Innovation Insight for Process Mining and Task Mining

Published 10 August 2021 - ID G00755303 - 19 min read

By Analyst(s): Marc Kerremans, Tushar Srivastava

Initiatives: Technology Innovation

Confusion surrounds technology that enables the discovery of processes, operations and interactions. To improve the success of mining initiatives, enterprise architecture and technology innovation leaders should distinguish between process and task mining — different but complementary disciplines.

Additional Perspectives

- Summary Translation: Innovation Insight for Process Mining and Task Mining (10 September2021)

Overview

Key Findings

- Different mining techniques and tools are assessed by client organizations for process discovery. Driven by the language of, and tools marketed by, vendors, these organizations often fail to recognize the underlying use cases for each of these mining techniques and end up with unsuccessful initiatives.
- All mining techniques and related technologies have their strengths and weaknesses. When applied as complementary disciplines, they provide very accurate insights into the operations of any organization.
- Because of the emerging nature of some of the use cases and underlying disciplines, most vendors offer capabilities related to only one mining technique. In order to support a wider set of use cases, some vendors started offering more than one mining technique.

Recommendations

To create continuous visibility and better insights into business operations, and to drive adoption, enterprise architecture (EA) and technology innovation leaders should:

- Select and implement process discovery and analysis initiatives by clarifying the differences between the process and task level. This will better match the right mining technique and technology to the expected scope and outcome of the defined use case.

- Combine process and task mining techniques as complementary approaches. Combining them will create more accurate visibility and understanding, driving the success of the mining at both the process and task level.

- Improve long-term sustainability by embedding these techniques and related technologies in a competence center. Move beyond the technology level to ensure business governance over opportunity discovery, architectural choices and effective change management.

Introduction

These days a lot of confusion arises around tools and technology that enable the discovery of processes, operations and interactions. Depending upon the scope, use case and goals, it is imperative to distinguish between these different but complementary disciplines, techniques and related tools. Based on the notion of a task and a process, we will explore the two major categories of mining — process mining and task mining. We are not redefining "task" and "process" in general, but will describe how we differentiate these notions within the context of this mining research (see Figure 1).

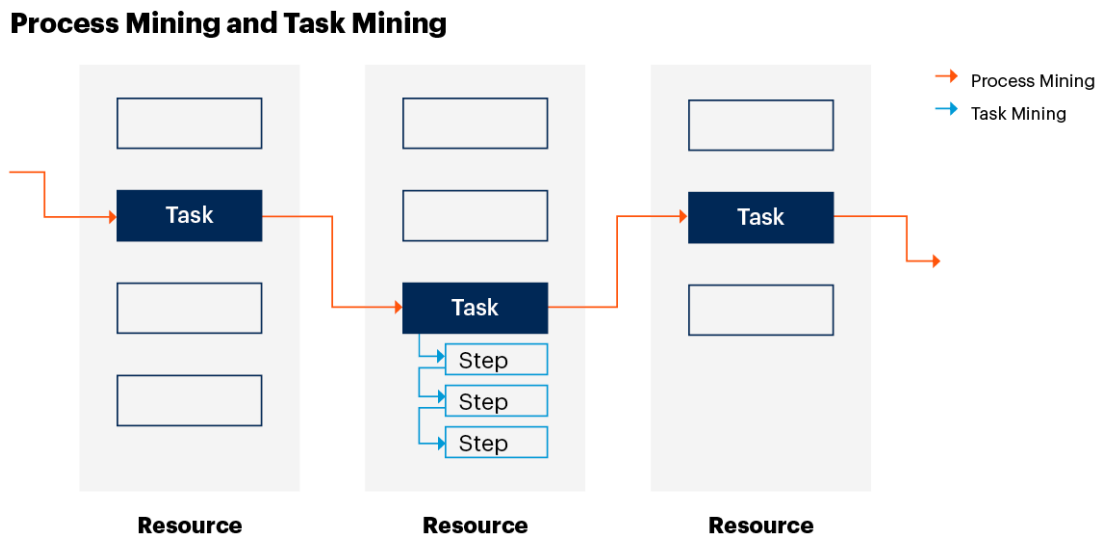
Process

Processes are end-to-end groupings of logically related work, such as order to cash (OTC). Most of these end-to-end processes can be divided into subprocesses that chunk this logically related work into functions. In the case of OTC, subprocesses may include sales or accounts receivable. Each of these processes contain a number of tasks.

Task

Tasks are the groupings of work within a process. For example, activities in the accounts receivable subprocess may include recording sales, generating invoices and tracking customer balances. Each of these tasks contains a number of steps that are detailed instructions to perform the work. Under the "check customer credit" task, examples of tasks include opening the customer credit screen, reviewing any outstanding balance, adding this to an Excel spreadsheet, subtracting that from the credit line and reviewing outstanding orders.

Figure 1. Process Mining and Task Mining



Source: Gartner
723821_C

Gartner

Description

Process Mining

Process mining is designed to discover, monitor and improve actual processes (that is, not assumed processes) by extracting knowledge from event logs readily available in today's information systems.

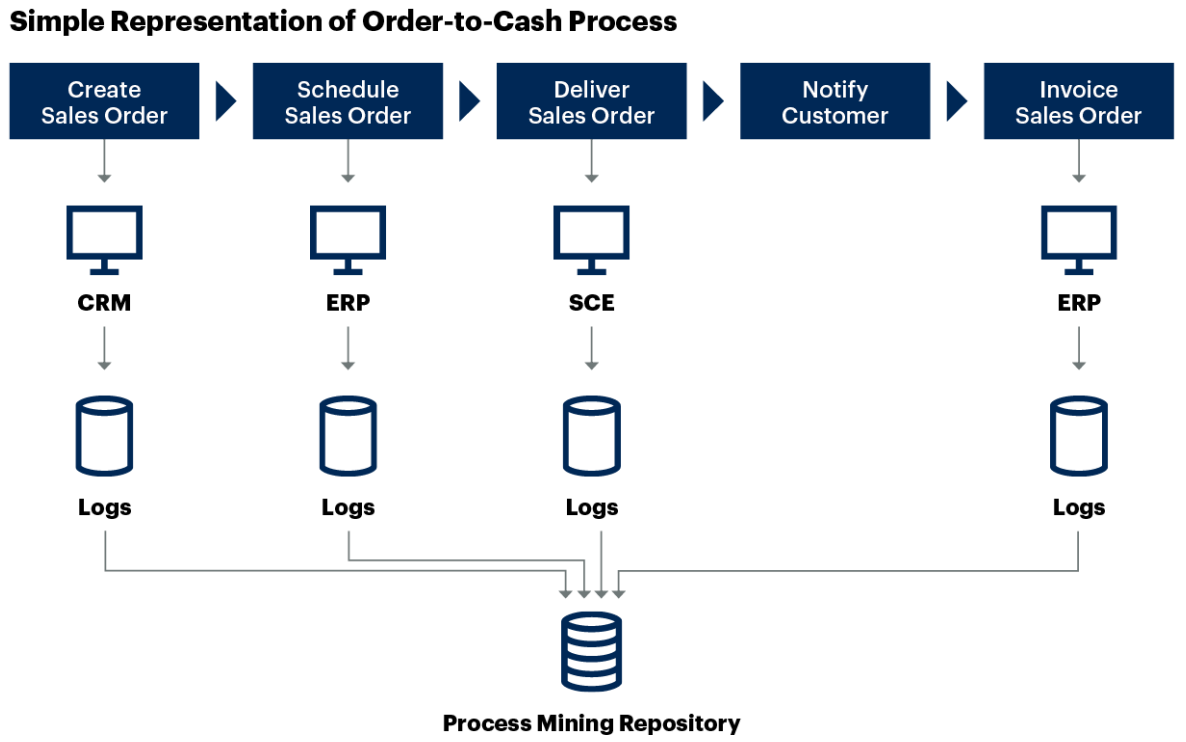
Process mining includes:

- Automated process discovery (extracting process models from an event log)
- Conformance checking (monitoring deviations by comparing model and log)
- Social network/organizational mining (how people interact and work together)
- Automated construction of simulation models
- Model extension
- Model repair
- Case prediction
- History-based recommendations [1,2]

See Market Guide for Process Mining.

The starting point for any process mining initiative is an event log. An important observation considering event logs, is that these events need not be present in a log file from an enterprise application. The advantage of having this application log file available is that the events are already structured. Events can come from anywhere (desktop applications, Internet of Things [IoT] devices, documents) and, in that case, need to be structured to be ready for ingestion by the process mining engine. Each event found in a log refers to an activity (that is, a well-defined step in some process) and is related to a particular case (that is, a process instance). The events belonging to a specific case are ordered and can be seen as one run of the process (see Figure 2).

Figure 2. Simple Representation of an Order-to-Cash Process



Source: Gartner

723821_C

Gartner

The tasks in Figure 2 (create sales order, schedule sales order, etc.) represent the “happy path” of the process, when there is no delay and no back and forth between tasks. The real issue is that this is generally not the case, and there are a lot of business outcomes that take a hit due to delays, assumptions and not adhering to the process.

Process mining is a technique that marries the concept of data science with process science to discover, monitor and improve actual processes (not assumed processes) by extracting knowledge from event logs that are available in the different systems.

Each task in the business process in Figure 2 has an event log associated with it. In the example, there are entries in an ERP system, a CRM system and a supply chain execution system. The notification of the client is a manual task, and if this task shows a lot of variance in execution, it should be investigated.

Once these event logs are put through the process mining algorithms, one can easily have a look at exactly what happened during the execution of the process. The possibilities include discovering a hidden aspect of the process, a bottleneck, conformance checking, or even finding opportunities to automate and enhance the processes.

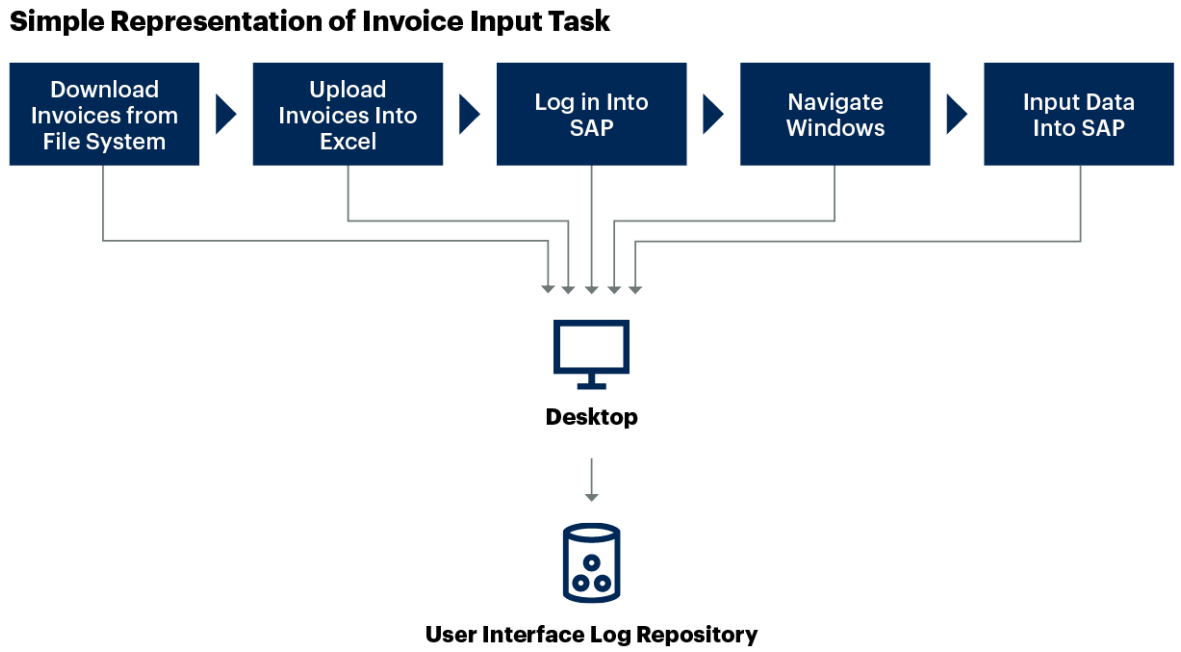
Task Mining

Task mining is a complementary approach to process mining that infers useful information from low-level event data available in UI logs or captured through computer vision. These UI logs describe the single steps done by a user (for example, in using a workstation) based on keystrokes, mouse clicks and data entries. Additional mining capabilities interpret the data by applying natural language processing (NLP) and optical character recognition (OCR) to correlate data in different ways.

The starting point for any task mining initiative is a user interface log. Each event found in a log refers to a step that is related to a particular task. The events belonging to a task are ordered and can be seen as one sequence or routine of a task (see Figure 3).

Consider a scenario where a user is responsible for pulling in invoices from an external system and navigates through multiple applications and screens to accomplish the task.

Figure 3. Simple Representation of an Invoice Introduction Task



Source: Gartner

723821_C

Gartner

The steps that the user takes are:

1. Download invoices from the directory of incoming invoices.
2. Upload the details of the invoices into an Excel sheet.
3. Log into the SAP system.
4. Navigate through three different screens on the SAP system.
5. Introduce the invoice data into SAP.

All of these steps are completed by an employee at his or her desktop. Task mining will analyze the data collected during the execution of this task, collect variants, and try to discover hidden aspects of the task and opportunities to enhance the task.

Table 1 details the differences between process mining and task mining.

Table 1: Differences Between Process Mining and Task Mining

(Enlarged table in Appendix)

	Process Mining	*Task Mining*
Scope	End-to-end process	Completion of single task or a set of single tasks on a desktop
Granularity, abstraction	Tasks	Steps and actions within a task
Sequencing	Control flow	Control flow plus data flow
Instance	Unique identifier of instance can be found in each tasks	No unique identifier can be found in each steps of a task
Detailed data	Not all info in log (e.g., address)	All I/O data attributes
Attributes of events	Incomplete	Complete
Discovery of events	Byproduct of transactions	Explicitly recorded for analysis
Underlying techniques	Algorithms	Algorithms, machine learning (ML), rules mining, computer vision

Source: Gartner (August 2021)

To summarize the main differences between process mining and task mining: [3]

- Subject matter:
  - Process mining and task mining have very different scopes. The former targets end-to-end business processes (such as procure to pay and quote to cash) or parts of these processes, where different resources work together to deliver the process outcome — a product, service or information. In this case, the most granular unit is a task. Process mining is mainly interested in the order in which individual tasks are executed or evaluated, usually called the control flow.
  - On the contrary, task mining targets a task consisting of different actions or steps. The granularity is a mouse click, keystroke, an entry of data or a desktop operation, such as copy/paste. Task mining is not only interested in the control flow (that is, the sequence of the steps), but also in the data flow.

- Data:
  - The starting point for any process mining task is an event log. Each event in such a log refers to an activity (that is, a well-defined step in some process) and is related to a particular case (that is, a process instance). For each of these single events, the time stamps captured are used for the sequencing, so basically you need an ID of the case, a task description and a time stamp. Other data (such as amount, area and client) is contextual data and can be used to do segmentations and refined analysis.
  - Task mining also needs step information and time stamps for the sequencing, but it does not have a unique identifier to bundle steps together. In order to provide the control flow, the first stage is to do a segmentation on all the captured information on the desktop or device to define repetitive routines (i.e., sequences of steps). Furthermore, once the control flow has been derived, it also has to look at the data introduced and how this data is composed. In order to do this, task mining requires more complete data than process mining, especially in what is called the “process discovery” or “robotic process mining (RPM)” use case (covered in the Use Cases for Process Mining and Use Cases for Task Mining sections).

- Technique:
  - Finally, the techniques in both mining categories are somewhat different. Process mining is a special case of data mining. Unlike data mining, process mining focuses on the process perspective. It includes the temporal aspect and looks at a single process execution as a sequence of activities that have been performed. Process mining bridges the gap between data mining and the modeling, control, and improvement of business processes. [4]
  - The underlying technology powering task mining is a combination of some process mining techniques, but it also needs some specific algorithms for the segmentation of cases and pattern recognition for the data transformation. Additionally, it may be complemented by using NLP to model topics and to identify and enrich key data (such as a username, keywords, labels, customer name and order IDs). In some use cases, such as the use case for robotic process automation (RPA), computer vision and screen capturing are used to create or enhance UI logs.

Benefits and Uses

Use Cases for Process Mining

Process mining can be applied in different areas of an organization. Therefore, it can have many use cases, and a multitude of stakeholders may depend on the use case and the scope of the process mining initiative.

Improving Processes by Algorithmic Process Discovery and Analysis

Process mining provides visibility and understanding of actual business operations and processes by applying a set of algorithms to events, resulting in highly adaptable, highly maintainable and validated process models. Besides identifying process inefficiencies, this technique delivers insight into where to improve operations (for example, in a digitalization initiative) and how to attain targeted business outcomes. By supporting process efficiency and effectiveness, process mining tools are key enablers of process improvement initiatives and their related disciplines. In this use case, data scientists typically support process improvement teams to discover, analyze, improve and optimize processes. They do this by using methodologies (such as Kaizen; lean; plan, do, check, act [PDCA]; total quality management [TQM], or define, measure, analyze, improve and control [DMAIC] and Six Sigma).

Improving Auditing and Compliance by Algorithmic Process Comparison, Analysis and Validation

Most organizations have internally defined standard operating procedures, policies, work instructions or best practices baked into enterprise applications, such as ERP and CRM. In practice, many of these predefined operations are often complemented by shadow operations. Process mining helps validate or audit whether actual operations are in conformance with these defined operations. More accurate assessments of process deviations and compliance issues, such as segregation of duties, help manage these risks and communicate findings in an audit report. This also helps ensure the reliability of external financial reporting and compliance with external laws and regulations. In these cases, the stakeholders of process mining can be found in financial management and in risk and compliance teams.

Another variation of this use case is the comparison and analysis of the same processes across different business units, subsidiaries or even countries (i.e., comparative process mining).

Improving Process Automation by Discovering and Validating Automation Opportunities

Process automation initiatives have been delivering highly improved efficiencies, but they have resulted in limited sustainable business value because of the dilemma presented by the assumption that the actual data is available only after automation has been completed. Process mining delivers the actual operational data before turning to automation and supports making a more accurate business case on the improvement of business outcomes. Process mining will help business analysts discover and assess the opportunities for process automation. Recently, task-level process automation, such as RPA, has often missed the process context and has not provided awareness of the changes that occur to the processes in which these tasks are embedded. Moreover, process models are often used to configure enterprise applications and their integrated processes, such as ERP and CRM. Finally process mining can be regarded as an essential building block for hyperautomation.

Supporting Digital Transformation by Linking Strategy to Operations

If we shift the scope of process mining to the operational and organizational level, it is a small step to link these operational insights to the big strategic initiatives, such as digital transformation or digital business. This is still a fairly new use case that is getting more attention because of the theme of digital business and, more recently, digital transformation. It provides visibility, analysis and understanding around business operating models that represent a way of doing business by providing near-real-time information to all end users on how they are performing and what could be improved. Furthermore, it supports finding opportunities for improvement. In some cases, it has been reported that it even helps end users think about digitalization opportunities.

Improving IT Operations Resource Optimization by Algorithmic IT Process Discovery and Analysis

This use case is not a direct goal or result of business-outcome-driven process improvements. However, the discovery, analysis and optimization of IT operations and processes is a new use case that can certainly deliver a lot of benefit to IT organizations and widely adopted practices, such as DevOps. Consider its use in the improvement of development and testing processes, preparation of migration processes, system error diagnostics, technology usage and postimplementation/transformation stabilization. This stabilization centers on usability issues, client service delays and reproduction scenarios.

Use Cases for Task Mining

Task mining also has different use cases, and a multitude of stakeholders may depend on the use case and the scope of the task mining initiative.

Improving Tasks by Task Discovery and Analysis

Task mining provides visibility and understanding on the actual steps of a task, switching between these steps, as well as switching between tasks and activities of personal interest. Besides identifying task inefficiencies, this technique delivers insight into how to improve tasks (for example, in a digitalization initiative) and how to attain targeted outcomes. By supporting task efficiency and effectiveness, task mining tools are key enablers of task improvement initiatives and their related disciplines. In this use case, data scientists typically support improvement teams that are taking care of digital workplace initiatives. Most organizations have defined standard procedures, policies and task instructions. In practice, many of these predefined operations are often complemented by shadow tasks.

Identifying Opportunities for Task-Level Automation — Robotic Process Mining

This is the use case that many RPA-oriented vendors refer to as process discovery.

RPM is a class of techniques and tools to analyze data collected during the execution of user-driven tasks. Its purpose is to support the identification and assessment of candidate routines for automation and the discovery of routine specifications that can be executed by RPA bots.

Given a UI log, RPM tools aim to:

- Identify automatable routines and their boundaries
- Collect variants of each identified routine
- Standardize and streamline the identified variants
- Discover an executable specification corresponding to a streamlined and standardized variant of the routine

The routines produced as the output should be defined in a platform-independent language that can be compiled into a script and executed in an RPA tool. [5]

Improving Workforce Optimization

Recently, workforce performance optimization tools received renewed attention because of the incorporation of AI and machine learning techniques (see Workforce Planning — How to Use Technology to Support Planning Processes).

As such, many organizations are searching and exploring the use of these tools to discover opportunities for task improvement by implementing task mining across different workplaces. However, it should be noted that task automation is limited to the task and that task automation optimizes within this task context.

An example is that monitoring worker A on a certain task can reveal that this worker is underperforming on loan assessment because, two activities upstream, worker B always performs very fast but forgets to introduce some relevant information. Thus, tasks should always be seen in the context of the larger process in which the task belongs.

Risks

Right Tools for the Right Purpose

Process mining and task mining have different scopes and, as such, are well equipped with the necessary capabilities to deliver on their specific use cases. Trying to apply process mining for discovering and analyzing tasks will not perform well in accomplishing the goal because UI logs miss some necessary data (such as a case ID) and contain too much noise and waste. Noise indicates events that capture actions that are irrelevant to the task. Noise may occur, for example, when the user is interrupted or gets distracted during the execution of a task, leading to performing activities that are not relevant to the task in question. [5] Waste refers to events that have no effect on the final outcome of the task. Waste may occur, for example, by typing data in a field, deleting this and typing something different. As such, process mining will not reveal the reason why a certain task can have many variations during execution.

Task mining will discover and analyze events that are performed on a single desktop. Vendors sometimes claim that combining several discovered tasks over several desktops and applying machine learning will discover the end-to-end business process. This underestimates the importance of discovering the business process exceptions that exist, which is one of the outcomes of process mining.

Process Mining and Task Mining — Better Together

Consider the following simplified examples.

Task mining complementing process mining: In the morning, at 8 a.m., I start introducing a new purchase requisition into a transactional system. Ten minutes later my colleague comes to my desk and asks me to complete something I forgot the day before. After finishing this task, I remember that I forgot that my wife asked me to bring her something from the store and I have a short messaging session. Finally, at 9:30 a.m., I complete my purchase requisition transaction. According to the events in the log file of the transactional system, process mining will make the assumption that the task has taken one hour and 30 minutes to complete. Here task mining can indicate the necessary corrections.

Process mining complementing task mining: Half an hour before office hours close I still have to input 20 invoices and I decide to leave out a control field that requires me to check on still open invoices in an aging list. I know the system will accept these invoices. The morning after, another colleague who performs invoice verification detects many invoices that are not complete and works the entire morning in completing these invoices. Learning from these two desktops will replicate and confirm some of this behavior. Here process mining could have indicated that these tasks were so-called shadow processes and tasks.

Governance

In this research we have covered techniques and technology. We do not need to point out the risk of not providing some governance mechanisms. Therefore, for the long-term success of these initiatives, it is imperative to set up a competence center approach that takes into account organization, roles and responsibilities, architecture, metrics, and change management (see Develop 3 Levels of Service for Your Center of Expertise to Scale DigitalOps and Robotic Process Automation).

Recommendations

To create continuous visibility and better insights into business operations and to drive adoption, EA and technology innovation leaders should:

- Select and implement process discovery and analysis initiatives by clarifying the differences between the process and task level. This will better match the right mining technique and technology to the expected scope and outcome of the defined use case.
- Combine process and task mining techniques as complementary approaches. Combining them will create better visibility and understanding, driving the success of the mining at both the process and task level.

- Improve long-term sustainability by embedding these techniques and related technologies in a competence center. Move beyond the technology level to ensure business governance over opportunity discovery, architectural choices and effective change management.

Representative Providers

Process Mining

- Apromore — Apromore
- BusinessOptix — BusinessOptix
- Celonis* — Celonis process mining
- Fluxicon — Disco
- IBM* — myInvenio
- Minit* — Minit
- PAFnow — PAFnow
- QPR Software — ProcessAnalyzer
- SAP** — SAP Process Intelligence by Signavio
- Software AG** — ARIS Process Mining

*Covering both technologies; **Covering both technologies through partnership

Task Mining

- ActiveOps* — Workware+
- EdgeVerve Systems, a wholly owned subsidiary of Infosys — AssistEdge Discover
- FortressIQ — Process Intelligence
- Kryon Systems — Kryon Process Discovery
- Pegasystems — Pega Workforce Intelligence
- Soroco* — Scout Enterprise

- StereoLOGIC* — StereoLOGIC
- UiPath* — UiPath Automation Hub
- UltimateSuite — UltimateSuite
- WorkFusion** — Intelligent Automation Cloud

*Covering both technologies **Covering both technologies through partnership

Evidence

[1] "The Process Mining Manifesto," IEEE Task Force on Process Mining.

[2] W. van der Aalst, "Process Mining, Data Science in Action," Springer Verlag, 2016.

[3] V. Leno, M. Dumas, M. La Rosa, F.M. Maggi and A. Polyvyanyy, "Automated Discovery of Data Transformations for Robotic Process Automation."

[4] Process Mining Book

[5] V. Leno, A. Polyvyanyy, M. Dumas, M. La Rosa and F.M. Maggi. "Robotic Process Mining: Vision and Challenges."

Document Revision History

Discover the Differences and Use Cases of Process Mining Versus Task Mining - 22 April 2020

Recommended by the Authors

Some documents may not be available as part of your current Gartner subscription.

Market Guide for Process Mining

How to Start Up a Business Process Competency Center to Deliver Better Business Outcomes

Develop 3 Levels of Service for Your Center of Expertise to Scale DigitalOps and Robotic Process Automation

Workforce Planning — How to Use Technology to Support Planning Processes

Predicts 2021: Accelerate Results Beyond RPA to Hyperautomation

Table 1: Differences Between Process Mining and Task Mining

↓	*Process Mining* ↓	*Task Mining* ↓
Scope	End-to-end process	Completion of single task or a set of single tasks on a desktop
Granularity, abstraction	Tasks	Steps and actions within a task
Sequencing	Control flow	Control flow plus data flow
Instance	Unique identifier of instance can be found in each tasks	No unique identifier can be found in each steps of a task
Detailed data	Not all info in log (e.g., address)	All I/O data attributes
Attributes of events	Incomplete	Complete
Discovery of events	Byproduct of transactions	Explicitly recorded for analysis
Underlying techniques	Algorithms	Algorithms, machine learning (ML), rules mining, computer vision

Source: Gartner (August 2021)