

Prüfungen mit ProcessMining

Zusammenfassung

Rechnungsprüfer:innen orientieren sich an einem klar definierten Prozess, mit dem sie eine Prüfung durchführen den sogenannten "Risikoorientierten Prüfungsprozess". Process Mining ersetzt diesen traditionellen Prüfungsansatz nicht. Es erfordert jedoch einige Anpassungen und bewusste Bemühungen, Process Mining in die bestehende Arbeitsweise zu integrieren. An manchen Stellen ist mehr Aufwand erforderlich. An anderen Stellen wird es einfacher. In diesem Artikel beschreiben wir detailliert, wie sich Process Mining anhand eines konkreten Projekts in die verschiedenen Phasen des Prüfungszyklus einfügt. Wir beschreiben die notwendigen Änderungen am Prüfungsprozess sowie die Vorteile und Herausforderungen.

Landesrechnungshof der Stadt Wien

Der Landesrechnungshof der Stadt Wien ist eine autonome und unabhängige öffentliche Finanzkontrollbehörde. Er prüft die Institutionen und Einrichtungen der Stadt Wien hinsichtlich ihrer Finanzgebarung und -sicherheit. Er unterstützt die Verantwortlichen in Politik und Verwaltung mit Prüfungsberichten und Empfehlungen. Er überprüft im Rahmen seiner Prüfungstätigkeit die Verwendung öffentlicher Mittel der Stadt Wien. Er überwacht auch die Einhaltung der Sicherheitsvorschriften zum Schutz der Wiener Bürgerinnen und Bürger sowie der Besucherinnen und Besucher.

Die Ressourcen für diese Aufgaben sind jedoch begrenzt. Um die Prüfungsschwerpunkte zu setzen, führt der Landesrechnungshof der Stadt Wien ein Prüfungsauswahlverfahren in Form einer Risikoanalyse durch. Für die ausgewählten Prüfungen wendet das Team des Wiener Landesrechnungshofs je nach Prüfung unterschiedliche Prüfverfahren an. Entscheidender Erfolgsfaktor war die Zeit, die Daten wirklich zu verstehen.

Seit 2016 wurde Process Mining als Prüfmethode in mehreren Prüfungen eingesetzt. 2020 wurde der Purchase-to-Pay-Prozess der Wiener Stadtwerke, einem der größten Infrastrukturkonzerne Österreichs, analysiert. Am Beispiel der Wiener Stadtwerke zeigt dieser Artikel detailliert, wie Process Mining zur Durchführung einer datengetriebenen Prüfung eingesetzt wurde. Die Ergebnisse dieser Prüfung sind im Prüfungsbericht [1] hier [2] öffentlich zugänglich. Der folgende Artikel konzentriert sich auf die Methode "Process Mining" im Rahmen einer Prüfung. Er beschreibt detailliert, wie Process Mining in den verschiedenen Prüfungsphasen eingesetzt wurde und welche Herausforderungen und Vorteile sich dabei ergaben.

Datengetriebener Prüfungsansatz mit Process Mining

Der Landesrechnungshof Wien orientiert sich bei seiner Prüfungstätigkeit an international anerkannten Prüfungsstandards. Ziel ist es, die bestehenden Standards in Zusammenarbeit mit nationalen und internationalen Prüfungsinstitutionen weiter zu verbessern und den Erfahrungsaustausch zu pflegen. Die Prüfungen erfolgen nach dem standardisierten Prüfungsprozess (siehe Abbildung 1).



Figure 1: Overall audit process

Jeder Schritt des oben dargestellten Prozesses umfasst mehrere Aufgaben. Abbildung 2 zeigt die Aufgaben im Zusammenhang mit dem Schritt „Durchführung der Prüfung“:

- Zunächst wird ein Prüfkonzept erstellt
- Anschließend werden die Daten erhoben.
- Auf Basis dieser Daten wird eine Situations- und Abweichungsanalyse durchgeführt,
- aus der die Auditergebnisse und Empfehlungen abgeleitet werden.
- Darüber hinaus werden der Auditpfad und die Nachweise dokumentiert und die Auditakte erstellt.

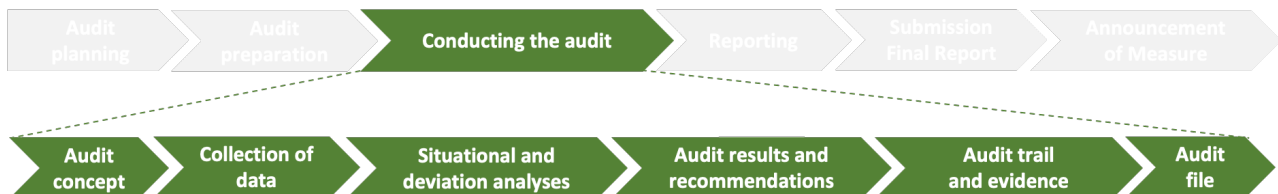


Figure 2: Conducting the audit

Wenn wir Process Mining in unserem allgemeinen Prüfungsansatz einsetzen, müssen wir unsere Arbeitsweise anpassen. Insbesondere die Art und Weise der Datenerhebung und -analyse ändert sich innerhalb des Prüfungsprozesses im Vergleich zu anderen Prüfungsmethoden.

Um Process Mining in die Phasen „Datenerfassung“ und „Situations- und Abweichungsanalyse“ einzubinden, haben wir die in Abbildung 3 dargestellten neun Schritte befolgt.

Dieses Modell hat uns sehr geholfen, unseren Ansatz beim Einsatz von Process Mining in Prüfungen zu standardisieren. Es fasst die wesentlichen Ergebnisse der Datenerhebung und der Durchführung der Situations- und Abweichungsanalyse zusammen. In den folgenden Abschnitten erläutern wir jeden Schritt und jedes Ergebnis des in Abbildung 3 dargestellten Modells detaillierter.

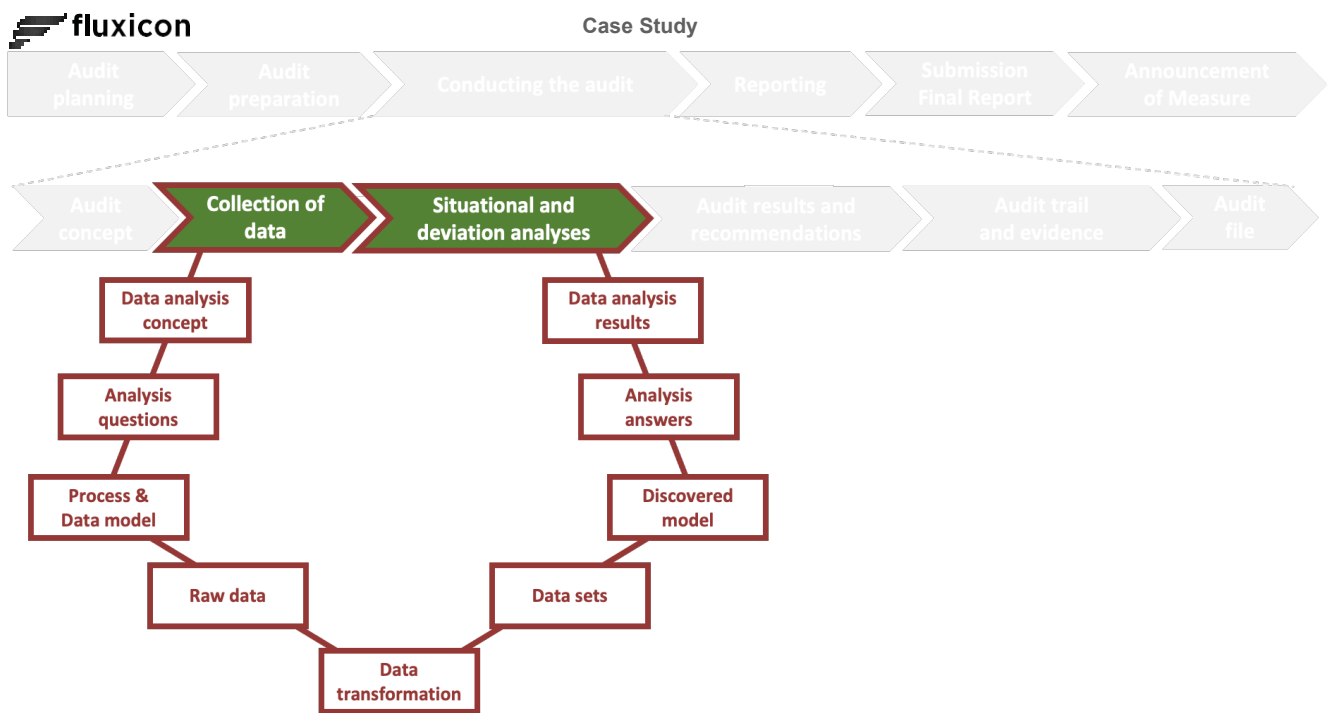
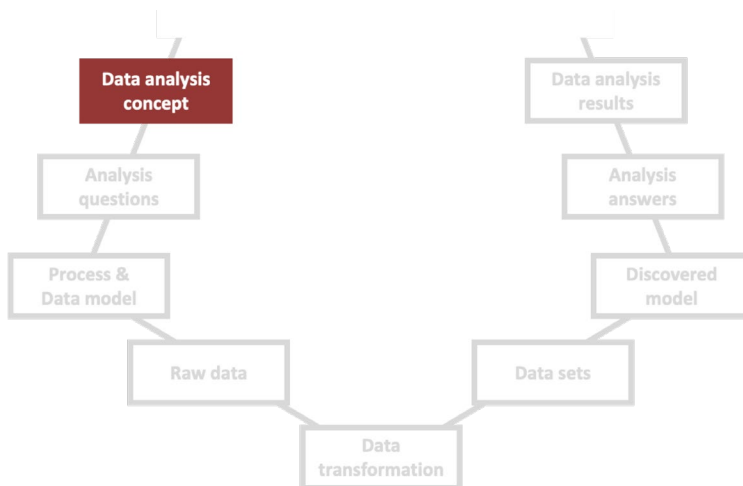


Figure 3: Data-driven audit approach with process mining

Step 1: Konzept der Datenanalyse



Der Landesrechnungshof Wien verfolgt ein risikobasiertes Prüfungsauswahlverfahren. Daher wurde die Prüfung bereits im Rahmen der jährlichen Prüfungsplanung definiert. Das „Datenanalysekonzept“ legt den Prüfungsumfang detailliert fest. Es gibt einen Überblick über die geprüfte Stelle, den relevanten Prozess, die IT-Infrastruktur und das Hauptprüfungsziel.

Die geprüfte Einrichtung – die Wiener Stadtwerke – ist einer der bedeutendsten Infrastrukturkonzerne Österreichs mit rund 15.000 Mitarbeiter:innen. Ihre Geschäftstätigkeiten lassen sich wie folgt kategorisieren:

Energie (Strom, Gas, Wärme, Kälte)

- Erzeugung
- Verteilung
- Netzbetrieb

Öffentlicher Personennahverkehr (U-Bahn, Straßenbahn, Bus).

- Verkehrsmanagement und -planung

- Betrieb
- Marketing
- Vertrieb
- Bestattung
- Friedhöfe
- Friedhofsgärtnerei
- Steinmetzarbeiten
- Parkplätze

Die Bilanzsumme der Wiener Stadtwerke belief sich zum 31. Dezember 2020 auf rund 13,9 Milliarden Euro. Der Wiener Stadtwerke-Konzern befand sich zu 100 Prozent im Eigentum der Stadt Wien.

Wir haben den relevanten Prozess detaillierter definiert, indem wir den allgemeinen Prozessumfang beschrieben haben. Da wir den Purchase-to-Pay-Prozess (=Beschaffung) prüfen wollten, haben wir diesen Prozess wie in Abbildung 4 dargestellt abgegrenzt. Der geprüfte Prozessumfang umfasste die Beschaffung und Rechnungsverarbeitung, beginnend mit der Bedarfsmeldung und endend mit der Bezahlung der entsprechenden Rechnung.

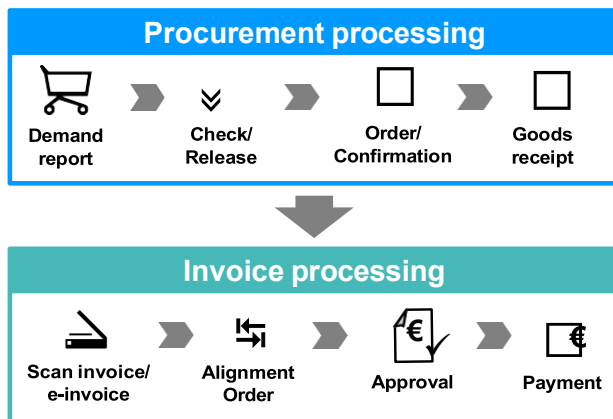


Figure 4: Purchase-to-pay process (Beschaffungswesen)

Dieser allgemeine Prozessumfang diente später als Referenzpunkt für weitere Untersuchungen und vermittelte einen ersten Eindruck von den Start- und Endpunkten des jeweiligen Prozesses.

Der Prüfungszeitraum wurde auf das Jahr 2019 festgelegt. Genauer gesagt berücksichtigten wir alle Bestellungen, die zwischen dem 1. Januar 2019 und dem 31. Dezember 2019 versandt wurden. Wir untersuchten die IT-Infrastruktur des Stadtrechnungshofs Wien und der geprüften Einheit, um uns ein Bild davon zu machen, welche Art von Daten und welche Tools für deren Verarbeitung und Analyse zur Verfügung stehen würden.

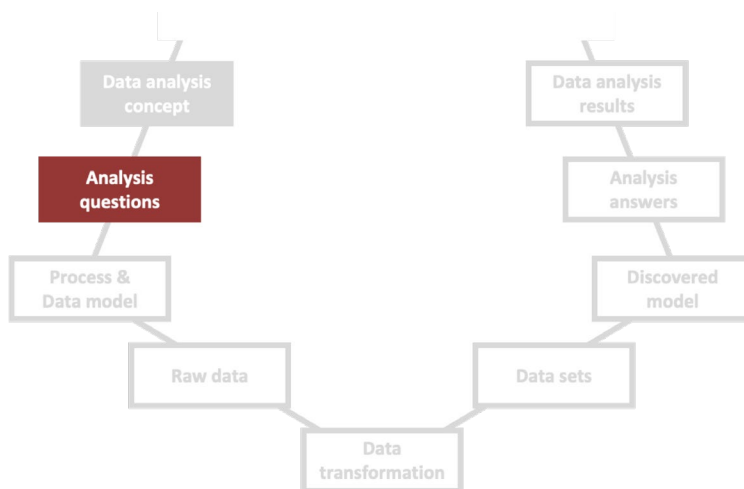
Aus unseren Voruntersuchungen wussten wir, dass die geprüfte Einheit SAP zur Abwicklung des Purchase-to-Pay-Prozesses nutzte. Wir gingen davon aus, dass wir die Daten nach dem Export aus SAP transformieren müssten. Als ETL-Tool für diese Transformationen wählten wir die KNIME Analytics Plattform [3], da wir diese Software bereits in früheren Process-Mining-Projekten für Datentransformationen eingesetzt und gute Ergebnisse erzielt hatten.

Das primäre Prüfungsziel bestand in der Durchführung eines Compliance-Audits. Ziel war es, den Purchase-to-Pay-Prozess der Wiener Stadtwerke hinsichtlich seiner Ordnungsmäßigkeit und der Einhaltung organisationsspezifischer Rahmenbedingungen zu analysieren.

Aufgrund dieses Hauptprüfungsziels planten wir, Aspekte wie die Vollständigkeit des Prozesses, die Aufgabentrennung, die Einhaltung des Vier-Augen-Prinzips und die Wirksamkeit des internen Kontrollsystems zu prüfen. Darüber hinaus wollten wir die Durchlaufzeit und das Auftreten von Engpässen aus Leistungssicht berücksichtigen. Fragen zur Benutzererfahrung lagen außerhalb des Prüfungsumfangs.

Nachdem wir den allgemeinen Rahmen und das primäre Prüfungsziel definiert hatten, ging es an die Detaillierung der Prüfungsschwerpunkte. Im nächsten Schritt identifizierten wir die konkreten Analysefragen, die wir im Rahmen unserer Process-Mining-Analyse beantworten wollten.

Step 2: Analysefragen im Einzelnen



Wir haben die folgenden zwölf Analysefragen definiert (siehe Tabelle 1). Wie aus Tabelle 1 hervorgeht, beziehen sich die meisten Fragen auf Compliance-Probleme (unser primäres Prüfungsziel), nur zwei Fragen beziehen sich auf die Leistung und keine auf die Benutzererfahrung.

Neben der Formulierung der allgemeinen Analysefragen haben wir versucht, diese möglichst präzise zu definieren und messbar zu machen. Daher haben wir für jede Analysefrage die Kennzahl, den Zielwert, den relevanten Prozessumfang und die Einflussfaktoren festgelegt. Tabelle 2 zeigt, wie wir Analysefrage Nr. 2 durch die Definition dieser Aspekte konkretisiert haben.

Die Frage „Werden alle Aufträge freigegeben?“ erscheint zunächst einfach, doch es ist sinnvoll, genauer darüber nachzudenken, wie wir die Antwort auf diese Frage messen können. Wir gingen davon aus, dass Auftragsfreigaben im Informationssystem erfasst werden. Daher wählten wir das Vorhandensein der Freigabeaktivität als Kennzahl. Darüber hinaus gingen wir davon aus, dass alle Aufträge ausnahmslos freigegeben werden müssen. Daher setzten wir den Zielwert auf 100 %, was bedeutet, dass für jeden Auftrag eine Freigabeaktivität dokumentiert werden muss.

Anschließend definierten wir den Prozessumfang, um zu zeigen, welcher Teil des Purchase-to-Pay-Prozesses für die Beantwortung der Analysefrage relevant ist. Bei Frage Nr. 2 umfasste der Prozessumfang alle Aktivitäten im Zusammenhang mit dem Freigabeschritt im Informationssystem. Schließlich haben wir die Einflussfaktoren erfasst, die wir bei der Datenanalyse und der Interpretation der Ergebnisse berücksichtigen mussten. Hinsichtlich der Auftragsfreigabe gingen wir davon aus, dass ein Vier-Augen-Prinzip sinnvoll sein könnte.

Analysefragen	Performance	Compliance
Entsprechen die tatsächlichen Abläufe den Soll-Vorgaben?		X
Existiert zu jeder Bestellung eine Rechnung?		X
Wurden alle Rechnungen vor der Bezahlung geprüft?		X
Wurden alle Rechnungen, die den Bestellwert überschritten, nach dem Vier-Augen-Prinzip freigegeben?		X
Wurden alle Bestellungen freigegeben?		X
Wurden alle Bestellungen bezahlt?		X
Wurden stornierte oder reklamierte Rechnungen bezahlt?		X
Wurden die Zahlungsziele beachtet?		X
Wurden alle Schulungen von der Personalabteilung genehmigt?		X
Wurden alle Rechnungen bezahlt?		X
Wurde das Vier-Augen-Prinzip eingehalten?		X
Gibt es Engpässe im Prozess?	X	

Table 1: Liste der Analysefragestellungen

Gegenstand der Prüfung	Performance		Compliance	X
Analysefrage	Wurden alle Bestellungen freigegeben?			
Messgröße/Kennzahl	Freigaben müssen dokumentiert werden			
Zielwert	100% (alle)			
Prozessumfang	PO - freigegebene Dokumente Ebenen 1 bis n, freigegebene Bestellpositionen			
Einflussfaktoren	Wertgrenzen (Vier-Augen-Prinzip)			

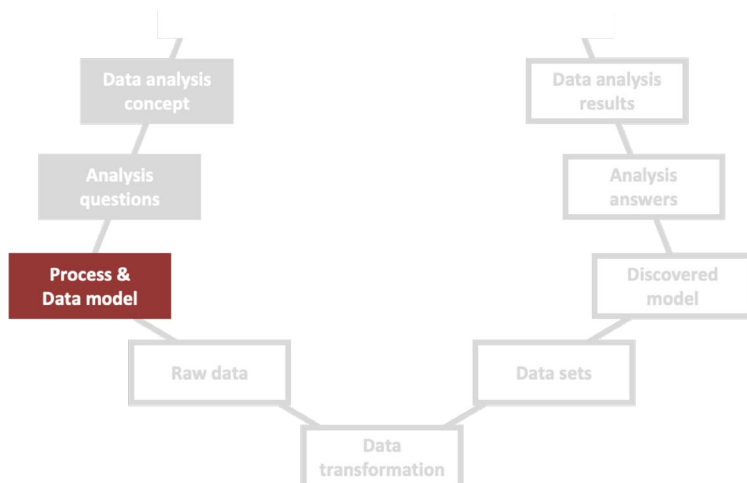
Table 2: Analysefrage No. 5 detailliert

Ab einer bestimmten Wertgrenze ist die Wertgrenze relevant. Daher haben wir die Wertgrenze als Einflussfaktor definiert, der bei späteren Analysen berücksichtigt werden muss.

Wir haben detaillierte Definitionen für alle Analysefragen festgelegt, ähnlich wie oben für Frage Nr. 5 beschrieben. Da wir die Analysefragen bereits in einer sehr frühen Phase des Process-Mining-Projekts definiert haben, mussten wir bestimmte Annahmen treffen, insbesondere hinsichtlich der Metrik und der Einflussfaktoren. In späteren Projektphasen haben wir diese Analysefragen mehrfach angepasst, da wir immer mehr Einblicke in den Prozess und die Daten gewonnen haben.

Dennoch war es sehr hilfreich, die Analysefragen bereits in dieser frühen Projektphase zu definieren. So erhielten wir einen guten Überblick über die benötigten Daten. Dadurch verringerten wir das Risiko, bei der Datenextraktion bestimmte Aspekte oder Datenfelder zu vergessen.

Step 3: Process and data model



Bei genauerer Betrachtung des Purchase-to-Pay-Prozesses wurde deutlich, dass dieser recht komplex war. Wir hatten eine detaillierte Prozessbeschreibung erhalten und beschlossen, den Prozess zu vereinfachen und ihn aus einer aggregierten Perspektive zu betrachten, um die Komplexität zu bewältigen. Der von uns definierte übergeordnete Referenzprozess umfasste nur die Schritte, die für die Beantwortung der im vorherigen Schritt beschriebenen Analysefragen wesentlich waren.

Wir gingen davon aus, dass für diesen Prozess erhebliche Datenmengen im Informationssystem generiert werden würden. Basierend auf dem übergeordneten Referenzprozess versuchten wir, die wesentlichen Datenfelder zu identifizieren, die während der Durchführung des Prozesses ausgefüllt wurden.

Der Purchase-to-Pay-Prozess wurde überwiegend in SAP abgewickelt. Für jeden Prozessschritt wurde die entsprechende Datenbanktabelle gesucht und das Prozessmodell mit diesen Informationen angereichert. Beispielsweise konnten die Daten für „Bestellanforderung erstellen“ in der EBAN-Tabelle in SAP gefunden werden (siehe Abbildung 5).

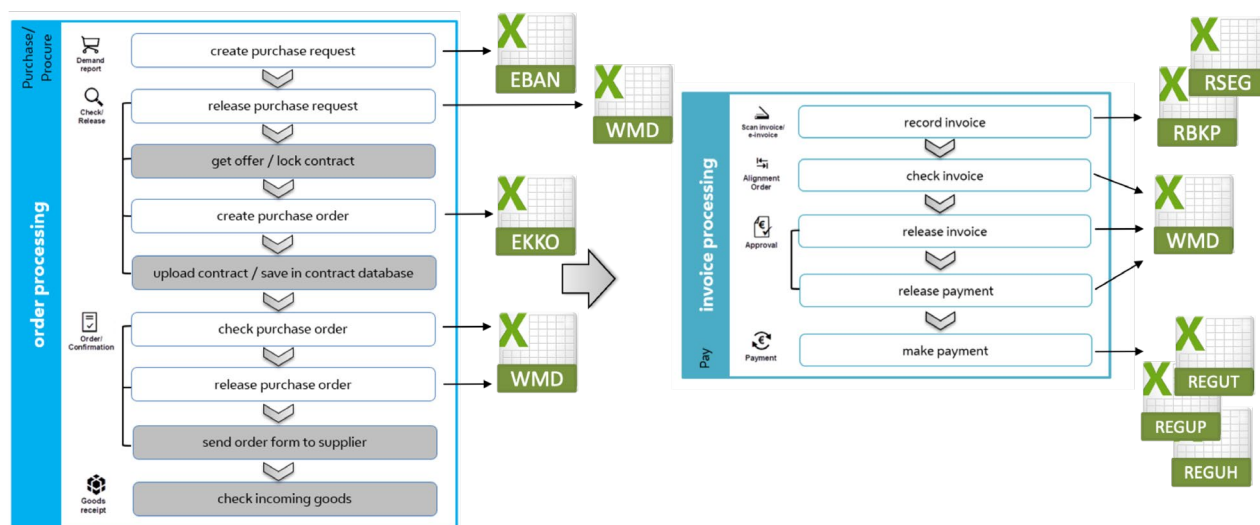


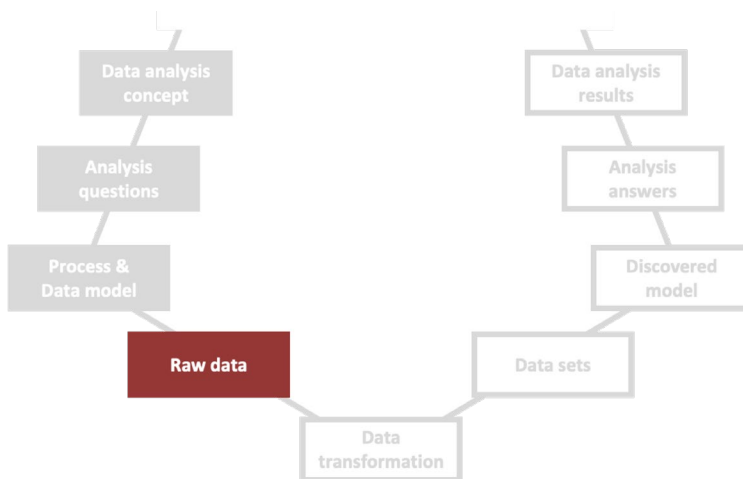
Figure 5: Process and data model

Bei der Erstellung des Prozess- und Datenmodells stellten wir fest, dass nicht alle Schritte des übergeordneten Referenzprozesses in SAP durchgeführt wurden. Beispielsweise stellten wir fest, dass die Genehmigungs- und Freigabe-Workflows mit einem SAP-Add-on namens WMD xSuite Feeder durchgeführt wurden. Da die Genehmigungs- und Freigabeschritte für das Compliance-Audit unerlässlich waren, haben wir diese Datentabellen in das Datenmodell und die spätere Datenextraktion integriert.

Weitere Schritte wie die Angebotseinholung, der Vertragsabschluss, das Versenden des Bestellformulars an den Lieferanten und die Wareneingangsprüfung wurden weder in SAP noch mit dem WMD xSuite Feeder durchgeführt. Diese Schritte (in Abbildung 5 grau dargestellt) wurden manuell oder per E-Mail durchgeführt. Aufgrund mangelnder Datenverfügbarkeit haben wir diese Schritte von unserer Process-Mining-Analyse ausgeschlossen.

Nachdem wir das Prozess- und Datenmodell definiert hatten, hatten wir einen guten Überblick über die verfügbaren Daten und wo wir diese finden konnten. Daher extrahierten wir im nächsten Schritt die Rohdaten für die weitere Verarbeitung.

Step 4: Raw data



Die Daten für unsere Process-Mining-Analyse waren in zwei verschiedenen Systemen gespeichert: SAP und dem WMD xSuite Feeder. Bei der Spezifikation des Datenmodells hatten wir die zu extrahierenden Datentabellen aus diesen Systemen identifiziert. Wir hatten keinen direkten Zugriff auf die Informationssysteme der Wiener Stadtwerke. Daher extrahierte die geprüfte Einheit die Datentabellen für uns und stellte uns die Rohdaten in CSV-Dateien zur Verfügung.

Im Datenmodell wussten wir bereits, in welchen Tabellen sich die Zeitstempel der einzelnen Aktivitäten befanden. Beispielsweise wussten wir, dass der Zeitstempel für „Kaufanfrage erstellen“ in der EBAN-Tabelle zu finden war. Der Zeitstempel für die Aktivität „Kaufanfrage freigeben“ befand sich in der WMD-Tabelle usw. Da die Rohdaten jedoch auf mehrere CSV-Dateien verteilt waren, mussten wir auch die Verbindungen zwischen den einzelnen Datentabellen finden, um die Dateien zu einer einzigen zusammenführen zu können (siehe Abbildung 6 auf der nächsten Seite für die Verbindungen zwischen den Tabellen)

Für jede Tabelle haben wir ermittelt, welche Informationen als Zeitstempel für eine Aktivität, Ressourcen und andere aktivitäts- oder fallbezogene Attribute verwendet werden könnten.

Basierend auf der Kenntnis der relevanten Datenfelder für die Aktivitätszeitstempel, Attribute und Ressourcen und mit diesem Verständnis der Zusammenhänge zwischen den Rohdatentabellen hatten wir nun die Grundlage für den Aufbau unseres Ereignisprotokolls [=Event-Log].

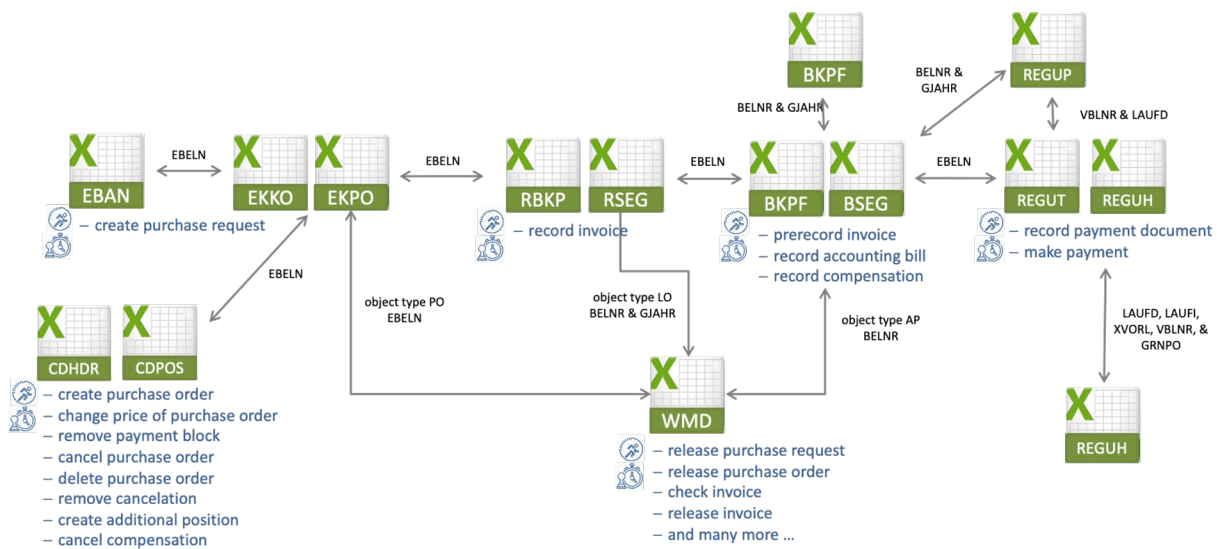
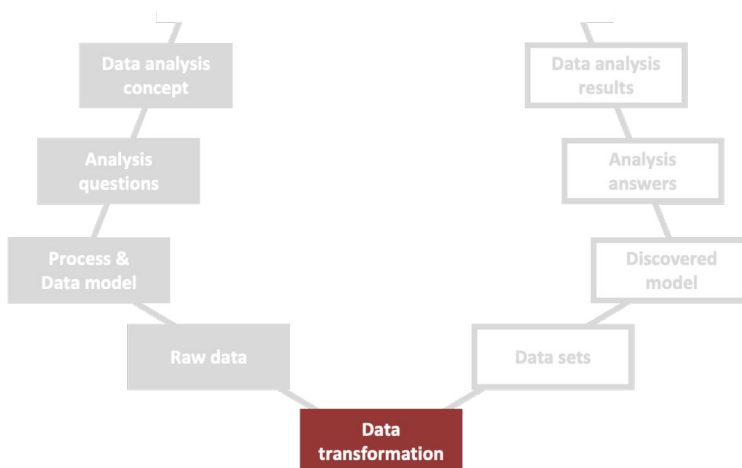


Figure 6: Datentabellen aus SAP und deren Verbindungen

Step 5: Data transformation



Ziel des nächsten Schritts war es, die Rohdaten in ein Format zu bringen, das wir in die Process-Mining-Software laden konnten. Wir filterten die relevanten Informationen aus den Rohdatendateien und verknüpften die Datentabellen anhand der zuvor definierten Verbindungen. Die Ausgabedaten wurden als Ereignisprotokoll formatiert, mit einer eindeutigen ID als Fall-ID, Aktivitätsnamen, Zeitstempeln, Ressourcen und Attributen für jedes Ereignis.

Die Datentransformation wurde mit der Open-Source-Software KNIME durchgeführt. Zur Validierung der transformierten Daten wurde bei jeder Änderung im Datentransformations-Workflow ein Abgleich mit dem Produktivsystem durchgeführt. Diese Validierungsschritte zeigten erhebliches Verbesserungspotenzial, und der Workflow wurde mehrfach angepasst, bis die Ausgabedaten schließlich den Daten des Produktivsystems entsprachen (siehe Abbildung 7 auf der nächsten Seite). Für jedes Ereignis wurde eine eindeutige ID als Fall-ID, Aktivitätsnamen, Zeitstempel, Ressourcen und Attribute verwendet.

Die Datentransformation war der zeitaufwändigste Schritt im Process-Mining-Projekt. Ein Grund dafür war, dass wir keinen direkten Zugriff auf das Produktivsystem hatten. Daher musste die geprüfte Stelle den

Datenvalidierungsprozess unterstützen und bei der Gegenprüfung helfen. Dies führte zu Wartezeiten und Verzögerungen im Projekt.

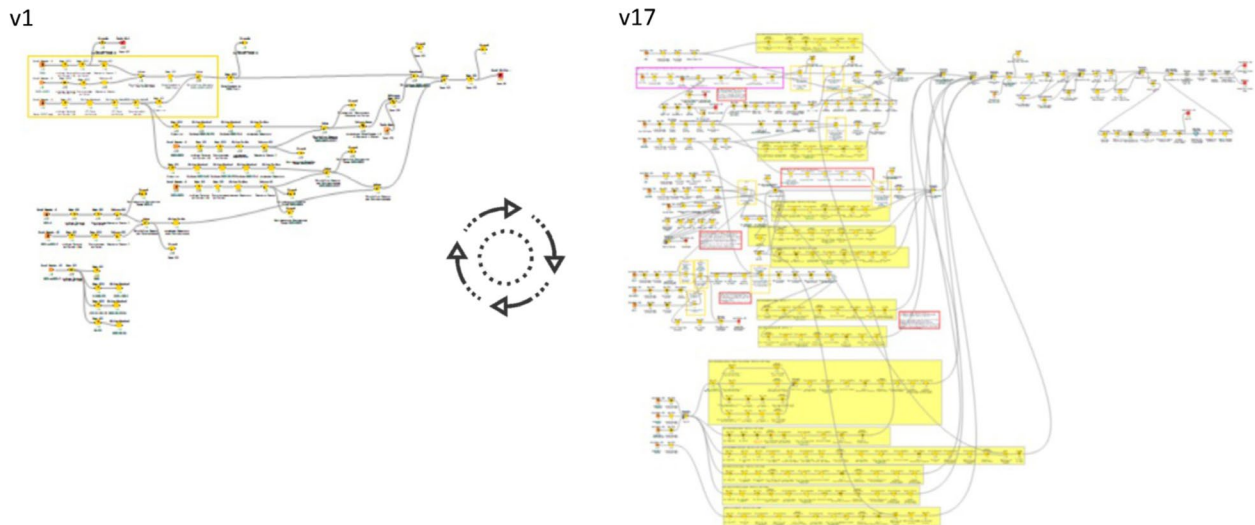
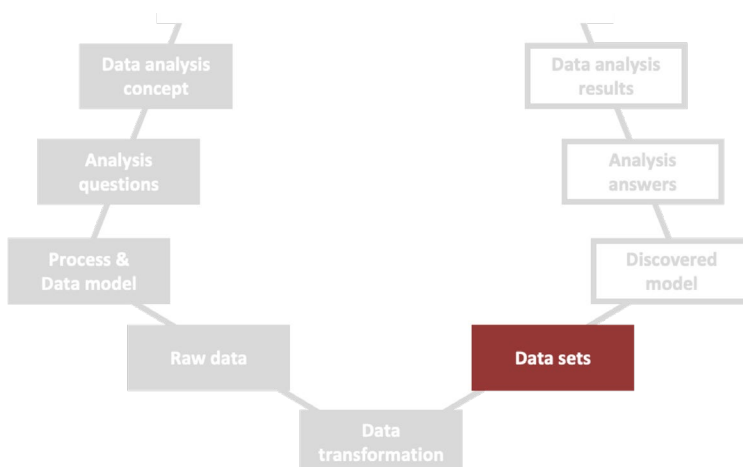


Figure 7: Die erste (links) und letzte (rechts) Version des Datentransformations-Workflows

Ein weiterer Faktor war, dass wir die 1:n- und n:m-Beziehungen bei der Ermittlung der Fall-IDs zunächst nicht ausreichend berücksichtigt hatten. Beispielsweise kann eine Bestellung zu mehreren Rechnungen und Zahlungen führen. Darüber hinaus kann eine Rechnung mehrere Bestellungen betreffen. Eine Zahlung kann mehrere Rechnungen abdecken usw. Diese Viele-zu-viele-Beziehungen [4] mussten bei der Datentransformation angemessen berücksichtigt werden.

Nach mehreren Anpassungen des Transformations-Workflows haben wir alle Validierungsschritte durchlaufen und einen Datensatz generiert, mit dem wir sicher arbeiten konnten.

Step 6: Data sets



Der Datentransformations-Workflow generierte einen Datensatz, den wir für unsere Process-Mining-Analyse nutzen konnten. Demnach wurden zwischen dem 1. Januar 2019 und dem 31. Dezember 2019 insgesamt 2.550 Bestellungen mit einem Bestellwert von ca. 21 Mio. EUR bearbeitet.

Als Fall-ID wurde zunächst die Bestellnummer gewählt. Daher wurden alle Fälle aus der Bestellperspektive analysiert (siehe Abbildung 8).

Case-ID	Timestamp	Activity	Invoice number
1030071289-10	10.01.2019 10:23:17	Create purchase order	
1030071289-10	11.01.2019 13:22:48	Release purchase order	
1030071289-10	23.01.2019 08:58:23	Record invoice	1230007
1030071289-10	26.01.2019 16:46:11	Record invoice	1230008
1030071289-10	25.02.2019 13:33:27	Check invoice	1230007
1030071289-10	25.02.2019 13:48:33	Release invoice	1230007
1030071289-10	01.03.2019 12:16:13	Check invoice	1230008
1030071289-10	01.03.2019 12:19:56	Release invoice	1230008
1030071289-10	14.03.2019 16:00:00	Make payment	1230007
1030071289-10	16.03.2019 14:16:30	Make payment	1230008

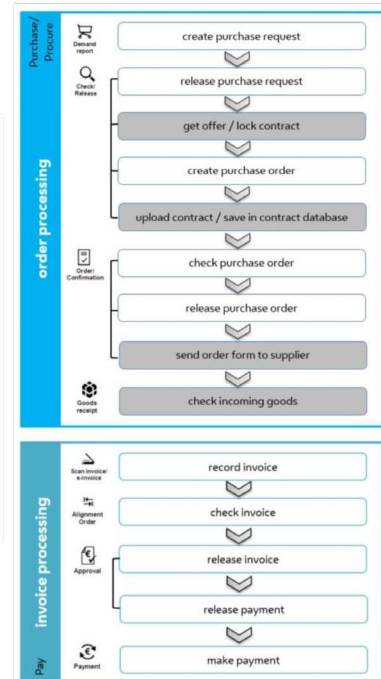


Figure 8: Order perspective (data set left, process view right)

Im Laufe der Analyse wurde jedoch deutlich, dass wir aufgrund der 1:n-Beziehung zwischen Bestellungen und Rechnungen nicht alle unsere Analysefragen zur Rechnungsverarbeitung mit diesem Datensatz beantworten konnten. Abbildung 8 zeigt beispielsweise, dass der Bestellung 1030071289-10 zwei Rechnungen (Rechnung 1230007 und Rechnung 1230008) zugeordnet sind. Für die Aktivitäten „Rechnung prüfen“ und „Zahlung leisten“ gibt es zwei Ereignisse (eines für jede Rechnung). Dies erschwert die Beantwortung von Fragen wie Analysefrage Nr. 7 („Wurden alle Rechnungen vor der Zahlung geprüft?“).

Daher haben wir uns entschieden, einen zweiten Datensatz zu generieren, der sich auf die Rechnungsperspektive konzentriert. Dies wurde erreicht, indem die Bestell- und Rechnungsnummern zu einer neuen Fall-ID zusammengefasst wurden. Der Umfang dieses zweiten Datensatzes ist kleiner (nur Rechnungsstellung und Zahlung). Der Vorteil besteht darin, dass die Aktivitäten im Zusammenhang mit Rechnung 1230007 und 1230008 nun in einem eigenen Fall erscheinen und separat analysiert werden können (siehe Abbildung 9).

Case-ID	Timestamp	Activity
1030071289-10-1230007	23.01.2019 08:58:23	Record invoice
1030071289-10-1230007	25.02.2019 13:33:27	Check invoice
1030071289-10-1230007	25.02.2019 13:48:33	Release invoice
1030071289-10-1230007	14.03.2019 16:00:00	Make payment
Case-ID	Timestamp	Activity
1030071289-10-1230008	26.01.2019 16:46:11	Record invoice
1030071289-10-1230008	01.03.2019 12:16:13	Check invoice
1030071289-10-1230008	01.03.2019 12:19:56	Release invoice
1030071289-10-1230008	16.03.2019 14:16:30	Make payment

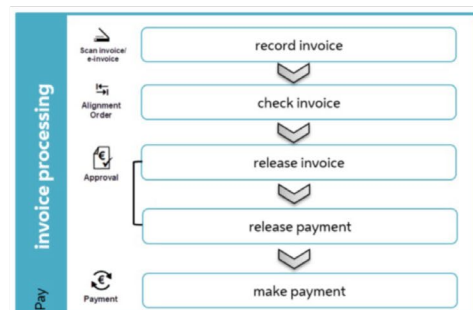
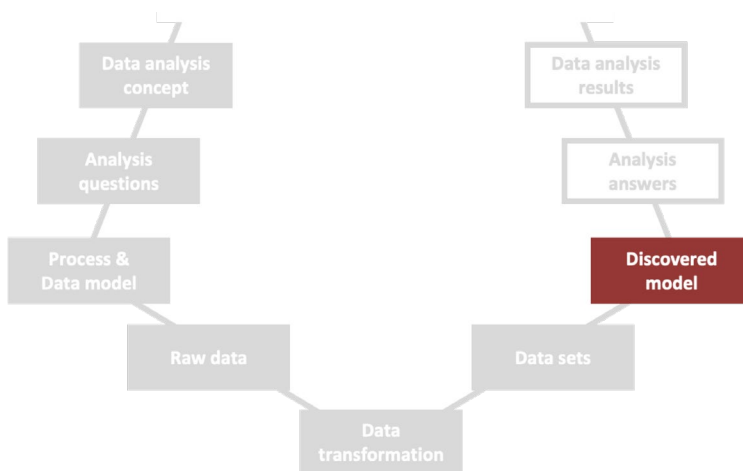


Figure 9: Rechnungsperspektive (Datensicht links, Prozesssicht rechts)

Auf Basis dieser beiden Datensätze – einer aus Auftragssicht und einer aus Rechnungssicht – konnten wir nun mit der Beantwortung unserer Analysefragen beginnen.

Step 7: Entdecktes Prozessmodell (Discovered model)



Nachdem wir Zugriff auf unsere transformierten Datensätze hatten, luden wir diese in die Process-Mining-Software Disco [5] und verschafften uns einen ersten Eindruck von der Komplexität des Prozesses.

Obwohl wir von Anfang an mit Vereinfachungsmethoden gearbeitet und uns auf die Aktivitäten des in Abbildung 5 dargestellten Referenzprozesses konzentriert hatten, um relevante Datentabellen zu identifizieren, war die Prozesslandkarte dennoch sehr komplex. Abbildung 10 zeigt das ermittelte Prozessmodell aus Auftragsperspektive.

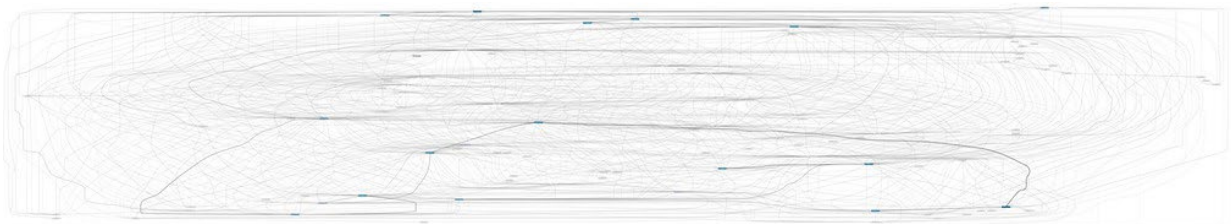


Figure 10: Discovered process model

Aufgrund der hohen Komplexität haben wir weitere Vereinfachungsstrategien angewendet, um eine explorative Analyse und einen Soll-Vergleich der realen Prozesspfade und des Referenzprozesses zu ermöglichen.

Erstens hatten wir durch die Einbeziehung der meisten auffindbaren Zeitstempelfelder eine hohe Anzahl von Aktivitäten aus den Rohdatendateien abgeleitet. Darunter befanden sich administrative Prozessschritte, die außerhalb unseres Referenzprozesses lagen. Wir reduzierten die Anzahl der Aktivitäten, indem wir nur diejenigen Prozessschritte beibehielten, die wir direkt dem übergeordneten Referenzprozess zuordnen konnten (Meilenstein-Vereinfachungsmethode [6]). Dadurch reduzierte sich die Anzahl der Aktivitäten von über 100 auf etwa 50. Zu beachten ist, dass die Daten im IT-System immer noch detaillierter waren als der übergeordnete Prozess. Beispielsweise konnte eine Bestellung auf verschiedenen Ebenen geprüft, abgelehnt und freigegeben werden (siehe Abbildung 1)

Zweitens bestand weiterhin eine hohe Variationsbreite hinsichtlich der Prozesspfade. Daher entschieden wir uns, die Daten in vier Gruppen zu clustern (Methode der semantischen Variantenvereinfachung [7]). Diese vier Gruppen waren:

- (1) stornierte Fälle,
- (2) Fälle ohne Rechnung,
- (3) Fälle mit genau einer Rechnung und
- (4) Fälle mit mehreren Rechnungen.

Durch die separate Betrachtung jedes Datensegments konnte die Anzahl der Prozessvarianten weiter reduziert werden.

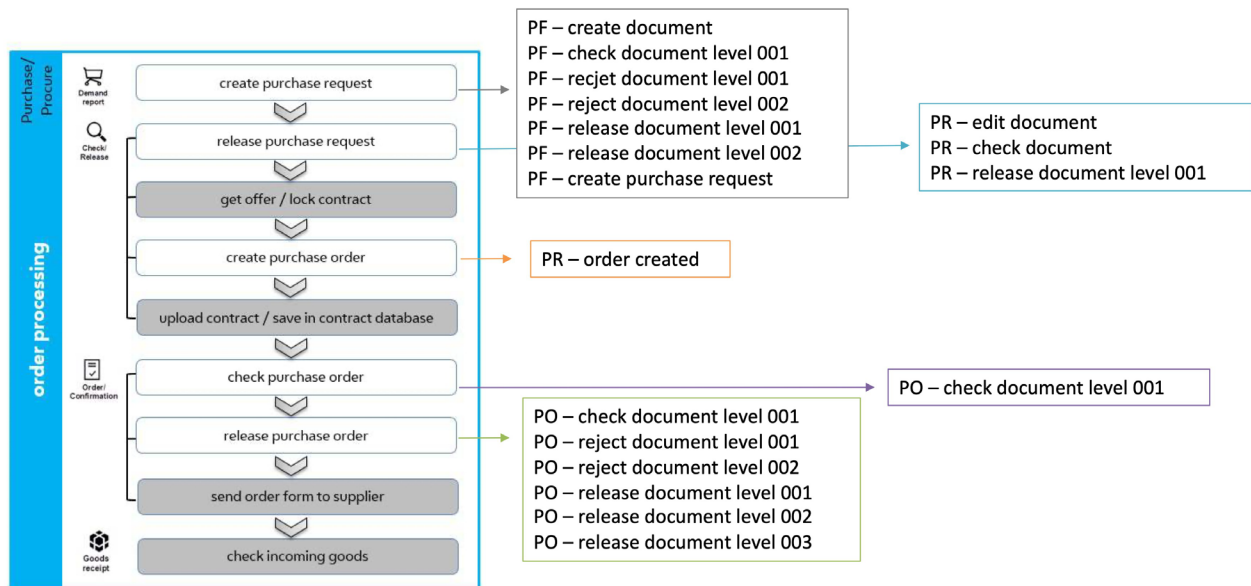


Figure 11: Mapping data to the high-level process

Case Study

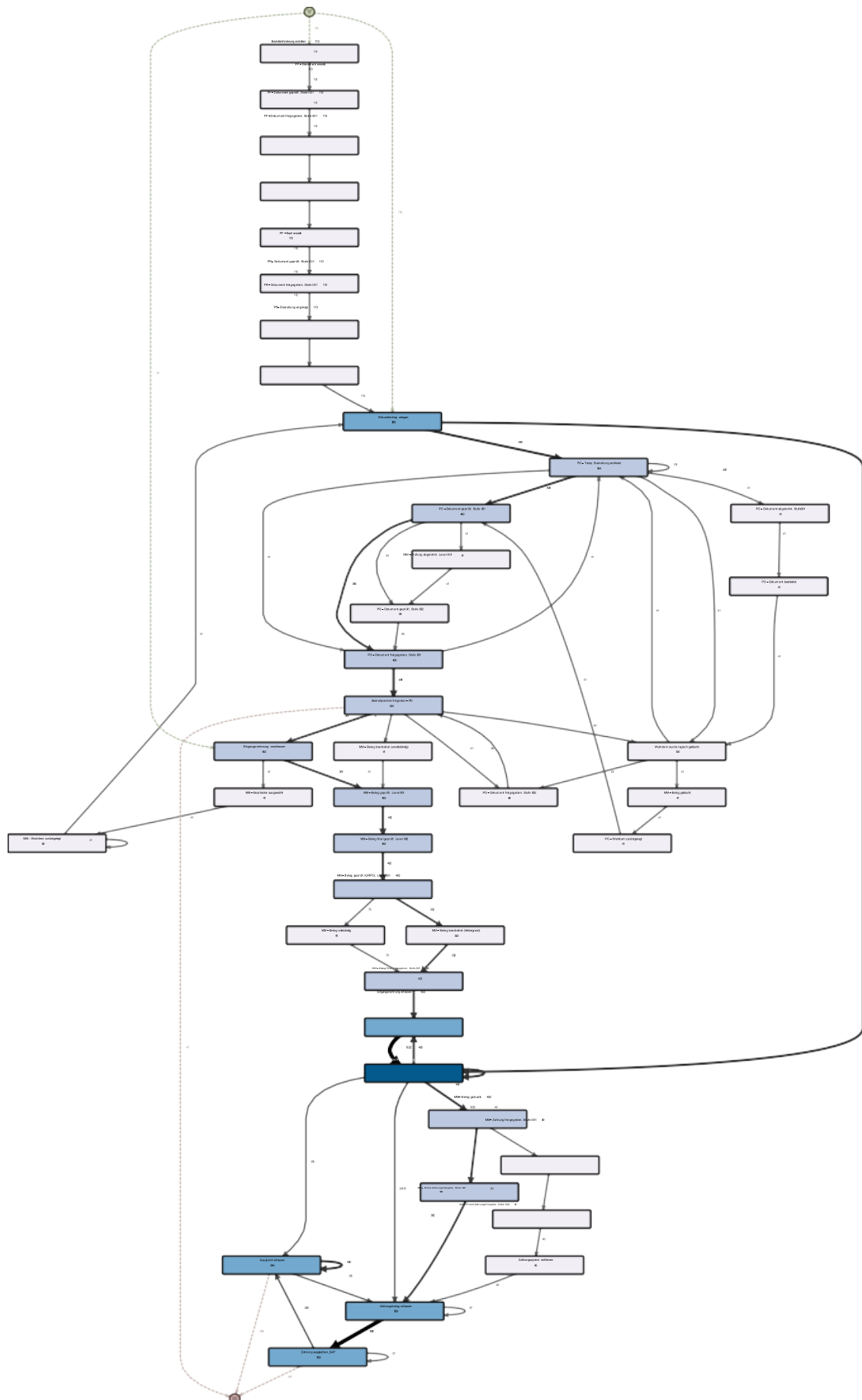
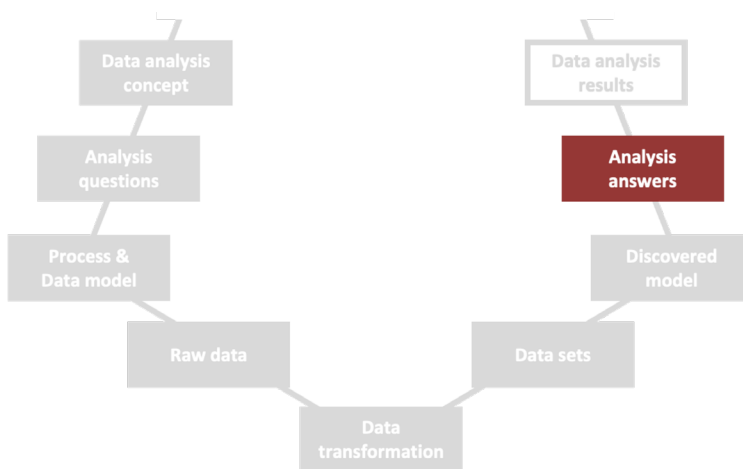


Figure 12: Entdecktes Prozessmodell nach Vereinfachungen

Schließlich konzentrierten wir uns auf die häufigsten Prozesspfade, um einen Überblick über das Mainstream-Verhalten zu erhalten (Methode der Variantenvereinfachung [8]). Abbildung 12 zeigt das ermittelte Modell basierend auf den zehn häufigsten Prozessvarianten. Dies half uns, einen Überblick über den Hauptprozess zu erhalten, bevor wir uns im Detail mit der Analyse der weniger häufigen Pfade und ihrer Abweichungen vom Mainstream-Verhalten befassten.

Durch die Komplexitätsreduzierung konnten wir nun eine explorative Analyse durchführen, nach Inkonsistenzen suchen und unerwartete Prozesspfade detaillierter analysieren.

Step 8: Beantwortung von Analysefragestellungen



Um unsere Analysefragen aus Tabelle 1 zu beantworten, haben wir sowohl eine explorative als auch eine gezielte Analyse verwendet.

Mit der explorativen Analyse untersuchten wir die entdeckten Modelle umfassender auf unerwartete oder ungewöhnliche Prozessverläufe, lange Wartezeiten und andere Auffälligkeiten. Mit dieser Analyse konnten wir die Fragen Nr. 1 („Entsprechen die realen Prozesse dem Soll-Prozess?“) und Nr. 12 („Gibt es Engpässe im Prozess?“) beantworten. Die Fragen Nr. 2 bis 11 wurden durch eine gezielte Analyse beantwortet. Wir übersetzten jede Analysefrage in einen individuellen Filtersatz, basierend auf den Definitionen, die wir im Schritt „Analysefragen“ erstellt hatten. Obwohl wir die Analysefragen bereits detailliert definiert hatten, verfeinerten wir diese Spezifikationen weiter, um sicherzustellen, dass wir nur Fälle erkannten, die gegen die Prozessanforderungen verstießen.

Ziel der Frage Nr. 2 („Werden alle Aufträge freigegeben?“) war beispielsweise herauszufinden, ob es Fälle mit fehlender Auftragsfreigabe gab. Der Zielwert für diese Analyse lag bei 100 %. Es könnte jedoch berechnete Gründe für eine fehlende Auftragsfreigabe geben. Wir kamen zu dem Schluss, dass nur ausgeführte Aufträge eine Freigabeaktivität benötigten. Darüber hinaus könnte es Aufträge geben, die nie ausgeführt wurden. Daher haben wir beschlossen, stornierte Fälle aus der Datenbasis für diese Frage auszuschließen. Daher haben wir zur Beantwortung von Frage Nr. 2 („Werden alle Bestellungen freigegeben?“) eine Kombination aus drei verschiedenen Filtern verwendet. Zunächst haben wir alle Bestellungen ausgeschlossen, die storniert wurden. In unserem Datensatz bedeutete dies, dass wir Fälle mit dem Attributwert L oder X im Attribut „Löschkennzeichen“ entfernt haben (siehe Abbildung 13 auf der nächsten Seite).

Zweitens war es nicht einfach festzustellen, ob eine Bestellung tatsächlich ausgeführt wurde. Es lagen uns keine Daten vor, die zeigten, wann ein Bestellformular an den Lieferanten gesendet wurde. Es gab auch keine Daten darüber, wann der Vertrag abgeschlossen oder die Wareneingänge geprüft wurden. Daher wählten wir die Rechnung als Messgröße, um zu messen, ob eine Bestellung tatsächlich ausgeführt wurde. Abbildung 14 zeigt den zweiten Filter, der nur Bestellungen mit einer Eingangsrechnung erfasst.

Objective	Performance		Compliance	x	Experience	
Analysis question	Are all orders released?					
Metric	Release activity needs to be documented					
Target value	100 % of orders are released					
Process scope	PO – document released level 1 to n, order position released					
Data set used	Order perspective					
Influencing factors	- Value limits (four-eyes principle) - Cancellations					

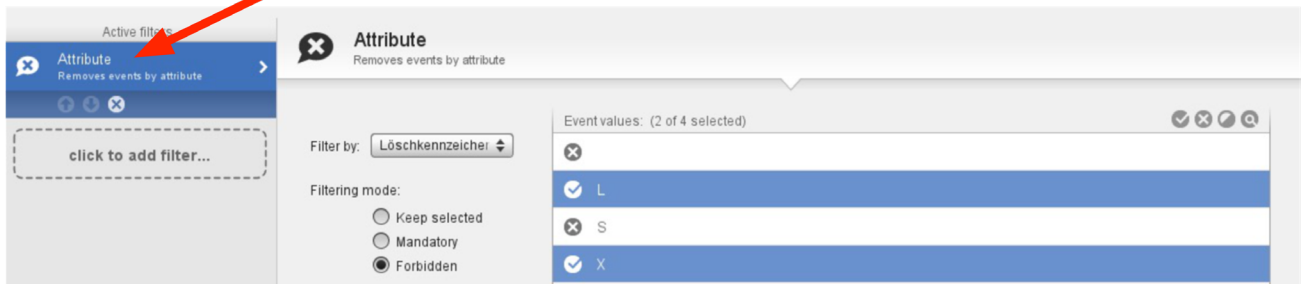


Figure 13: Excluding canceled orders

Objective	Performance		Compliance	x	Experience	
Analysis question	Are all orders released?					
Metric	Release activity needs to be documented					
Target value	100 % of orders are released					
Process scope	PO – document released level 1 to n, order position released					
Data set used	Order perspective					
Influencing factors	- Value limits (four-eyes principle) - Cancellations - Order was really placed					

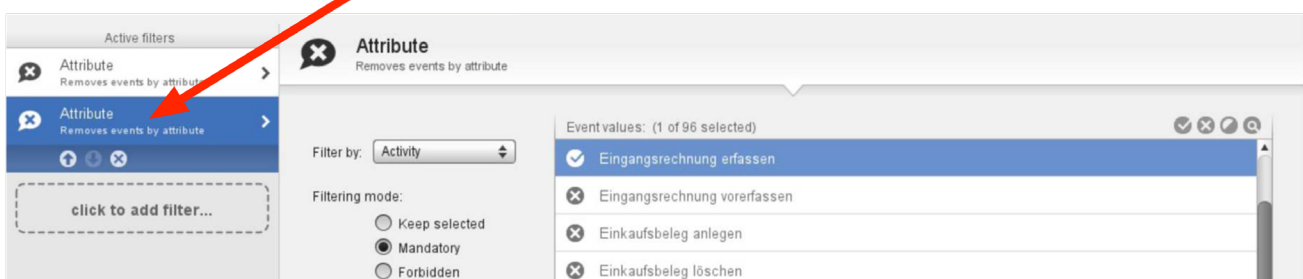


Figure 14: Including only orders with an invoice

Drittens, da wir Verstöße gegen Frage Nr. 2 finden wollten, schlossen wir alle Aufträge mit einer Auftragsfreigabeaktivität aus. Wir suchten also nach dem Gegenteil des geforderten Verhaltens im Prozess. Abbildung 15 zeigt diesen dritten Filter, der alle Fälle ausschließt, die eine der ausgewählten Auftragsfreigabeaktivitäten („PO - Dokument freigeben. Stufe 001“ - „... 003“) enthalten.

Objective	Performance		Compliance	x	Experience	
Analysis question	Are all orders released?					
Metric	Release activity needs to be documented					
Target value	100 % of orders are released					
Process scope	PO – document released level 1 to n, order position released					
Data set used	Order perspective					
Influencing factors	- Value limits (four-eyes principle) - Cancellations - Order was really placed					

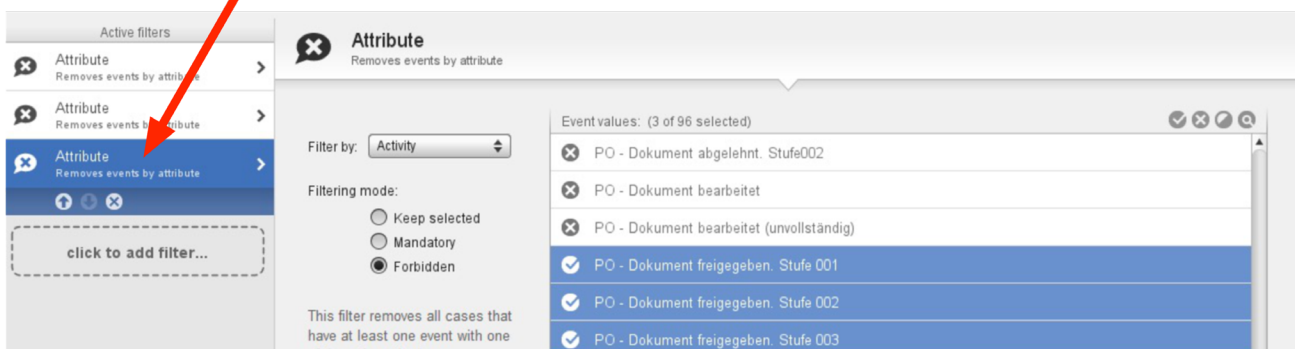
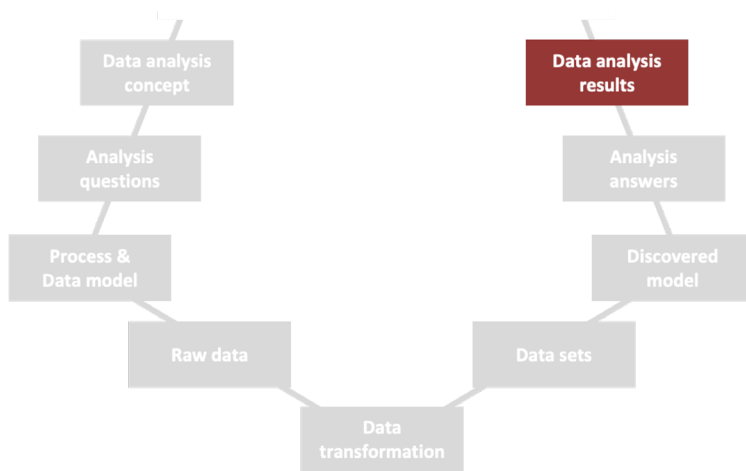


Figure 15: Excluding orders with an order release activity

Durch die Anwendung dieser drei Filter stellten wir fest, dass in 17 % aller nicht stornierten Fälle eine Auftragsfreigabe fehlte, obwohl eine Rechnung erfasst war. Bei genauerer Betrachtung dieser Fälle stellten wir fest, dass sich alle Bestellungen auf die Warengruppe Bürobedarf bezogen. Die geprüfte Einheit erklärte, dass Bürobedarf nicht den regulären Auftragsfreigabe-Workflow durchlaufen müsse und ohne vorherige Auftragsfreigabe bestellt werden könne. Somit entsprachen auch diese Bestellungen trotz der festgestellten Unregelmäßigkeiten den Beschaffungsrichtlinien.

Analog zum obigen Beispiel beantworteten wir die verbleibenden Analysefragen jeweils durch die Übersetzung in individuelle Filtereinstellungen. Zusätzlich besprachen wir alle festgestellten Unregelmäßigkeiten mit der geprüften Stelle, um zu klären, ob es einen tatsächlichen Grund für die Abweichung gab oder ob wir einen Compliance-Verstoß im Prozess festgestellt hatten.

Step 9: Ergebnisse der Datenanalysen



Die Ergebnisse der Datenanalyse fassten die Antworten auf die Analysefragen zusammen und bildeten die Grundlage für unseren Prüfungsbericht, einschließlich der Feststellungen und Empfehlungen.

Abbildung 16 zeigt, dass wir alle unsere Fragen beantworten konnten. Lediglich die Fragen Nr. 1 und Nr. 12 konnten nicht vollständig beantwortet werden, da wir einerseits aufgrund fehlender Daten einige blinde Flecken im Prozess hatten und andererseits den Prozess aufgrund seiner Komplexität deutlich vereinfachen mussten. Daher haben wir ihren Status auf Gelb gesetzt.

Nr.	Analysis questions	Performance	Compliance	Could be answered
1	Do the real processes fit the should-process?		x	● ● ●
2	Are all orders released?		x	● ● ●
3	Have all training orders been approved by the HR department?		x	● ● ●
4	Is an invoice recorded for every order?		x	● ● ●
5	Are all orders payed?		x	● ● ●
6	Are all invoices payed?		x	● ● ●
7	Have all invoices been checked before payment?		x	● ● ●
8	Are orders payed although they were cancelled or deleted?		x	● ● ●
9	Is the four-eyes principle maintained?		x	● ● ●
10	Were all invoices that exceeded the order value released according to the four-eyes principle?		x	● ● ●
11	Are the payment targets observed?	x		● ● ●
12	Are there any bottlenecks within the process?	x		● ● ●

Figure 16: All our analysis questions could be answered

Im Rahmen der explorativen Analyse stellten wir fest, dass im Jahr 2019 etwa 1 % aller Bestellungen storniert wurden. Darüber hinaus wurde in 1,4 % aller Fälle eine Bestellung im Informationssystem erfasst, ohne dass weitere Aktivitäten erfasst wurden. Auch diese Fälle schienen storniert zu werden, wurden jedoch nicht im Informationssystem dokumentiert. Obwohl der Gesamtanteil der stornierten Bestellungen

Im normalen Bereich lag, identifizierten wir Verbesserungspotenzial hinsichtlich der Dokumentationsstornierter Fälle.

Darüber hinaus stellten wir fest, dass etwa 7 % aller Fälle mit einer Rechnungserfassung begannen (die entsprechende Bestellung wurde erst später aufgegeben). Diese Fälle deuteten darauf hin, dass der formale Beschaffungsprozess nicht immer eingehalten wurde. Stattdessen tätigten einige Mitarbeiter Einkäufe ohne vorherige Genehmigung – ein unerwünschtes Verhalten, das als „Maverick Buying“ bezeichnet wird.

Im Rahmen der gezielten Analyse stellten wir fest, dass 17 % aller fakturierten Bestellungen nicht freigegeben wurden. Diese Abweichung war jedoch tolerierbar, da sie alle zur Produktgruppe Bürobedarf gehörten, für die keine Auftragsfreigabe erforderlich war.

Die Beschaffungsrichtlinien der geprüften Einheit sahen zudem vor, dass alle Bestellungen mit einem Bestellwert von über 20.000 EUR nach dem Vier-Augen-Prinzip freigegeben werden mussten. Unsere Analyse ergab, dass das Vier-Augen-Prinzip bis auf einen Fall eingehalten wurde, sobald der Bestellwert von 20.000 EUR überschritten wurde. In diesem Fall war eine zweistufige Freigabe im Informationssystem dokumentiert, jedoch war eine Person nicht berechtigt, eine Bestellung dieses Werts freizugeben. Diese Abweichung vom definierten Freigabeprozess deutete auf eine mögliche Schwachstelle im internen Kontrollsystem hin. Daher empfahlen wir, die Freigabeberechtigungen regelmäßig zu überprüfen.

Die Beschaffungsrichtlinien verlangten für die Rechnungsverarbeitung eine Prüfung aller Rechnungen vor der Zahlung. Unsere Process-Mining-Analyse zeigte, dass diese Prüfungen im Informationssystem nicht einheitlich dokumentiert waren. Somit war nicht transparent, ob die erforderlichen Prüfungen vor der Zahlung durchgeführt wurden. Darüber hinaus war bei der Rechnungsfreigabe ein Vier-Augen-Prinzip vorgeschrieben, sobald der Rechnungsbetrag mehr als 10 % bzw. 1.000 EUR vom Bestellwert abwich. In Einzelfällen wurde dieses Vier-Augen-Prinzip nicht eingehalten. Daher empfahlen wir die Implementierung weiterer Kontrollmaßnahmen, um sicherzustellen, dass alle Rechnungen vor der Freigabe zur endgültigen Zahlung ordnungsgemäß geprüft wurden.

Unsere Process-Mining-Analyse zeigte außerdem, dass etwa die Hälfte aller Rechnungen sofort fällig war, obwohl das allgemeine Zahlungsziel 30 Tage nach Rechnungserhalt betrug. Somit bestand Verbesserungspotenzial bei der Dokumentation der Zahlungsziele.

Schließlich wurde das Zahlungsziel in rund 6 % aller Fälle nicht eingehalten. Die geprüfte Partei begründete dies damit, dass in einigen Fällen noch Klärungsbedarf hinsichtlich des Wareneingangs oder der Leistungskonformität zum Zeitpunkt der Fälligkeit einer Rechnung bestand. Weitere Untersuchungen zeigten zudem, dass es bei Nichteinhaltung der Zahlungsziele häufig zu Verzögerungen bei der Freigabe der Rechnung zur Schlusszahlung kam. Wir empfahlen, die Einhaltung der Zahlungsziele zu überwachen und gegebenenfalls Maßnahmen zur Verkürzung der Wartezeiten zu ergreifen.

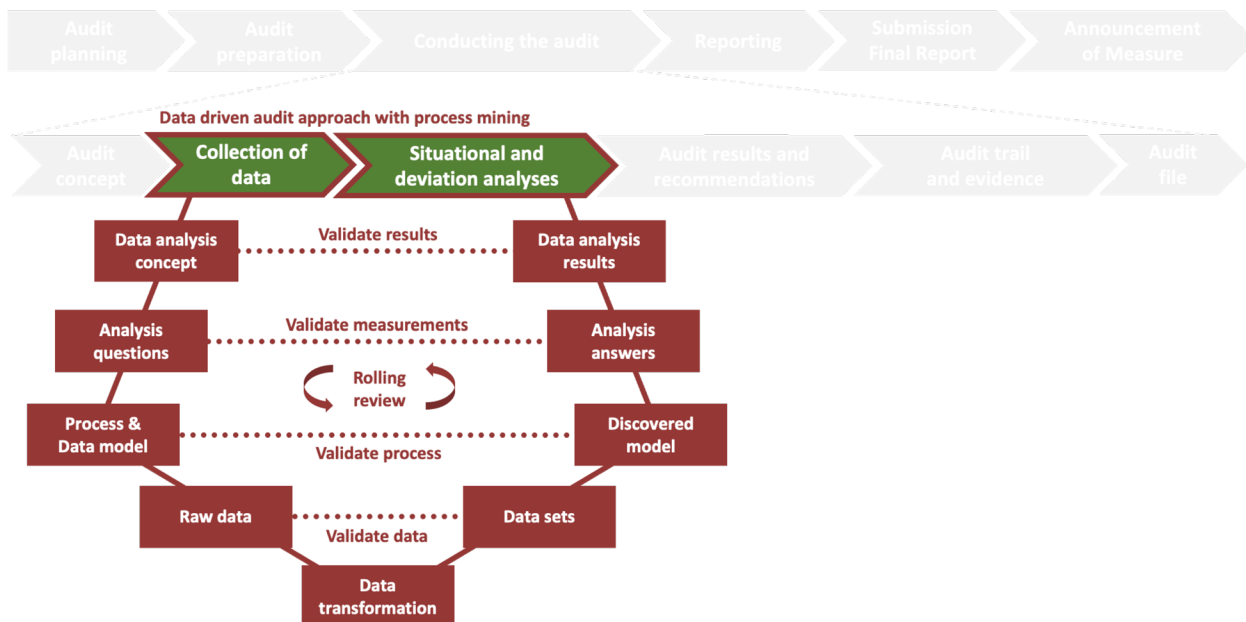


Figure 17: Rolling Review mit Validierungsschritten auf mehreren Ebenen

In dieser Fallstudie folgten wir dem 9-Schritte-Modell aus Abbildung 3, um Process Mining in unserem Audit anzuwenden. Im Laufe unserer Arbeit stellten wir fest, dass die neun Schritte nicht in einer strengen Reihenfolge abliefen. Die Erkenntnisse aus späteren Projektphasen konnten wir häufig nutzen, um die Ergebnisse früherer Schritte zu verbessern. Beispielsweise überarbeiteten wir unsere Analysefragen mehrfach, da wir in späteren Projektphasen neue Informationen zur Datenverfügbarkeit und -qualität erhielten. Das Modell ist daher als iterativer Ansatz zu verstehen, bei dem eine rollierende Überprüfung Validierungsschritte auf mehreren Ebenen umfasst. Abbildung 17 zeigt die wichtigsten Validierungsschritte:

- Zunächst überprüfen wir, ob die als Input für die Process-Mining-Analyse verwendeten Datensätze mit den Rohdaten und den Daten des Produktsystems übereinstimmen. In diesem Validierungsschritt stellten wir sicher, dass unsere Datenanalyse auf einer zuverlässigen Datenquelle basierte.
- Anschließend validierten wir den Prozess selbst, indem wir überprüften, ob das ermittelte Modell dem in einer früheren Projektphase definierten Prozessmodell entspricht. Dabei untersuchten wir, ob wir die korrekten Daten verwendet hatten oder ob Anpassungsbedarf bestand.
- Als Nächstes überprüften wir, ob die Analyseantworten alle Analysefragen abdeckten. So konnten wir feststellen, ob noch unbeantwortete Analysefragen – sei es aufgrund fehlender Daten oder aufgrund fehlender Antworten – unbeantwortet blieben.
- Abschließend validierten wir die Ergebnisse der Datenanalyse, indem wir überprüften, ob sie den Anforderungen des Datenanalysekonzepts entsprachen und das primäre Prüfungsziel ausreichend berücksichtigt hatten. Anhand dieser Bewertung konnten wir feststellen, ob die finale Analyse die von uns geplanten Prüfungsziele abdeckte.
- All diese Validierungsschritte halfen uns, zuverlässige Ergebnisse mit einem gewissen Maß an Sicherheit und Qualität zu erzielen und die Ergebnisse des Process-Mining-Projekts zu verbessern.

Herausforderungen und Einschränkungen

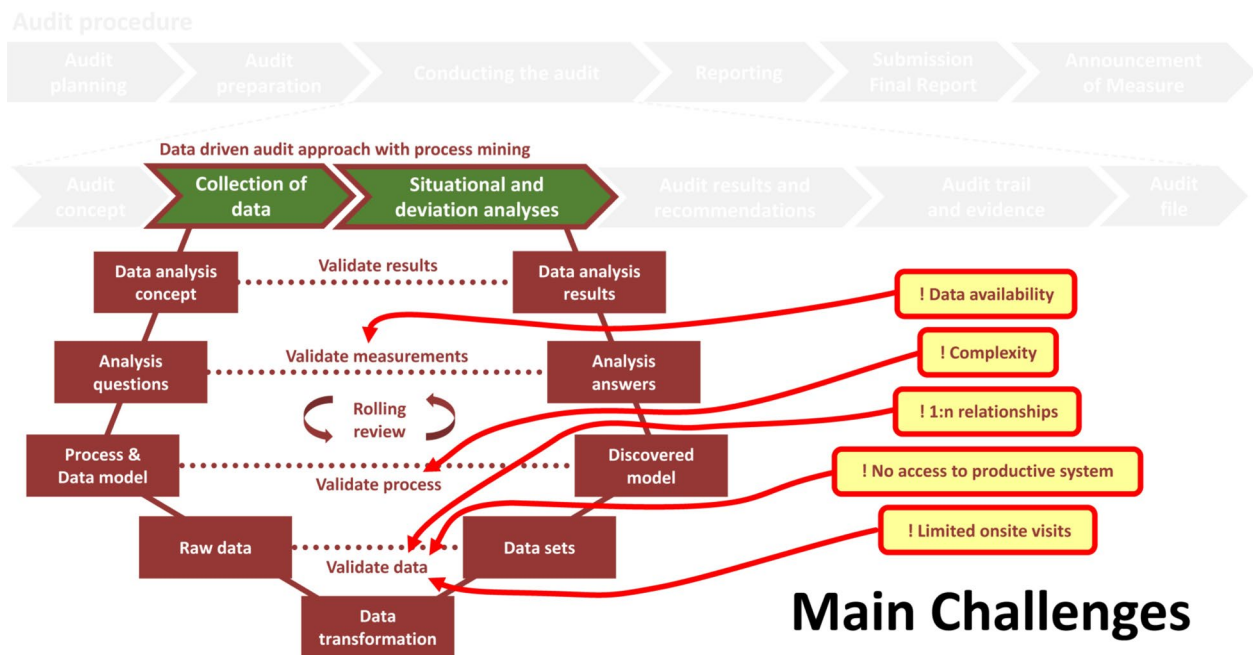


Figure 18: Herausforderungen und Einschränkungen

Während unseres Process-Mining-Projekts sind wir auf eine Reihe von Herausforderungen gestoßen (siehe Abbildung 18).

Die Datenaufbereitung stellte eine der größten Herausforderungen dar. Wie bereits erwähnt, lagen nur für einige der von uns betrachteten Prozessschritte Daten vor. Daher konnten wir nicht alle Analysefragen beantworten, die uns ursprünglich vorschwebten.

Die verfügbaren Daten waren auf verschiedene Tabellen verteilt, die miteinander verknüpft werden mussten. 1:n- und n:m-Beziehungen erschwerten die Datenaufbereitung, und wir mussten mehrere Versionen des Datentransformations-Workflows implementieren, bevor dieser zuverlässige und validierte Daten lieferte, denen wir vertrauen konnten.

Der fehlende direkte Zugriff auf die Produkktivsysteme machte die Datenvalidierung zusätzlich zeitaufwändig, da wir von der geprüften Partei abhängig waren, die die Daten zur Gegenprüfung bereitstellen musste. Da wir das Audit während der COVID-19-Pandemie durchführten, waren aufgrund von Kontaktbeschränkungen nur eingeschränkte Vor-Ort-Besuche möglich.

Die hohe Komplexität des analysierten Prozesses und die große Anzahl von Ereignissen in den Datensätzen machten die explorative Analyse zu einer großen Herausforderung. Daher waren die Analysefragen Nr. 1 („Entsprechen die realen Prozesse dem Soll-Prozess?“) und Nr. 12 („Gibt es Engpässe im Prozess?“) schwer zu analysieren. Wir mussten die Daten vereinfachen, um die Komplexität auf ein Niveau zu reduzieren, das den Prozess analysierbar machte. Letztendlich war es unmöglich, einen konkreten Prozentsatz zu nennen, wie viele Fälle dem Soll-Prozess entsprachen bzw. nicht.

Die 1:n- und n:m-Beziehungen machten es notwendig, für unterschiedliche Analysefragen mit unterschiedlichen Datensätzen zu arbeiten. Daher führten wir unsere Process-Mining-Analysen sowohl aus der Auftrags- als auch aus der Rechnungsperspektive durch.

Trotz der oben genannten Herausforderungen und Einschränkungen lieferte uns die Process-Mining-Analyse einen hervorragenden Einblick in die tatsächliche Durchführung des Purchase-to-Pay-Prozesses. Somit übertraf der Nutzen des Process-Mining-Einsatzes im Audit (siehe Abbildung 19) die Herausforderungen, die wir bewältigen mussten.

Vorteile des Process-Mining-Einsatzes im Prüfungswesen

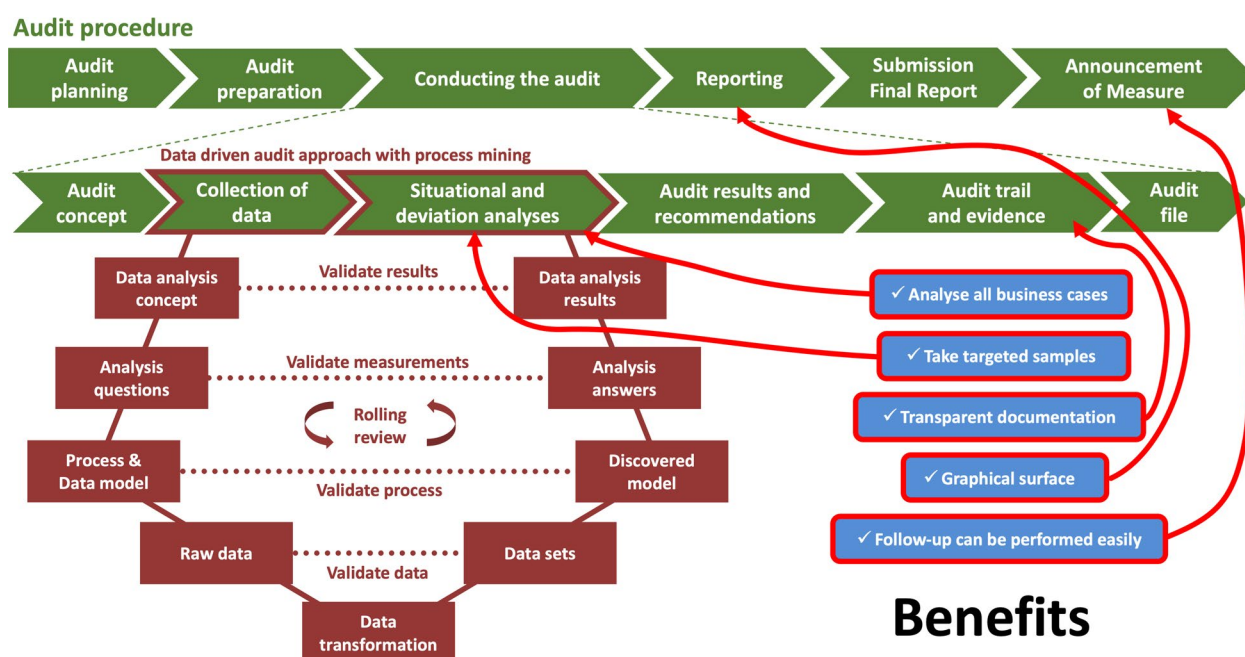


Figure 19: Vorteile des Process-Mining-Einsatzes im Audit

Im Rahmen unserer Prüfungstätigkeit sind wir häufig mit riesigen Datenmengen und einer hohen Anzahl von Fällen konfrontiert. Da unsere Ressourcen begrenzt sind und wir unsere Prüfungen innerhalb eines bestimmten Zeitrahmens abschließen müssen, wählen wir traditionell eine Stichprobe von Fällen aus den relevanten Prozessdaten aus und untersuchen diese Fälle detaillierter. Mithilfe von Process Mining müssen wir keine Stichprobe mehr ziehen. Stattdessen können wir alle Fälle unabhängig von der Gesamtzahl analysieren. So konnten wir beispielsweise im Audit dieser Studie eine vollständige Untersuchung aller 2.550 Fälle durchführen und Aussagen zu Bestell- und Rechnungsfreigaben für alle Bestellungen aus dem Jahr 2019 treffen.

Wir setzten weiterhin Stichprobenverfahren ein, um gezielt Fälle zu untersuchen, in denen wir Unregelmäßigkeiten entdeckten. Dadurch konnten wir unsere Ressourcen effizienter einsetzen. Unserer Erfahrung nach führte dies zu einer höheren Qualität der Prüfungsergebnisse und einer höheren Sicherheit bei der Erkennung potenzieller Schwachstellen in den int. Kontrollsystemen.

Dank enormer Datenmengen ist es zudem oft schwierig, die Analyseergebnisse für alle verständlich darzustellen. Die grafische Oberfläche der Process-Mining-Software ist in dieser Hinsicht sehr hilfreich. Wir konnten die Analyseschritte problemlos im Beisein des geprüften Partners durchführen und so transparent machen, welche Maßnahmen wir ergriffen haben, um zu dem jeweiligen Ergebnis zu gelangen. Dieser Vorteil ist natürlich noch größer, wenn der geprüfte Partner ebenfalls Erfahrung mit Process Mining hat. In diesem Fall kann er die Analyse nachvollziehen und messen, ob die aufgrund der Empfehlungen des Prüfers vorgenommenen Änderungen die erwartete Wirkung haben.

Schließlich kann der Einsatz von Process Mining zur Bewertung der Auswirkungen von Prozessänderungen auch aus Sicht des Prüfers sehr hilfreich sein. Bei einer Folgeprüfung kann Process Mining erneut eingesetzt werden, um alle Fälle umfassend zu untersuchen und zu überprüfen, ob die implementierten Änderungen die Prozessqualität verbessert haben.

Wie dieser Artikel zeigt, kann der Einsatz von Process Mining in einer Prüfung sehr vorteilhaft sein und tiefere Einblicke in den jeweiligen Prozess ermöglichen. Der Einstieg in Process Mining in der Prüfungsarbeit ist zweifellos eine Herausforderung, wird aber mit zunehmender Erfahrung einfacher. Wir setzen Process Mining seit 2016 in unseren Prüfungen ein und arbeiten seitdem kontinuierlich an der Verbesserung unserer Praxis. Jedes Process-Mining-Projekt bietet uns die Chance, unseren Ansatz zu verbessern und den Prüfpfad transparenter zu gestalten.

Mit diesem Artikel möchten wir andere Prüfer ermutigen, mehr über Process Mining zu erfahren und es in ihre Prüfungsmethodik zu integrieren. Wir schätzen den Austausch mit unseren Kollegen und laden Sie ein, uns über die unten stehenden Kontaktdaten zu kontaktieren, um Ihre Erfahrungen mit uns zu teilen.

Autoren

Jasmine Handler, MA MSc - jasmine.handler@wien.gv.at (corresponding author)

Ing. Dipl.Ing.(FH) Andreas Preslmayr, MSc - andreas.preslmayr@wien.gv.at

Referenzen

- [1] Stadtrechnungshofausschuss Periode 2021. Sitzung vom 02. Dezember 2021
URL: <https://www.stadtrechnungshof.wien.at/ausschuss/05/inhaltsverzeichnis.htm>
- [2] StRH I - 8/20. WIENER STADTWERKE GmbH, Prüfung von SAP Beschaffungsprozessen.
URL: <https://www.stadtrechnungshof.wien.at/berichte/2021/lang/05-11-StRH-I-8-20.pdf>
- [3] KNIME Analytics Platform. URL: <https://www.knime.com/knime-analytics-platform>
- [4] Process Mining in Practice. Many-to-many relationships in process mining:
<https://fluxicon.com/book/read/datasuitability/#many-to-many-relationships-between-different-case-ids>
- [5] Process mining software Disco. URL: <https://fluxicon.com/disco/>
- [6] Process Mining in Practice. Milestone simplification method: <https://fluxicon.com/book/read/simplification/#strategy-9-focusing-on-milestone-activities>
- [7] Process Mining in Practice. Semantic variant simplification method: <https://fluxicon.com/book/read/simplification/#strategy-5-semantic-process-variants>
- [8] Process Mining in Practice. Variant simplification method: <https://fluxicon.com/book/read/simplification/#strategy-2-focusing-on-the-main-variants>