

# Process Mining für jedermann

Stephan La Rocca

PITSS GmbH

Bielefeld

## Schlüsselworte

Process Mining, Prozess Discovery, Business Process

## Einleitung

Process Mining versteht sich in der Evolution des Business Intelligence Ecosystems als legitimer Nachfolger von „Data Mining“ und „Predictive Analysis“, führt aber noch ein deutliches Schattendasein. Dabei steckt in dem Ansatz, die Prozesse, die sich hinter der Änderung von unternehmensrelevanten Daten verbergen, genauer betrachten, analysieren und vergleichen zu können, enormes Potential. Es ist eine Art Schulterchluss von IT und Business Abteilung, der im Zeitalter der Digitalisierung notwendig ist.

## Motivation

Warum eigentlich Process Mining? Sollte sich ein Unternehmen nicht bewusst sein, welche Prozesse systemtechnisch unterstützen werden und welche nicht? In den meisten Fällen sicherlich, aber sind Sie wirklich sicher? Könnten Sie adhoc Fragen beantworten wie:

Ist meine Bestellung in 85% aller Fälle innerhalb von zwei Tagen erfüllt? Und was sind die Gründe, in den Ausnahmen? Gibt es andere Möglichkeiten, um den Auftrag zu erfüllen, und werden diese verwendet? Wenn ja, wie oft? Sind alle Alternativen im System abgebildet?

Process Mining hilft solche und ähnliche Fragestellungen zu beantworten und kontinuierlich zu monitoren und deren Antworten zu verbessern. In größeren Unternehmen kommt dann recht schnell die Fragestellung hinzu, ob Prozesse standortübergreifend harmonisiert sind? Und auch hier, wenn nicht, folgt die Frage nach dem Warum.

Die ersten Erfahrungen zeigen darüber hinaus sehr interessante weitere Einsatzgebiete. Als Software-Hersteller z.B. möchten Sie wissen, wie ihre unterschiedlichen Kunden ihre Software verwenden um dadurch auf vermeintliche Schwachstellen, Schulungsbedarfe oder Potential für das nächste Release rückschließen zu können. Als Inhouse-Lösungsanbieter hingegen eruieren Sie vielleicht Standard-Lösungen oder planen einen Invest in eine neue Software-Entwicklung – hier drängt sich eine Betrachtung auf Prozessebene nahezu auf. Sind alle bestehenden Prozesse in der neuen Welt abgedeckt? Was ist die Grundlage für eine Gap-Analyse oder die Requirement Definition? Diese Fragen sollten mit den bestehenden und verwendeten Systemen beantwortet werden - nicht mit den alten Pflichtenheften in Papierform, die den Staub der Jahre tragen.

Ein letztes Beispiel für Process Mining möchte ich mit der gerade nicht behaglichen Situation eines Audits anführen. Um wieviel einfacher würden solche Prüfungssituationen ablaufen, wenn Sie mit Process Mining aufzeigen könnten, dass alle Prozesse in Ihrem Unternehmen in 99% aller Fälle die vorgeschriebenen Wege aus den Qualitätsnormen durchlaufen. In den Process Mining Typen bezeichnet man diesen Vergleich zwischen einem vorhandenen Modell und dem tatsächlichen Ablauf der Prozesse als „Conformance“. Das wäre die Aufgabe eines Mausclicks.

Weitere Typen sind „Discovery“, der aktuell weit verbreitete Typ, der dazu dient, aus den Daten die dahinterliegenden Prozesse zu identifizieren. Weiterhin wird vereinzelt nach der Analyse und dem Vergleich auch noch auf den Typ „Enhancement“ zurückgegriffen, der dann die identifizierten Verbesserungen modelliert und zurückführen soll.

## Was sind die tatsächlichen Grenzen im Process Mining?

Anwendungen sind auf die Bedürfnisse des Benutzers, des Anwendungsfalls und der Unterstützung der Geschäftsprozesse implementiert. Sie sind nicht darauf ausgelegt, sinnvolle Daten mit sehr hoher

Qualität für ein Process Mining bereitzustellen, damit in diesem Tool Analysen und Visualisierungen durchgeführt werden können.

Um an Rohdaten für das Process Mining zu kommen, haben sich drei verschiedene Wege etabliert:

Der einfachste Ansatz besteht darin, bestehende Logfiles zu konsumieren, um mit Audit-Daten oder Workflow-Tracking-Informationen, den zugrundeliegenden Prozess aus diesen Audit-Daten zu rekonstruieren. Gerade workflowbasierte Systeme, z.B. auf Basis von BPEL sind perfekt dafür geeignet. Die Qualität der Ergebnisse ist direkt mit der Existenz und der Datenqualität dieser Audit-Protokolle verbunden. Leider sind nicht alle Legacy-Anwendung in der Lage, diese Daten zu liefern. Sogar Datenbankzentrierte Anwendungen liefern in der Regel keine Audit-Daten auf Transaktionsebene in dieser notwendigen Art der Granularität.

In diesem Fall greift der zweite Ansatz, bei dem alle Key-User und Prozessleiter in umfangreichen Interviews gebeten werden, den Idealzustand der Prozesse zu beschreiben. Parallel dazu wird jedes Mal das Datenmodell geparkt, um die Auswirkung des beschriebenen Prozesses innerhalb des Modells zu erkennen. Wird also zu einem späteren Zeitpunkt das Datenmodell analysiert, können an Hand der Daten die dahinterliegenden Prozesse rückwirkend angenommen werden. Dieser Ansatz ist sehr zeitaufwendig und gepflastert mit Missverständnissen. Daten könnten durch nicht analysierte Ereignisse anderweitig geändert worden sein. Zusätzlich stellen die Interviews keine Vollständigkeit der relevanten Prozesse sicher, ganz zu schweigen von der Tatsache, dass vielfach Prozesse keine signifikanten, eindeutigen Änderungen im Datenmodell hervorheben. Es ist also mehr ein Process Guessing als ein Process Mining.

Dann hilft vielleicht der dritte Ansatz, der ein bisschen die „Beweislast“ umdreht. Selbst noch recht jung in der ohnehin schon neuen Disziplin Process Mining ist der Versuch, ähnlich wie beim Data Mining mit noch weiter entwickelten Algorithmen aus dem Bereich der künstlichen Intelligenz Prozessmuster in den verschiedenen Pools von Daten zu identifizieren. Dafür benötigen diese Algorithmen eine tiefe Analyse und Lernphase, bevor sie mit Process Mining auf den Rohdaten losgelassen werden können. Ein sehr innovativer aber auch sehr zeitaufwändiger und kostenintensiver Ansatz.

Alle Methoden haben als Zwischenziel, dass die zu analysierenden Systeme einen möglichst akkuraten Eventstream an das Process Mining Tool übergeben. Als Format für einen Eventstream hat eine Arbeitsgruppe der IEEE-Organisation 2010 XES (Extensible Event Stream) als Beschreibung übernommen und im letzten Jahr als Standard verabschiedet. (<http://www.xes-standard.org/>). Dieser wird von allen gängigen Tools unterstützt und erlaubt eine freie Erweiterbarkeit auf Basis eigener Extensions um Spezifikas des Unternehmens und/oder Prozesses auch aus der Applikation an das Process Mining Tool zu übergeben. Darüber hinaus unterstützen die meisten Tools aber auch den Import via CSV-Dateien oder auch den direkten Zugriff auf Datenbank-Objekte.

Für die großen Systeme mit konstanten Datenmodellen sind bereits vorkonfigurierte Exportstrecken erstellt, die z.B. für einen klassischen Beschaffungsprozess alle notwendigen Eventstream Informationen aus der Applikation extrahieren. Hier gibt es für SAP, Microsoft Dynamics, Oracle E-Business Suite bereits für eine Reihe von Lösungen.

Jedoch gibt es keine simplen Exportstrecken für Jedermanns-Software.

### **Wie könnte das allerdings erreicht werden?**

Da in den meisten Fällen die Individual-Lösung nicht zwingend eine Workflow-Engine im Hintergrund benutzt, die mit wenigen Erweiterungen dazu genötigt werden kann, hilfreiche Audit-Daten auszugeben, bleibt vielfach nur der Blick auf die Veränderungen im Datenmodell. Die sind

allerdings, wie bereits oben argumentiert, mit viel Zeitverlust durch Interviews, Research und Vermutungen verbunden.

Eine Chance dem zu entgehen, ist es die Applikationssoftware mehr in die Verantwortung zu nehmen. Sollte es gelingen, den Anwender bei der Abarbeitung des Prozesses durch die Software „zu beobachten“ und diese Informationen feingranular in einen Eventstream zu schreiben, wäre die Brücke zwischen Prozess und Daten sehr schnell genommen.

Definieren wir einen Prozess als eine Abfolge von Interaktionen des Anwenders mit der Software, die einen Rahmen z.B. innerhalb einer Transaktion haben. Alternativ kann der Prozess durch eine Initialaktion eingeleitet werden, z.B. ein bestimmter Button in der Applikation, die Wahl eines Menüeintrags, etc. und durch eine passende, gleichwertige Aktion beendet werden.

Sollte es also nun gelingen, diese Aktionen in den Eventstream mit dem verbundenen primären Applikationsobjekt (eindeutige Bestellnummer, o.ä.) einzutragen, hat jedes Process Mining Tool im Anschluss ein leichtes Spiel. Idealerweise wird der Eventstream nicht direkt innerhalb der Applikation abgebildet, sondern vielmehr die notwendigen Daten zunächst über die üblichen Log-Mechanismen weitergegeben. Das hat den Vorteil, dass das Framework sich bereits um die Aufgaben kümmert, die Dateien zu verwalten, Rotating sicherzustellen und integrative Methoden zum Schreiben dieser Informationen auf unterschiedlichem Loglevel performant ermöglicht. So ist mit minimalen Footprint eine Übergabe möglich. Ein Verändern des Loglevels erlaubt darüber hinaus, ein Process Mining gezielt ein- oder auszuschalten.

Betrachten wir in der Welt von Oracle z.B. Applikationen auf dem Weblogic-Server, so bietet dieser bereits mit dem Oracle Diagnostic Logging umfangreiche Möglichkeiten an, Informationen aus unterschiedlichen Applikationen über die Administrationskonsole verwalten zu können. Spezielle Logger-Klassen, z.B. in ADF können mit einem konfigurierten ODL-Handler diese Möglichkeiten nutzen. Mit Oracle Forms 12c gelingt es gleichfalls, dass Forms Applikationen mit dem bekannten Message-Built-IN direkt Informationen an das ODL schicken können.

Etwas umfangreicher wird es bei Javascript basierten SPA, die z.B. mit dem JET-Toolkit erstellt wurden. In der Regel reicht es aus, die Prozesse beginnend mit dem Zugriff auf den Service-Layer der Applikation zu identifizieren. Der JavaScript-Part auf dem Client ist primär für die UI Anpassungen verantwortlich, die für einen Prozess vernachlässigt werden können. Mit Javascript auf dem Server, bzw. WebServices kann dann in gewohnter Weise wieder in das Logfile des Application Servers geschrieben werden. Hier ist das primäre Objekt des Prozesses bekannt, allerdings besteht die Gefahr, dass je nach Implementierung, durch die lose Kopplung von User Session und Services, der Context für die Zusammengehörigkeit von einzelnen Prozess-Schritten verloren geht. Hier sind ggfls. Anpassungen an der Software notwendig, die wir auf Grund des Low-Footprint-Ansatzes auf jeden Fall vermeiden wollen.

Die Nutzung von Logging-Funktionalitäten flächendeckend in einer Applikation sind durchweg von dem eingesetzten Framework abgedeckt und hinterlassen keine proprietären Spuren fremder Hersteller. Soll das Event-Stream-Logging nachträglich in eine Applikation eingefügt werden, greifen Vererbungsmechanismen, Parser und Codegeneratoren.

Basiert ihr System auf Java, können Sie mit klassischen Methoden der Vererbung in den Basisklassen Methoden einfügen, die die Eventstream Informationen ausschreiben. In Oracle ADF bieten sich beispielsweise die Framework-Extension-Classes dazu an.

Für JavaScript gibt es reichlich Code-Generatoren, die immer wieder und sogar kontextbezogen Codefragmente erstellen und einbinden können. Basieren dabei die Backend-Services auf Java-WebServices ist ein Java-Logger ohnehin verfügbar.

Selbst für Oracle Forms basierte Applikationen können für jede PL/SQL-Unit nachträglich passende MESSAGE-Built-In Aufrufe mit den relevanten Informationen eingetragen werden. Sollten sie bereits Objektklassen in Oracle Forms nutzen, können sie hier leicht Trigger-Vererbungen einsetzen.

Alternativ können Sie mit dem Forms-API-Master selber strukturiert Message-Builtins dem PL/SQL-Code hinzufügen. Die Firma PITSS GmbH bietet darüber hinaus weitere Möglichkeiten an, diese Informationen auch strukturiert und kontext-abhängig in die Forms- und Reports- und sogar Datenbank-Anwendung einzufügen.

Durch die Nutzung der systematischen Logging-Funktionalität erhöhen Sie automatisch die Qualität der Aufzeichnung gegenüber einem Mining auf dem Datenmodell, da sie zweifelsfrei die Aktion des Benutzers mit dem Prozess und den zugehörigen Daten verknüpfen. Ein Reverse-Engineering um von den Daten an die Ursachen zu kommen, ist nicht mehr notwendig. Ganz nebenbei erzielen Sie dabei auch einen wertvollen zusätzlichen Effekt, der den drei ursprünglichen Methoden verborgen bleibt: Das Fehlen von Einträgen im Eventstream deutet darauf hin, dass diese Prozessschritte, obwohl implementiert und vorhanden, nie benutzt wurden. Auf Grund von Unwissenheit, Fehlprogrammierung oder saisonaler Tätigkeiten? Nun sind Sie in der Lage, mit diesen Fragen zurück in die Fachabteilung zu gehen und über diesen Weg die Prozesse weiter verbessern.

### **Fazit**

Der Vortrag soll Ihnen aufzeigen, dass es nicht länger den großen Applikationen vorbehalten ist, die neuesten Ergebnisse der Forschung für Process Mining nutzen zu können. Werfen Sie einen Blick auf Ihre Applikation und planen Sie die strukturierte Ausgabe prozessrelevanter Informationen.

Viel Spaß und Erfolg bei ihrem ersten Schritt ins Process Mining!

### **Kontaktadresse:**

Stephan La Rocca  
PITSS GmbH  
Otto-Brenner-Str. 209  
D-33604 Bielefeld

Telefon: +49 (0) 521-54679507  
Fax: +49 (0) 521-54679501  
E-Mail: [slarocca@pitss.com](mailto:slarocca@pitss.com)  
Internet: [www.pitss.de](http://www.pitss.de)