

A Process Mining Software Comparison

Daniel Viner, Matthias Stierle, Martin Matzner
Institute of Information Systems, University of Erlangen-Nuremberg
Nuremberg, Germany
{daniel.viner, matthias.stierle, martin.matzner}@fau.de

Abstract—www.processmining-software.com is a dedicated website for process mining software comparison and was developed to give practitioners and researchers an overview of commercial software available on the market. Based on literature review and experimental software testing, a set of criteria was developed in order to assess the tools' functional capabilities in an objective manner. With our publicly accessible website, we intend to increase the transparency of software functionality. Being an academic endeavour, the non-commercial nature of the study ensures a less biased assessment as compared with reports from analyst firms.

I. INTRODUCTION

Starting in the late nineties as an academic research project, the discipline of process mining enjoys an increasing penetration in various industries over the past few years [1]. Process mining helps organisations leverage event log data stored in databases or IT systems with the objective to discover, monitor and enhance processes [2].

The diversity of real-world applications is exemplified by the use of process mining software in banking, manufacturing, online gaming, healthcare, public service and many more industries [3]. Use cases include compliance checks, continuous process improvement (CIP) and the assessment of robotic process automation (RPA) initiatives, to name a few. With the rise of use cases and continuous adoption of process mining in various industries, dozens of commercial tools have emerged on the process mining software market. The changing dynamics of the software market are marked by acquisitions, the increasing number of solutions and continuous releases of new features. Looking forward, the global process analytics market, which includes the discipline of process mining, is expected to grow at a rate of around 50% annually from 2018 to 2023 to reach USD 1.42 billion by 2023 [4].

Depending on the scope and intended scalability, process mining initiatives may require high investments in terms of cost and stakeholder involvement, thus underscoring the danger of selecting the wrong software. Considering the academic context, process mining researchers are often not fully aware of practitioners' needs and the developments in the software market. An overview of available tools and their capabilities is essential to address these issues. While several analyst firms such as Gartner [5] published market studies that deliver an overview of the software landscape, we provide a more detailed analysis of process mining software with tangible criteria that examines functional capabilities. We conducted a non-commercial process mining software analysis and published the results on www.processmining-software.com. Besides serving

as an independent software selection support for practitioners, the website also intends to help researchers understand the state-of-the-art in practice, allowing them to evaluate the usefulness of their work in regards to practical utility.

Based on literature review and experimental software testing, a set of criteria was derived in order to compare the features and functional capabilities of process mining software. This paper describes the underlying methodology and presents nine criteria categories with a brief description for each category and criterion.

II. METHODOLOGY

A. Software Selection

In order to ensure a comprehensive and representative listing, the most recent process mining-related reports of three analyst firms were taken as a basis to identify relevant software. Reports from Gartner [5], Everest Group [6], [7] and Forrester [8] were analysed accordingly. Taking into consideration all software vendors stated in the commercial reports, a list of 34 potential tools was derived and further refined in three steps. First, three vendors not granting access to a demo environment were excluded from the study as we did not want to rely on information provided by the vendors. As the reports do not exclusively cover process mining software but also the software of related disciplines such as task mining or documentation, the respective tools were identified and excluded from the analysis in the second step, reducing the number of relevant vendors to 19. Third, three open-source tools were neglected. The ProM framework offers a comprehensible library of scientific techniques and algorithms but is geared towards academic scholars. PM4Py is an open-source Python library that currently does not provide a graphical user interface, making the solution difficult to use in the organisational context. Apromore was not tested either, however the tool will be considered in the second testing cycle due to the availability of a commercial license. Finally, 16 tools were tested, see Table I.

B. Evaluation Criteria

In order to create a list of relevant criteria, a two-sided approach was followed. First, a literature review was undertaken to identify potential criteria from previous studies. The following terms were searched in the academic search engine sites WorldCat, SpringerLink and Google Scholar: “process mining software”, “process mining software comparison”, “process mining tools”, “process mining tool comparison” and “process mining criteria”. Second, the software was experimentally tested

upfront to better understand what features and capabilities the vendors offer. Vendors were asked to grant access to all features in the demo environment to ensure all available features can be explored. The experimental approach also included the screening of all available knowledge bases and product documentations made accessible by the vendor. The derived criteria set was applied in three steps. In Phase 1, a test scenario was conducted for every tool using the same logs and files. In Phase 2, the results were compared with each other to identify inconsistent terminology and discrepancy in the level of detail. The final assessment was conducted in Phase 3. After testing, follow-up workshops were conducted with every vendor to clarify open questions and to get additional context for features. The exchange with the vendors also served as a quality gate for the correctness of the test results.

C. Testing Setup

The software testing was conducted primarily using event logs of Purchase-to-Pay (P2P) processes with their respective “happy path” reference models in BPMN format.

III. SOFTWARE ANALYSIS

A. Analysed Software

In the course of the study, 16 tools capable of mining event log files were analysed, see Table I. The study was carried out in spring 2020.

TABLE I
COMMERCIAL PROCESS MINING SOFTWARE

Tool Name (Vendor)	
ABBYY Timeline (ABBYY)	MEHRWERK ProcessMining (Mehrerwerk GmbH)
ARIS Process Mining (Software AG)	Minit (Minit j.s.a.)
BusinessOptix (BusinessOptix)	myInvenio (myInvenio Srl)
Celonis Process Mining (Celonis SE)	PAFnow (Process Analytics Factory GmbH)
Disco (Fluxicon BV)	ProDiscovery (Puzzle Data Co., Ltd.)
EverFlow (EverFlow)	QPR ProcessAnalyzer (QPR Software Plc)
LANA Process Mining (Lana Labs GmbH)	Signavio Process Intelligence (Signavio GmbH)
Logpickr Process Explorer 360 (Logpickr)	UiPath Process Mining (UiPath Inc.)

B. Website

The website is mainly built on three layers. While the homepage (first layer) introduces the discipline of process mining, typical use cases and our criteria overview, the “Tools” page (second layer) lists brief profiles of all tools which are linked to the detailed tool profile pages (third layer). An introductory paragraph briefly describes the vendor and the strengths of its software. Eight criteria categories examine the availability and extent of tested functionality while one

criteria category provides general information. The “Distinctive Focus and Features” section provides additional context by highlighting outstanding functionality. In order to offer users visual impressions of a tool, every profile is enriched with a “featured video” provided by the vendor and up to seven screenshots, of which five are defined and two undefined (proprietary). Also, any two selected tool profiles can be contrasted with each other through a side-by-side comparison.

C. Software Criteria

The software criteria derived from the literature review and experimental software testing represents the core of this study. The criteria were grouped into nine categories depicted in Tables II - X in the appendix.

Category *General* gives a brief overview of the vendor and key aspects of the tool. *Data Management* examines functionalities and factors related to the extraction, transformation and loading (ETL) of process data into the process mining tool. The *Process Discovery* category examines process graph capabilities and process analysis features such as benchmarking and rework analysis. *Conformance Checking* is a fundamental process mining feature to identify deviations between the actual “as is” process and an “a-priori” reference model. This category considers all relevant factors pertaining to conformance checking. The *Operational Support* criteria examine the availability of forward-looking capabilities to help users anticipate the outcome of running cases and facilitate decision making with the help of intelligent recommendations. *Views, Monitoring and Reporting* addresses the ability to monitor processes with the help of metrics and visualisations to support decision making. Additional criteria examine available languages and means of collaboration to share insights with other users. While process enhancement functionality such as performance metrics in the process graph is partly covered in the aforementioned criteria categories, *Advanced Enhancement Capabilities* investigates further capabilities that add a new perspective to the graph or the overall process. Lastly, *Security & Compliance* addresses role-based access control and the availability of audit logs.

IV. CONTRIBUTIONS, LIMITATIONS AND OUTLOOK

The study of 16 process mining solutions with commercial licenses showed that the maturity level of the investigated software is highly varying. While some vendors offer basic discovery functionality without conformance checking in some cases, other vendors offer more elaborate features such as process simulation, predictive analytics and decision rule mining. We observe a potential trend: The boundaries between mere process mining functionality and other disciplines such as process modelling (BPMN), business intelligence and Machine Learning become more and more blurred.

The software selection is based on software listed in commercial reports and hence reflects a non-exhaustive picture of the market. Further, open-source software was not analysed. It is important to note that the software listing represents only a snapshot of the tools’ capabilities and features in terms of information timeliness. Vendors are continuously improving

their products and extend the functionalities with periodic releases.

A follow-up study could examine the perspective of organisations on the relevance of the suggested criteria. Interviews may be conducted with organisations interested in process mining as well as organisations with already implemented process mining software.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to all vendors that participated in this study for their time and effort.

REFERENCES

- [1] W. van der Aalst, "Academic View: Development of the Process Mining Discipline," in *Process Mining in Action. Principles, Use Cases and Outlook*, L. Reinkemeyer, Ed. Cham: Springer, 2020, pp. 181–196.
- [2] W. van der Aalst, *Process Mining: Data Science in Action*, 2nd ed. Berlin: Springer, 2016.
- [3] S. Sortino, M. Dell'Orto, G. Cotroneo, R. Carbone and P. Riccardi, "Process Mining: A Database of Applications," 2020, HSPI Report. [Online]. Available: https://www.hspi.it/wp-content/uploads/2020/01/HSPI_Process_Mining_Database_2020.pdf
- [4] MarketsandMarkets Research, "Process Analytics Market - Global Forecast to 2023," 2018. [Online]. Available: <https://www.marketsandmarkets.com/Market-Reports/process-analytics-market-254139591.html>
- [5] M. Kerremans, "Market Guide for Process Mining," 2019, Gartner Report. [Online]. Available: <https://www.gartner.com/en/documents/3939836/market-guide-for-process-mining>
- [6] S. Burnett, A. Vijayan and H. Makan, "Process Mining – The New Juggernaut Driving Digital Transformation," 2019, Everest Group Report, Research Note EGR-2019-38-V-3085.
- [7] S. Burnett, A. Modi, H. Makan and U. Shahdeo, "Process Mining – Technology Vendor Landscape with Products PEAK Matrix® Assessment 2020," 2020, Everest Group Report. [Online]. Available: <https://www2.everestgrp.com/reportaction/EGR-2020-38-R-3576/Marketing>
- [8] R. Koplowitz, "Now Tech: Process Mining And Documentation, Q1 2020. Forrester's Overview Of 21 Process Mining And Documentation Providers," 2020, Forrester Report. [Online]. Available: <https://www.forrester.com/report/Now+Tech+Process+Mining+And+Documentation+Q1+2020/-/E-RES158955>
- [9] A. Rozinat and W. van der Aalst, "Decision Mining in ProM," in *Business Process Management*, S. Dustdar, J. L. Fiadeiro and A. Sheth, Eds., vol. 4102. Berlin, Heidelberg: Springer, 2006, pp. 420–425.

APPENDIX

TABLE II
GENERAL INFORMATION

Criterion	Brief description
Company Size	1-10, 11-50, 51-100, 101-250, 251-500, 501-1000, 1001-5000, 5000+ employees
Free Trial	"Immediate access" or "Upon request"
Licenses	List of all available licenses types, e.g. Academic, Commercial
Deployment	List of all available deployment options, e.g. On-Premises
Embedded In	If applicable: Name of external system/platform that the software is embedded in, e.g. Qlik Sense
Tested Version	Version/build number and month/year of testing

TABLE III
CRITERIA CATEGORY "DATA MANAGEMENT"

Criterion	Brief description
Import File Types	Supported file types for event log upload, e.g. CSV, XES
Database Connections	Available connectors to source data from databases, e.g. ODBC or JDBC drivers
Adapters/Connectors	Available connectors to source data from IT systems (e.g. ERP, CRM) or via APIs
Integrated ETL Functionality	Yes/No - User can extract data from source system, perform ≥ 5 different transformation operations, and finally load data into the software
Data Pseudonymisation	Yes/No - Selected set of data can be pseudonymised, i.e. replaced with hash values
Data Loading	Data Refresh ✓X, Scheduled Jobs ✓X (User can append a new data set A to an existing data set B, i.e. incremental data loading; a time interval or specific dates for data extraction from a specified source can be configured)
Character Encodings	Test of UTF-8 compatibility incl. special characters from various non-Latin languages; List of all additional supported character encodings
Attribute Types	Case-level ✓X, Event-level ✓X
Specify Business Hours	Working week ✓X, Multiple shifts/day ✓X, Exclude days ✓X, Holiday calendar ✓X
Define Event Order	List of manual means to order events in case of identical timestamps, e.g. by selected column that contains sorting information
Start/End Timestamp	"1 timestamp" or "2 timestamps"

TABLE IV
CRITERIA CATEGORY "PROCESS DISCOVERY" (1/2)

Criterion	Brief description
As-Is Process Visualisation	List of all available visualisation types for the process graph, e.g. Directly-Follows Graph
Export As-Is Process Graph	Available export formats for the process graph
Performance Highlighting	Active time ✓X, Idle time ✓X (Visual bottleneck highlighting of activities, i.e. active time, and transitions, i.e. idle time)
Process Animation (Replay)	Adjust speed ✓X, Adjust timeframe ✓X, Switch time mode ✓X, Zoom in case ✓X (Animated replay of all process flows from a case perspective)
Search & Filter in Graph	Search ✓X, Filter ✓X (User can search for any activity in the process graph; User can filter by activities/transitions (nodes/arcs) directly from the process graph)
Graph Abstraction	Yes/No - Amount of displayed nodes and arcs in the process graph can be varied/adjusted
Frequency Metrics	List of all available frequency-related metrics
Time Metrics	List of all available performance-related metrics
Additional Graph Metrics	Cost metrics ✓X, Custom metrics ✓X

TABLE V
CRITERIA CATEGORY “PROCESS DISCOVERY” (2/2)

Criterion	Brief description
Process Benchmarking	Visual comparison ✓X, Metric comparison ✓X (2 filtered sets of the same process can be compared with each other visually and metrically)
Process Benchmarking (Different Logs)	Visual comparison ✓X, Metric comparison ✓X (Processes of ≥ 2 different event logs can be compared with each other visually and metrically)
Root Cause Analysis	Yes/No - The software delivers a list of root causes for selected or defined anomalies/symptoms
Variant Breakdown by	List of metrics by which the variants can be classified/sorted
Case and Activity List	Activity List ✓X, Case List ✓X, Case List for Variants ✓X
View Case Details	Yes/No - User can access a case view with respective case activities and metrics
Rework Analysis	Yes/No - User can identify rework, i.e. loops and self-loops, through pre-configured dashboards or filtering
Edge/Transition Details	List of all transitions ✓X, From-to activities ✓X

TABLE VI
CRITERIA CATEGORY “CONFORMANCE CHECKING”

Criterion	Brief description
Compare As-Is and Target Process	Yes/No - User can compare as-is process with a target process, e.g. happy path
Target Model Creation	Import model ✓X (<model types>), Auto-create from as-is ✓X, Create new ✓X
In-Graph Conformance Visualisation	Yes/No - Deviations from a target process can be visualised in the process graph
List of Conformance Violations	Yes/No - List of identified conformance violations for undesired activities, missing activities and non-compliant activity sequence
Four-Eyes Principle	Yes/No - Breach of the four-eyes principle can be detected for any 2 selected activities
Sequence Filtering	Yes/No - User can filter by the condition “activity A is (not) directly followed by activity B”
Conformance Root Cause Analysis	Yes/No - Root causes can be automatically identified for selected conformance violations

TABLE VII
CRITERIA CATEGORY “OPERATIONAL SUPPORT”

Criterion	Brief description
Alert Generation	Yes/No - Capability to trigger alerts defined by the user via query/filter, KPI threshold or a particular time interval
Predictive Analytics	Yes/No - Capability to predict the future outcome of a running case based on historic data [2]
Recommendations (Prescriptive Analytics)	Yes/No - Capability to suggest potential next actions in order to meet a particular business goal, e.g. minimising cycle time [2]

TABLE VIII
CRITERIA CATEGORY “VIEWS, MONITORING AND REPORTING”

Criterion	Brief description
Export Reports	Events (<formats>), Cases (<formats>), Variants (<formats>)
Export Charts and Tables	Yes/No
Custom Dashboards	Custom charts ✓X, Custom tables ✓X
Custom Metrics/KPIs	Yes/No - User can define custom metric/KPI through a formula using own syntax, or by selection of any imported numerical attribute with the option of at least 5 different aggregation types, e.g. mean, median and percentiles
KPI Thresholds	Yes/No - User can define thresholds for metrics/KPIs or charts to emphasise (non-)acceptable values by colour highlighting
Advanced Charts	Yes/No - User can choose from at least 5 different chart types
World Map	Latitude & longitude coordinates ✓X, Location by attribute (e.g. country codes, city names) ✓X (Visualisation of process-related locations in a world map graph)
Save Filter Settings	Yes/No - Applied filter settings can be reused at a later point in time
UI Languages	List of all available languages in the GUI
Share and Collaborate	Share selection ✓X, <collaboration features> (Sharing applied filter settings with other users; List of all additional means to collaborate and share insights, e.g. comment feature)

TABLE IX
CRITERIA CATEGORY “ADVANCED ENHANCEMENT CAPABILITIES”

Criterion	Brief description
Organisational Mining	Yes/No - Capability to visually add organisational perspective by grouping of activities and org. entities such as resources and departments [2]
Scenario Simulation	Yes/No - The impact of specific process alternations (e.g. adjusting resource allocation and work times for activities) on the overall process can be simulated
Decision Rule Mining	Yes/No - Automatic derivation of rules for decision points based on case-related data such as case-level attributes [9]

TABLE X
CRITERIA CATEGORY “SECURITY & COMPLIANCE”

Criterion	Brief description
Role-Based Access	Yes/No - Access to projects, dashboards or certain process data can be restricted for any user in the system via user roles or user-specific access permissions
User Authentication	List of all means of authentication for user login, e.g. 2FA, LDAP, Active Directory
Audit Logs	Yes/No - Audit logs can be produced which contain data of at least user identification, executed activity and corresponding timestamp