


Figure 6. Results of the counter account analysis

#### 5.5. Payments without Performance Record

Accounts payable deals with entering invoices and posting of payables to a particular vendor in the system. As long as these payables have not been paid, they are treated as open items. When they are get paid, the invoice will be cleared. In the context of the purchase-to-pay-process, there are system-documented liabilities in the form of invoice documents that precede the payment [21]. While payments with invoice reference can be considered less critical, invoices without a proof of performance have no assigned liabilities. In this context, such payments can be considered critical and relevant for the audit, since the transparency of the related business transaction is lost. Thus, there is no reference to an invoice or purchase order. Typical risks are insufficiently specified guidelines or poorly implemented controls [21]. Here, journal entry tests are helpful in order to analyze the outgoing payments. In the following, we conduct a test to identify open outgoing payments without a performance record, filtered by suppliers with the highest amount. Therefore, we filtered on the key *BSCHL*= "25" in *BSIK* to get outgoing payments. The results are displayed in Table 4.

LIFNR	BSCHL	NAME1	XCPDK	VBUND	DMBTR
0000140040	25	DFECCC485E6DC86BA54738D543BAEBEFDB7		399141	283492.00
0000140070	25	0C3C68661EACA0D17B05C5A794D3BB4C90E		399141	158069.72
0000140065	25	64240D2C9ECC967B2E2E31C4E3D842FD2E		399141	110134.16
0000140205	25	0657AEGBCA9D9EE6057114D320EF609A4D7		399141	55406.28
0000140130	25	084BD792C17758C31E58FD8D4EBA52F783		399141	30362.00

Table 4. Overview of top five suppliers with the highest amount of open outgoing payments

Within the analysis, we were unable to create an event log and thus to apply process mining. As a limitation of our proposed approach, we can see that depending on the type of journal entry test, a combination with process mining is not always possible. This is due the fact that no liabilities exit to the payments and thus necessary information is not provided in the SAP system. While journal entry tests can be used to identify and test risky areas like open liabilities, it can trigger further substantive tests and test of details of the corresponding payments. Furthermore, it is obvious that these payments are made manually. Thus, they can be treated similar like manual payments where only their accounting effects are subsequently mapped in the SAP system. However, taking into account the transaction data of third-party systems used (e.g. banking software), journal entry tests combined with process mining can be considered as a helpful tool to identify risky payments, to show the underlying processes in the third-party system and to check them along internal controls.

## 6. Discussion

In this paper, we successfully applied process mining and journal entry tests to a SAP dataset. Particularly, we implemented an approach to extract an event log to mine comprehensible process models. By narrowing down the data through journal entry tests, first, we were able to derive more detailed and specific process models according to the filtered booking transactions. This helps to generate a more comprehensive process understanding as well as to identify critical areas in the process-flow. Second, the developed approach can be used as a tool in the risk-oriented audit that can lead to efficiency gains and risk mitigation. By revealing conspicuous transactions through analyzing uncommon process behaviour, it can help the auditor to get a more comprehensive view on financial statements under consideration of the client's business, processes and internal controls. Of course, the proposed approach does not come without limitations. Currently, the event log extraction is centered around the idea of an already well known business process in which process activities can be directly mapped to specific SAP database tables. If the process of interest or SAP tables are not known in advance, the proposed event log extraction is not realisable. Another limitation is the statically defined granularity level. As described in [10], changing the granularity level of the considered purchase documents and underlying activities might be beneficial according to the audit objectives. In this study, the journal entries are used as filters before process mining methods were applied. However, other ways to integrate process

mining and journal entry tests need to be further elaborated. For example, evaluation process mining along different perspectives of an audit framework may be interesting. This includes, for example, the evaluation along the internal controls of the client which could not be considered during this case study. Furthermore, technical issues dealing with *convergence* and *divergence* to extract appropriate event logs have to be addressed in more detail. Here, approaches that focus on the object's life cycle like a *purchase order* or *invoice document* might be of special interest [31, 32].

## 7. Conclusion

The prototypical implementation addresses two issues for the application of process mining and journal entry tests. First, an approach on how relational data structures can be transformed to an event log was presented. Second, the audit of business processes in combination with journal entry tests were conducted. We showed how a process-oriented visualization supports the auditor during identification and assessment of a critical business process as well as conspicuous transactions, and thus pointed out the practical applicability within a risk-oriented audit. As future work, we plan to work on a dynamic event log extraction algorithm, which can be adapted to the specific requirements of the client at run-time and allows to evaluate preemptively unknown business processes. This goes hand-in-hand with the integration of the proposed approach in an audit framework and evaluating it along different journal entry test perspectives like *what, how, when, who* etc. [20]. Another promising topic is the further integration of journal entry tests and process mining. On the one hand, there are journal entry tests that offer potential to be used in conjunction with existing process discovery techniques. On the other hand, it might be advantageous to design novel process discovery and conformance checking methods that are linked with internal controls and analyze the process from different compliance perspectives. Additionally, the integration of data from third-party systems is subject of further research.

## References

- [1] H. Madani, "The role of internal auditors in ERP-based organizations," *Journal of Accounting & Organizational Change*, vol. 5, pp. 514–526, 2009.
- [2] G. Brösel, C. Freichel, M. Toll, and R. Buchner, *Wirtschaftliches Prüfungswesen*. München: Vahlen, 3 ed., 2015.
- [3] R. Hayes, P. Wallage, and H. Gortemaker, *Principles of Auditing: An Introduction to International Standards on Auditing*. Harlow: Pearson Education Limited, 3 ed.,

- 2014.
- [4] International Standard on Auditing (ISA) 330, "The Auditor's Responses to Assessed Risks." <https://www.ifac.org/system/files/downloads/a019-2010-iaasb-handbook-isa-330.pdf>, 2020. Retrieved: 18.11.2020.
  - [5] M. Werner, N. Gehrke, and M. Nüttgens, "Business Process Mining and Reconstruction for Financial Audits," in *Proceedings of the 45th Hawaii International Conference on System Sciences*, pp. 5350–5359, 2012.
  - [6] W. van der Aalst, K. van Hee, J. van der Werf, and M. Verdonk, "Auditing 2.0: Using Process Mining to Support Tomorrow's Auditor," *Computer*, vol. 43, no. 3, pp. 90–93, 2010.
  - [7] International Standard on Auditing (ISA) 315, "Identifying and Assessing the Risks of Material Misstatement through Understanding the Entity and Its Environment." <https://www.ifac.org/system/files/downloads/a017-2010-iaasb-handbook-isa-315.pdf>, 2020. Retrieved: 18.11.2020.
  - [8] J. E. Ingvaldsen and J. A. Gulla, "Preprocessing Support for Large Scale Process Mining of SAP Transactions," in *Business Process Management Workshops* (A. ter Hofstede, B. Benatallah, and H.-Y. Paik, eds.), (Berlin, Heidelberg), pp. 30–41, Springer, 2008.
  - [9] M. Krebs, F. Stadler, and J. Anke, "Vorbereitung von SAP Event Logs für Process Mining mit ProM. Preparation of SAP Logfiles for Process Mining Using ProM," *HMD Praxis der Wirtschaftsinformatik*, vol. 55, no. 1, pp. 104–119, 2018.
  - [10] M. Jans and P. Soffer, "From Relational Database to Event Log: Decisions with Quality Impact," in *Business Process Management Workshops* (E. Teniente and M. Weidlich, eds.), (Cham), pp. 588–599, Springer International Publishing, 2017.
  - [11] N. Gehrke and N. Mueller-Wickop, "Basic Principles of Financial Process Mining A Journey through Financial Data in Accounting Information Systems," in *AMCIS 2010 Proceedings*, vol. 289, 2010.
  - [12] N. Mueller-Wickop, M. Schultz, N. Gehrke, and M. Nüttgens, "Towards Automated Financial Process Auditing: Aggregation and Visualization of Process Models," in *Proceedings of the 4th International Workshop on Enterprise Modelling and Information Systems Architectures* (M. Nüttgens, O. Thomas, and B. Weber, eds.), (Bonn), pp. 121–134, Gesellschaft für Informatik e.V., 01 2011.
  - [13] M. Werner, M. Schultz, N. Müller-Wickop, N. Gehrke, and M. Nüttgens, "Tackling Complexity: Process Reconstruction and Graph Transformation for Financial Audits," in *Proceedings of the 33rd International Conference on Information Systems* (F. G. Joey, ed.), (Atlanta, GA), 2012.
  - [14] M. Werner and N. Gehrke, "Multilevel Process Mining for Financial Audits," *IEEE Transactions on Services Computing*, vol. 8, no. 6, pp. 820–832, 2015.
  - [15] M. Werner, "Financial process mining - Accounting data structure dependent control flow inference," *International Journal of Accounting Information Systems*, vol. 25, no. February, pp. 57–80, 2017.
  - [16] M. J. Jans, "Process Mining in Auditing: From Current Limitations to Future Challenges," in *Business Process Management Workshops* (F. Daniel, K. Barkaoui, and S. Dustdar, eds.), (Berlin, Heidelberg), pp. 394–397, Springer, 2012.
  - [17] M. Jans, M. Alles, and M. Vasarhelyi, "A Field Study on the Use of Process Mining of Event Logs as an Analytical Procedure in Auditing," *Accounting Review*, vol. 89, no. 5, pp. 1751–1773, 2014.
  - [18] T. Chiu and M. Jans, "Process Mining of Event Logs: A Case Study Evaluating Internal Control Effectiveness," *Accounting Horizons*, vol. 33, no. 3, pp. 141–156, 2019.
  - [19] M. Jans and M. Hosseinpour, "How active learning and process mining can act as Continuous Auditing catalyst," *International Journal of Accounting Information Systems*, vol. 32, no. November 2018, pp. 44–58, 2019.
  - [20] K. Droste and J. Tritschler, *Journal Entry Testing*. Düsseldorf: IDW Verlag, 1 ed., 2018.
  - [21] A. Bönner, M. Riedl, and S. Wenig, *Digitale SAP® - Massendatenanalyse: Risiken erkennen - Prozesse optimieren*. Berlin: Erich Schmidt Verlag, 2011.
  - [22] C. Durtschi, W. Hillison, and C. Pacini, "The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data," *Journal of Forensic Accounting*, pp. 17–33, 2004.
  - [23] A. Argyrou, "Auditing Journal Entries Using Extreme Value Theory," in *Proceedings of the 21st European Conference on Information Systems*, vol. 169, 2013.
  - [24] M. Schreyer, T. Sattarov, D. Borth, A. Dengel, and B. Reimer, "Detection of Anomalies in Large Scale Accounting Data using Deep Autoencoder Networks," 2018.
  - [25] M. Schultz and M. Tropmann-Frick, "Autoencoder Neural Networks versus External Auditors: Detecting Unusual Journal Entries in Financial Statement Audits," in *Proceedings of the 53th Hawaii International Conference on System Sciences*, 2020.
  - [26] S. F. Wamba, A. Gunasekaran, S. Akter, S. J.-f. Ren, R. Dubey, and S. J. Childe, "Big data analytics and firm performance: Effects of dynamic capabilities," *Journal of Business Research*, vol. 70, pp. 356–365, 2017.
  - [27] R. Brody and G. Kearns, "IT Audit Approaches for Enterprise Resource Planning Systems," *ICFAI Journal of Audit Practice*, vol. 6, no. 2, 2009.
  - [28] W. van der Aalst, A. Berti, H. Beyel, C.-Y. Li, R. Schimassek, D. Schuster, J. Xu, and S. J. van Zelst, "State-of-the-art-process mining in Python." <https://pm4py.fit.fraunhofer.de/>, 2020. 18.11.2020.
  - [29] S. Kothari, K. Ramanna, and D. J. Skinner, "Implications for GAAP from an analysis of positive research in accounting," *Journal of Accounting and Economics*, vol. 50, no. 2, pp. 246–286, 2010.
  - [30] International Standard on Auditing (ISA) 240, "The Auditor's Responsibilities Relating to Fraud in an Audit of Financial Statements." <https://www.ifac.org/system/files/downloads/a012-2010-iaasb-handbook-isa-240.pdf>, 2020. Retrieved: 18.11.2020.
  - [31] G. Li, E. de Murillas, R. de Carvalho, and W. van der Aalst, "Extracting Object-Centric Event Logs to Support Process Mining on Databases," in *Information Systems in the Big Data Era* (J. Mendling and H. Mouratidis, eds.), (Cham), pp. 182–199, Springer International Publishing, 2018.
  - [32] W. van der Aalst, "Object-Centric Process Mining: Dealing with Divergence and Convergence in Event Data," in *Software Engineering and Formal Methods* (P. C. Ölveczky and G. Salaün, eds.), (Cham), pp. 3–25, Springer International Publishing, 2019.