CHAPTER - 1

INTRODUCTION

Our world and the organizations in it are full of processes. From purchasing to order management, organizations deal with complex, global and sometimes faulty processes on a daily basis. Frictionless processes, on the other hand, ensure:

- that you can find the right groceries at the grocery store,
- that planes land on time,
- that patient waiting times at hospitals are kept to a minimum.

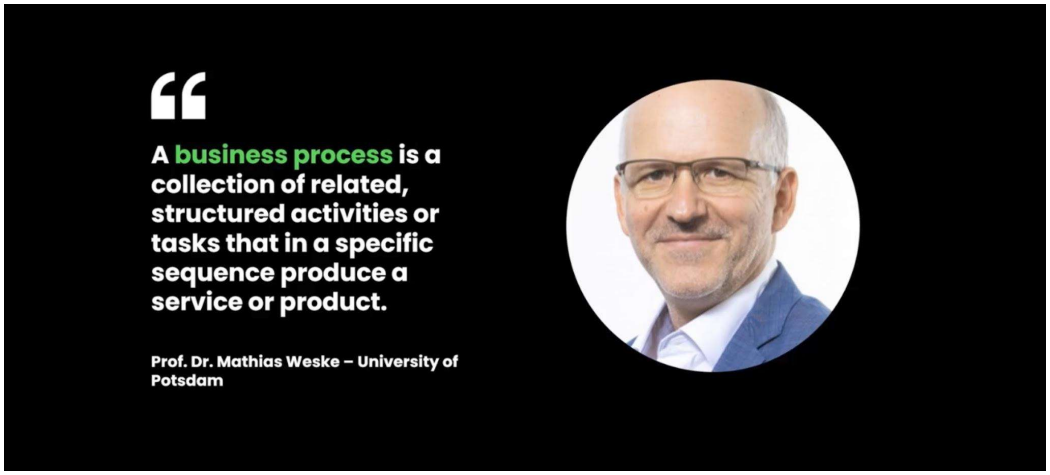


Fig. No. 1.1: Prof. Dr. Mathias Weske about a business process

1.1 The Process Movement

We learned that you can find processes everywhere in daily life and that they're crucial for frictionless operations. Business Process Management has traditionally examined these processes by talking to the people involved. The problem is that, by gathering information from people, you also gather their (false) assumptions and subjective or fractured observations. Take a look at the video below to understand the issues that this approach poses.

As you've seen, traditional approaches fail to understand the real-life complexity of processes and also struggle to provide complete insights and visibility given the vast amounts of data that are now available.

By contrast, Process Mining offers a data-driven and therefore more objective and holistic approach to understanding business processes. As a result, Process Mining has come to dominate a large majority of operational excellence, automation and digitalization ambitions within industry. Process Mining is the leading new technology when it comes to talking about algorithmic businesses - in other words, businesses that use algorithms and large amounts of real-time data to create business value. This has only become possible through the advent of information systems and administrative tools (e.g., Enterprise Resource Planning or Customer Relationship Management systems) which provide a good data source for process analytics. Process Mining is a solution to costly and time-intense efforts to get data-driven insights into a business, as acknowledged by the industry research firm Gartner.

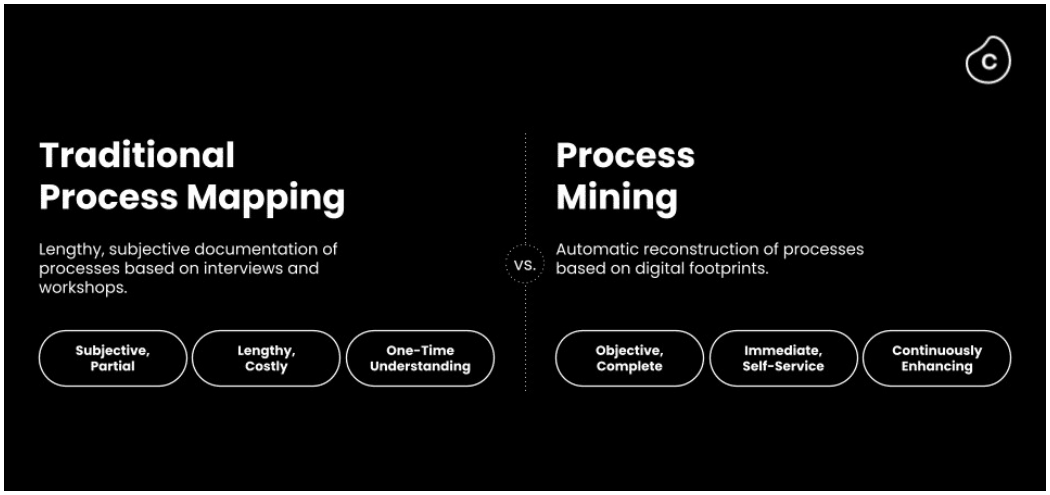


Fig. No. 1.2: Traditional Process Mapping vs. Process Mining

Compared to the traditional process mapping approaches, Process Mining technology solves the complexity and visibility problem. It's an x-ray for businesses that gives 100% transparency into processes, eliminating process blind spots, and quantifying the impact of process problems on core KPIs. It leverages data from a business's source systems and user desktops to map the processes, thus eliminating conjecture about how they're running. This allows businesses better field competing requests from their stakeholders because they can have confidence in the data model.

And lastly, because it's using the system data in real-time, it provides a living, breathing view of the processes that is generated immediately and is always up-to-date, substantially reducing cost and the time to value.

1.2 Early Stages

Process Mining originally emerged from academic research into how event log data retrieved from Information Systems could be used to discover, monitor and improve real processes. This real data can facilitate several aspects of Business Process Management including:

- Process discovery
- Conformance checking
- Organizational mining, i.e., using data to analyze the roles and people involved in a process
- Automation
- Simulation, i.e., foreseeing and testing the outcome of a process depending on the variation of variables
- Prediction
- History-based recommendations

1.3 Process

- A process is very simply a series of linked actions or steps taken in order to achieve a particular end. For customer service, these could be the steps to resolve a ticket.
- For sales, it could be the steps to progress an opportunity from a lead to closure.
- Take order management, for example. This could be the steps from a customer ordering goods, to you shipping, and then ultimately getting paid for them. You get the point!

1.4 Process Mining Enters the Business World

Since then, these ideas and the academic concepts behind Process Mining have bridged the gap to enter the business world. A variety of software vendors have ventured into the market and even expanded its capability from analytics to business

execution. They achieve this through stronger operational links to automation frameworks and IT source systems which allow daily users to receive prompts and take direct action to improve processes.

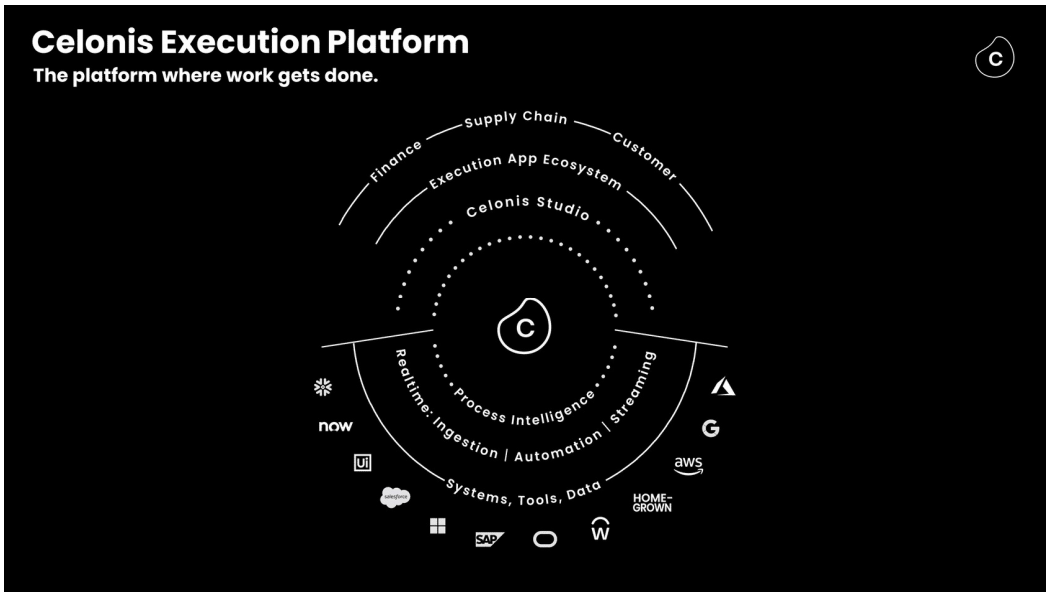


Fig. No. 1.3: Celonis Execution Platform

1.5 Process Mining Market Direction

The direction that the Process Mining market will develop into in the future is not just determined by technological capabilities, but also by development within the market itself. For example, recent mergers and acquisitions between software vendors Celonis and Make, Signavio and SAP or Process Gold and UiPath paint a clear picture: The Process Mining landscape is blending more and more into other intelligent business applications, developing from a pure analytics tool towards an operational tool which supports and improves daily business interactions. The global process mining software market size was valued at USD 1.13 billion in 2022 and is projected to grow from USD 1.66 billion in 2023 to USD 27.72 billion by 2030, exhibiting a CAGR of 49.4% during the forecast period.

Process mining creates an “MRI” of a business process that helps you gain visibility and uncover value opportunities hiding within core operations. Celonis process mining has quickly become the backbone of many companies' efforts to streamline and optimize processes. Celonis holds the highest reported capability scores

in seven out of eight assessment criteria including market adoption, portfolio mix, vision and strategy, process set-up and discovery, process intelligence, implementation and support, and commercial and sales model.

1.6 Data Mining and Process Mining

Data mining is the process of extracting and discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems. For example, sales and marketing teams can use data mining to predict which prospects are likely to become profitable customers. Process Mining is the combination of two disciplines: Data Science and Business Process Management. Process Mining essentially uses Data Science techniques, such as Big Data and AI, to address Process Science problems such as process improvement and automation.

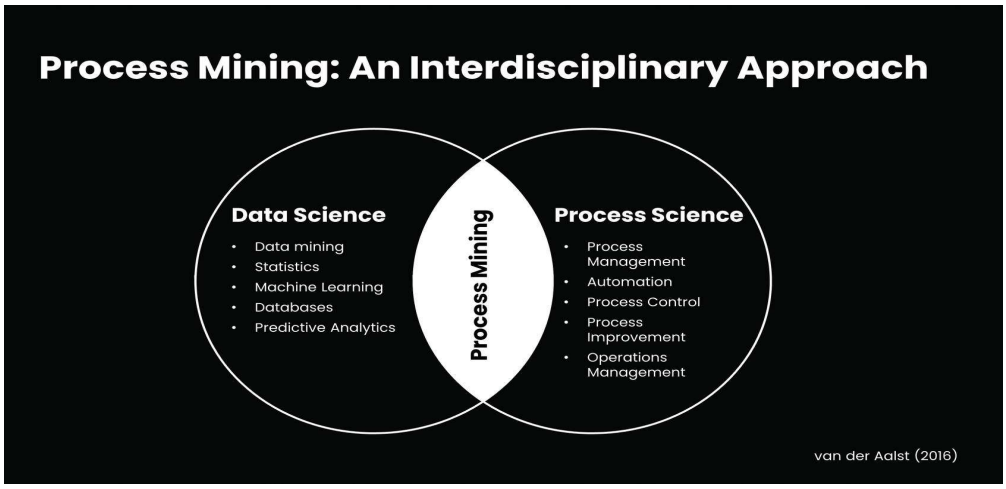


Fig. No. 1.4: Process Mining: An Interdisciplinary Approach

Process Mining achieves this union by taking the digital footprints that are created in IT systems and using them to reconstruct and visualize process flows. From here, Process Mining technology can identify patterns and deviations and ultimately eliminate bottlenecks. Now we will take a deeper look at what is required to reconstruct a process in this way. Digital footprint or digital shadow refers to one's unique set of traceable digital activities, actions, contributions, and communications manifested on the Internet or digital devices. Digital footprints can be classified as either passive or active.

CHAPTER - 2

TECHNOLOGY

Process mining is a set of techniques for the analysis of operational processes based on event logs extracted from company's databases, information systems, or business management software such as enterprise resource planning (ERP), customer relationship management (CRM), electronic health records (EHR), etc. In simple words, it's about finding out how the processes are actually performed to discover problems and areas for improvement. Process mining can also be described as a part of business process management (BPM) that applies data science (with its data mining and machine learning techniques) to dig into the records of the company's software, get the understanding of its process's performance, and support optimization activities.

There are different perspectives that process mining can be focused on.

1. The control-flow perspective is about the order of activities, and the goal here is to find the optimal path of performing a process.
2. The organizational perspective targets the resources involved in a process, i.e., roles, departments, etc. The general purpose is to come up with an optimal structure for the organizational units.
3. The time perspective is focused on the timing and frequency of events.
4. The case perspective considers the properties that are inherent in different cases or types of processes. Further analysis defines relations and hidden dependencies between these properties, giving a deeper understanding of the problem and its causes.

Process mining can be beneficial for businesses of any size and in any industry as it can be applied to any workflow, with the purpose of process analysis and optimization. Process mining applies data science to discover, validate and improve workflows. By combining data mining and process analytics, organizations can mine log data from their information systems to understand the performance of their processes, revealing bottlenecks and other areas of improvement.

2.1 Main stages of process mining

As people (and software) interact with business IT systems, their actions are captured by these systems and can then be transformed into event logs and visualized with the help of process mining. That's how it happens as shown in the below figure.

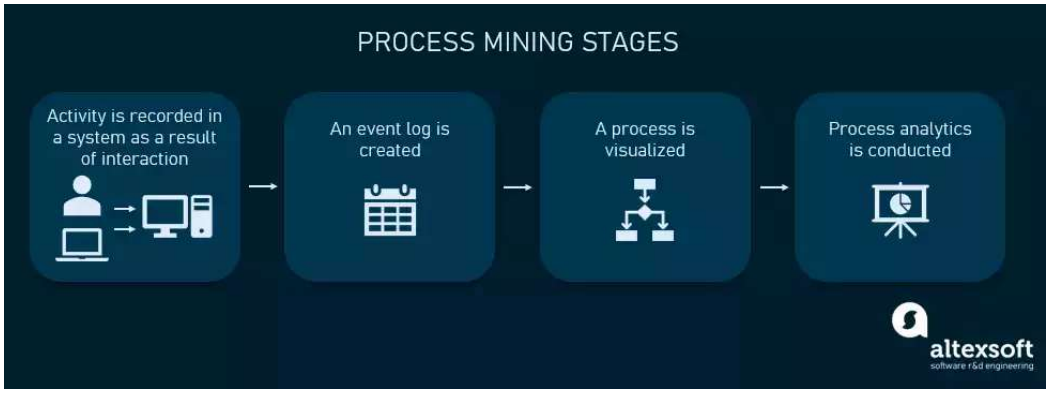
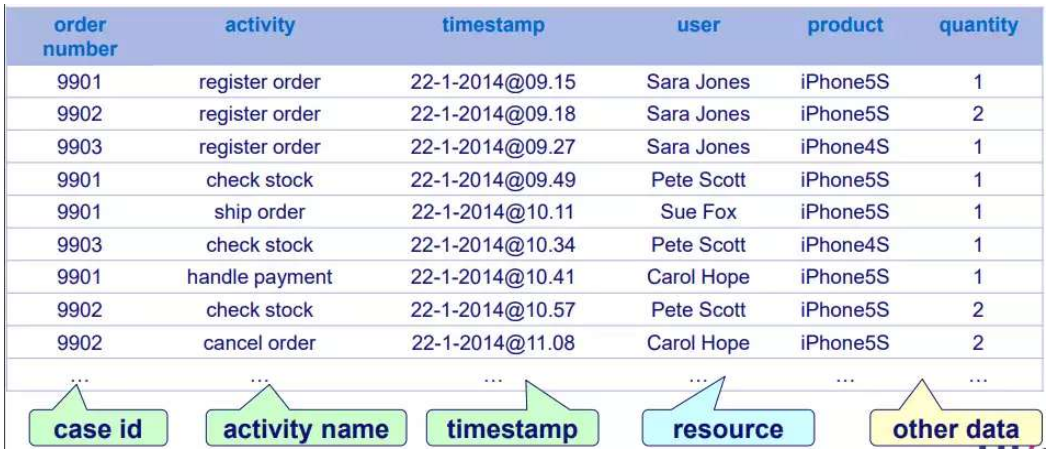


Fig. No. 2.1: Process mining stages

1. The activity or interaction with the system takes place, creating a digital record. Some examples of such activities are receiving an order, submitting a piece of documentation, approving a loan, entering information into a health record, etc.
2. Process mining software transforms the digital records into event logs. The most common format for these event logs is an XML-based format XES (eXtensible Event Stream) which was adopted by IEEE Task Force on Process Mining. Event logs have at least three main attributes: case ID, activity, and timestamp.
3. The visualization of a process is automatically created using event logs. It's important to understand that unlike traditional BPM techniques, process mining shows the real process as it's actually done, not the ideal model as it was meant to be.
4. Process analytics takes place. Here, KPIs can be created and monitored to uncover potential improvement areas; data mining and/or ML algorithms can be used to detect hidden patterns and dependencies; or conformance checking techniques can be applied to compare the process to a certain ideal model.



order number	activity	timestamp	user	product	quantity
9901	register order	22-1-2014@09.15	Sara Jones	iPhone5S	1
9902	register order	22-1-2014@09.18	Sara Jones	iPhone5S	2
9903	register order	22-1-2014@09.27	Sara Jones	iPhone4S	1
9901	check stock	22-1-2014@09.49	Pete Scott	iPhone5S	1
9901	ship order	22-1-2014@10.11	Sue Fox	iPhone5S	1
9903	check stock	22-1-2014@10.34	Pete Scott	iPhone4S	1
9901	handle payment	22-1-2014@10.41	Carol Hope	iPhone5S	1
9902	check stock	22-1-2014@10.57	Pete Scott	iPhone5S	2
9902	cancel order	22-1-2014@11.08	Carol Hope	iPhone5S	2
...	...	...	...	...	...

Fig. No. 2.2: A sample event log of order handling

2.2 Process mining techniques

According to the Process Mining Manifesto, there are three main groups of process mining techniques.

1. (Automated) Process discovery is a primary technique and implies extracting and visualizing process models from an event log like we described above.
2. Conformance checking compares the actual process with a predefined model to discover deviations. So, it's used to check if the reality conforms to an existing pattern.
3. Enhancement goes beyond measuring and comparing. It's focused on extending the model with additional information, such as location data, costs, timing, etc. Enhancing the model with such attributes helps improve its performance and conduct more advanced analysis.

Process mining approaches are roughly classified into three categories: process discovery, conformance checking, and process enhancement.

- Process Discovery
- Conformance Checking
- Enhancement and Extension

In the past the terms like Workflow Mining and Automated Business Process Discovery were used in the place of the below listed approaches. Let's know more about them going into the details. These approaches are as follows:

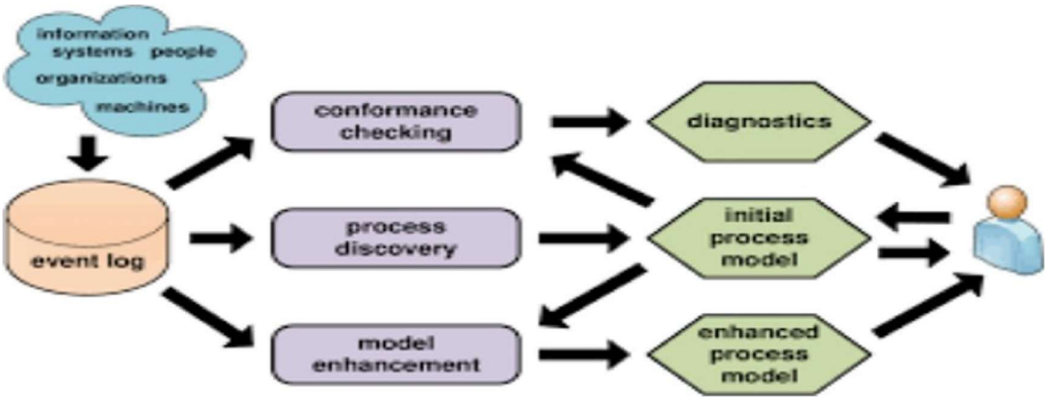


Fig. No. 2.3: Process Mining Techniques

These three types of process mining techniques: Process Discovery, Conformance Checking, and Enhancement/Extension to provide essential insights into understanding, analyzing, and optimizing business processes using event log data taken from the Information Systems.

2.2.1 Process Discovery:

The goal of process discovery is to extract a process model from event logs. Process discovery strategies include the following:

- **Alpha Algorithm**: The Alpha Algorithm constructs a process model based on the frequency of events. It begins with a direct succession link and gradually constructs a model.
- **Heuristic Miner**: This approach identifies common patterns in event logs to infer a process model. It uses heuristics to determine the most likely process.
- **Inductive Miner**: To develop a process model, this technique uses a divide-and-conquer strategy It divides event logs into smaller subsets and builds models for each one, which are then integrated to form a final process model.

2.2.2 Conformance Checking:

Conformance checking compares the discovered model with a reference or target model, which can be visualized as a superimposition of the two models. While it's ideal to see a high degree of similarity between the basic activities in both models, this isn't always the case. To find discrepancies and deviations, the identified process model is compared to the actual event logs in the IT systems.

The following are the primary techniques:

- **Token-Based Replay**: This technique uses tokens to recreate the execution of the process model and compares the generated traces to the observed event logs e Fitness Checking: It calculates the percentage of events that are appropriately aligned to determine how well the process model fits the event logs.
- **Precision and Generalization Checking**: These techniques assess process model accuracy by taking into account both overfitting and underfitting.

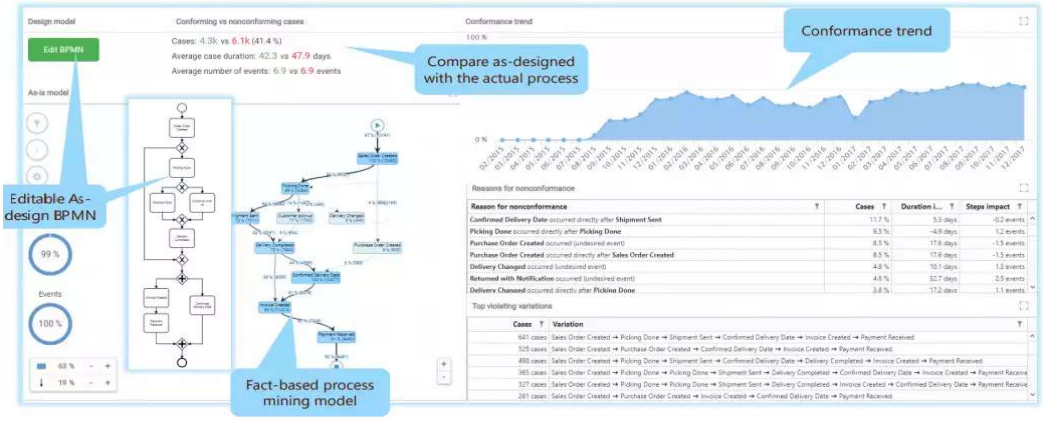


Fig. No. 2.4: Conformance Analysis

2.2.3 Enhancement and Extension:

Enhancement approaches are used to improve the quality of process models or expand their capabilities. Among the concepts covered in this lesson are discussed.

Subprocess Discovery discovers regularly occurring subprocesses inside a larger process, allowing for more extensive analysis and improvement. The concepts are like Multi-viewpoint Mining and Predictive Analytics.

CHAPTER - 3
APPLICATIONS

Process mining is a technology that uses event logs to discover, analyze, and improve real processes, such as business processes, information systems processes, and healthcare processes. It offers valuable insights into how processes actually work, allowing organizations to identify bottlenecks, inefficiencies, and compliance issues. Here are some key applications of process mining:

1. **Business Process Improvement:** Process mining can help organizations identify areas where their processes can be optimized or streamlined. By visualizing process flows and analyzing event data, organizations can pinpoint bottlenecks, inefficiencies, and areas for improvement.
2. **Process Discovery:** This involves automatically constructing process models from event logs. It helps organizations understand how processes are executed in reality, as opposed to how they were designed on paper. This can be especially useful when existing documentation is outdated or incomplete.
3. **Conformance Checking:** Process mining can be used to compare actual process executions with the prescribed or expected processes. This helps organizations identify deviations, compliance issues, and non-conformance with regulations or standards.
4. **Root Cause Analysis**: When issues or problems arise within a process, process mining can be used to trace back to their root causes. This is valuable for addressing the underlying problems rather than just addressing symptoms.
5. **Performance Monitoring:** Organizations can use process mining to continuously monitor the performance of their processes. This includes measuring key performance indicators (KPIs), analyzing cycle times, and identifying areas where service levels are not met.
6. **Auditing and Compliance:** Process mining can assist with audit processes by providing an objective view of process compliance. It helps ensure that processes adhere to legal requirements, industry standards, and internal policies.

7. **Resource Allocation:** By analyzing the data generated during process executions, organizations can make informed decisions about resource allocation. This includes workforce optimization, budget allocation, and equipment utilization.
8. **Supply Chain Optimization:** In supply chain management, process mining can be used to track the movement of goods, identify delays, and optimize inventory management, helping organizations reduce costs and improve efficiency.
9. **Healthcare Process Optimization:** Process mining is increasingly used in healthcare to analyze patient journeys, optimize hospital workflows, and improve the overall quality of care. It can help reduce waiting times, streamline administrative processes, and enhance patient outcomes.
10. **Customer Journey Analysis:** Organizations can use process mining to gain insights into the customer journey. This includes understanding how customers interact with digital platforms, identifying drop-off points in online processes, and improving user experiences.
11. **IT System Analysis:** Process mining can be applied to analyze how IT systems interact with each other and with users. It can help identify IT-related bottlenecks, failures, and opportunities for system improvements.
12. **Fraud Detection:** Process mining can be employed to detect fraudulent activities within organizations. By analyzing process data, it can identify unusual patterns or deviations that may indicate fraudulent behavior.
13. **Predictive Analytics:** Advanced process mining techniques can be used for predictive analytics, helping organizations forecast future process behavior and trends based on historical data.

Process mining offers a versatile set of tools and techniques that can be applied across various industries and domains to improve efficiency, compliance, and decision-making. Its applications continue to expand as organizations recognize its value in optimizing processes and gaining deeper insights into their operations.

CHAPTER - 4

MODULES

4.1 Module 1: Introduction to Process Mining:

Process Mining is a data-driven approach to understanding business processes, dominating operational excellence, automation, and digitalization ambitions in the industry. It is a leading technology for algorithmic businesses using algorithms and real-time data. Process Mining provides a solution to costly and time-consuming efforts to gain data-driven insights, as acknowledged by Gartner. Look at the picture below and know the differences.

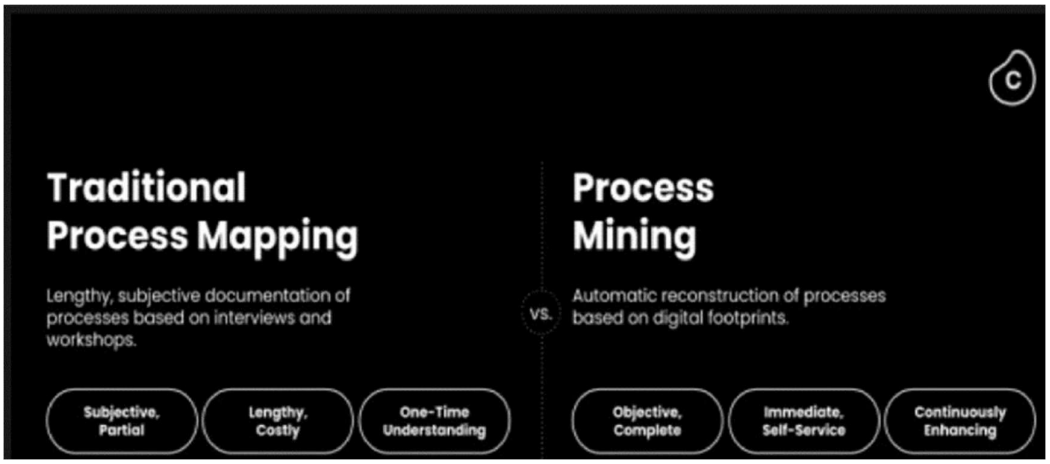


Fig. No. 4.1: Traditional Process Mapping vs. Process Mining

The complexity and visibility issues are resolved by Process Mining technology as opposed to conventional process mapping methodologies. A living, breathing perspective of the processes is provided that is generated instantly and is always up-to-date thanks to the real-time usage of system data, significantly lowering cost and time to value. It is like an x-ray for firms since it provides complete transparency into processes, removes process blind spots, and calculates the effect of process issues on key performance indicators. It uses information from user desktops and a company's source systems to map the processes, removing the need for speculation about how they function Because they can trust the data, organizations are better able to handle conflicting requirements from their stakeholders 4.1 Emergence of Process Mining: Formative Concepts:

Process Mining, a data-rich approach, emerged from scholarly investigations into the potential applications of event log data extracted from Information Systems.

It aimed to uncover, monitor, and enhance operational processes within organizations, laying the groundwork for optimizing various dimensions of Business Process Management (BPM). Process Mining provided valuable insights into the true paths and sequences of processes, conformance checking, organizational mining, automation, simulation, prediction, and history-based recommendations.

By analyzing patterns and trends, organizations could foresee potential process behaviors and outcomes, enabling proactive responses to challenges or opportunities. Process Mining heralded a paradigm shift in Business Process Management, bridging the gap between theoretical process models and real-world execution, facilitating enhanced process understanding, optimization, compliance, and overall organizational agility.

4.1.1 Event Logs:

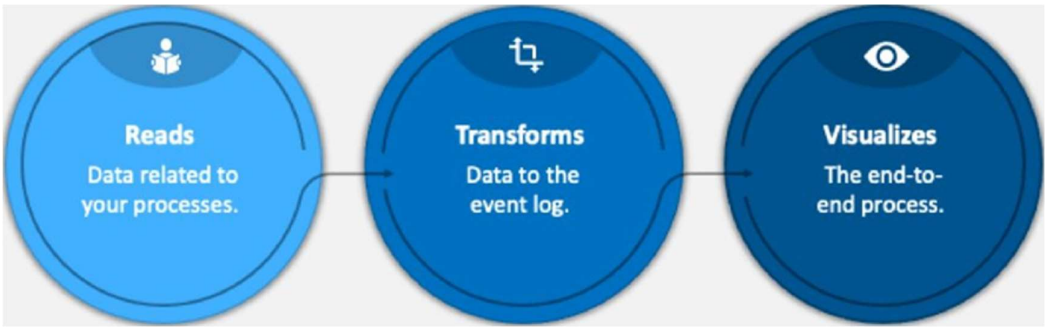


Fig. No. 4.2: Event Logs

Event logs are a format for obtaining digital traces from underlying IT systems like customer process mining (CRM), supply chain management (SCM), and enterprise resource planning (ERP) systems. These systems capture and instantly store events for every Case ID. A distinct Case ID, an activity description, and a timestamp are the three main pieces of information found in Event Logs. These data points can combine information across all Case IDs and reconstruct the process flow for a specific Case ID. A process is a chain of interconnected actions or stages carried out to accomplish a

certain goal, such as resolving a ticket in customer service, advancing a sales opportunity, or managing orders in order management. Event Logs can be examined and analyzed to help identify problems and incidents.

4.2 Module 2: Fundamentals of Process Mining:

Process mining is an analytical discipline that extracts knowledge from event logs in information systems to find, monitor, and improve processes. By autonomously executing insights and orchestrating current technologies, the Celonis Execution Management System (EMS) extends process mining. Organizations can employ transactional system digital footprints to acquire an objective, real-time view of their processes. Organizations can use dynamic visual representation and drill down tools to use exploratory or confirmatory approaches.

- **Exploratory Approach**: An exploratory approach involves just exploring the data to discover what value opportunities emerge. You are approaching the data with an open mind and no preconceived notions. Process Explorer, Variant Explorer, and Conformance Checker analysis tools are perfect for this.
- **Confirmatory Approach:** With the confirmatory approach, you're examining the data to see if it confirms or denies a hypothesis. Using your Celonis Analysis, specifically by filtering on attributes and using drill down tables, you can find out whether the data confirms or denies that these perceived pain points exist and have a significant impact.

With Celonis Analysis, our customers choose to use Celonis tools such as Action Flows (process automation) and Celonis Apps to maximize their organization's performance capacity. In this sense, they don't stop at Process Mining and leverage all that the Celonis Execution Management System (EMS) has to offer.

- **Process:** A process is a sequence of interconnected stages taken to attain a specific purpose.
- **Activity:** An activity is a stage in the process that occurs. Process activities are actions that begin or end a process or occur during it. Each activity consists of one or more tasks that, when completed, represent a milestone in the process.

- **Case:** An "item" or "object" that you follow through the procedure is referred to as a case. Even for the same business process, the case varies for each organization, depending on how detailed they wish to get.

    1. Events that take place during an activity process.
    2. An item or object you follow Case through the process.
    3. A series of linked steps taken in order to achieve a goal.

4.2.1 Variant Explorer:

The Variant Explorer is one of several analysis tools in an exploratory approach to uncovering insights. The variant Explorer shows you all the different ways the process flows in your organization, each called a variant. In other words, a variant represents one or more cases that have taken that exact set and sequence of activities. In the images so far and in the guided tour, the Variant Explorer was set to the Case Frequency RPI. Represented by a number, a Key Performance Indicator (RPI) allows you to quickly assess how your process is performing.

- **Case Frequency**: the frequency with which a case occurs. The number of unique cases linked to an activity or connection is shown in the RPI. Naturally, the number is the same across all activities and connections in a single version.

  Case Frequency shows 118 cases associated with this variant. Therefore, the count is across this entire single variant**.**

- **Activity Frequency:** Activity Frequency shows how many times each activity is performed. Occurred in total (236 times) for the 118 cases in the variant.

Activities means the number of different activities. Minimal frequency is how often the least frequent activity has occurred. Median frequency is the median of how often each activity has occurred. Mean frequency: How often each activity has occurred on average. Activities can be classified into types, there are Send, Receive, User, Manual, Business Rule, Service, and Script activity types.

In the Activities chart, each circle on the activities chart represents a certain activity that occurs in the analyzed process. The size and tone of the circle are associated with the frequency of events per day in which this activity occurs (the bigger and darker the bubble, the higher the frequency). By clicking on an Activity Circle, you are able to see metrics or filter your analysis to this activity. Variant Explorer shows all the different paths (or variants) a specific process takes within an organization. Process Explorer, on the other hand, does not show a specific process path (or variants) for a group of cases, but instead shows the most common activities and connections filtered on case frequency by default.

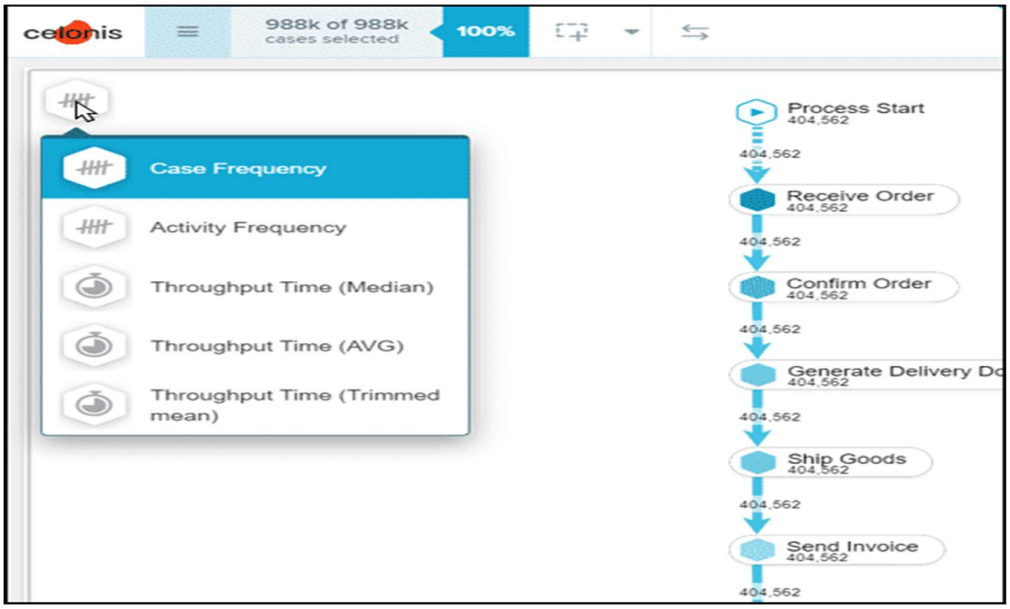


Fig. No. 4.3: Case Frequency

The activity frequency on "Generate Delivery Document" (236) is exactly double the case frequency (118); this represents the fact that each case in this variant, as indicated by the loop, runs through "Generate Delivery Document" twice the same can be said of the operation "Send 2nd Payment Reminder. " A side-by-side comparison of the identical version with Case Frequency and Activity Frequency KPIs is shown above take note of the difference in the "Generate Delivery Document" activity count.

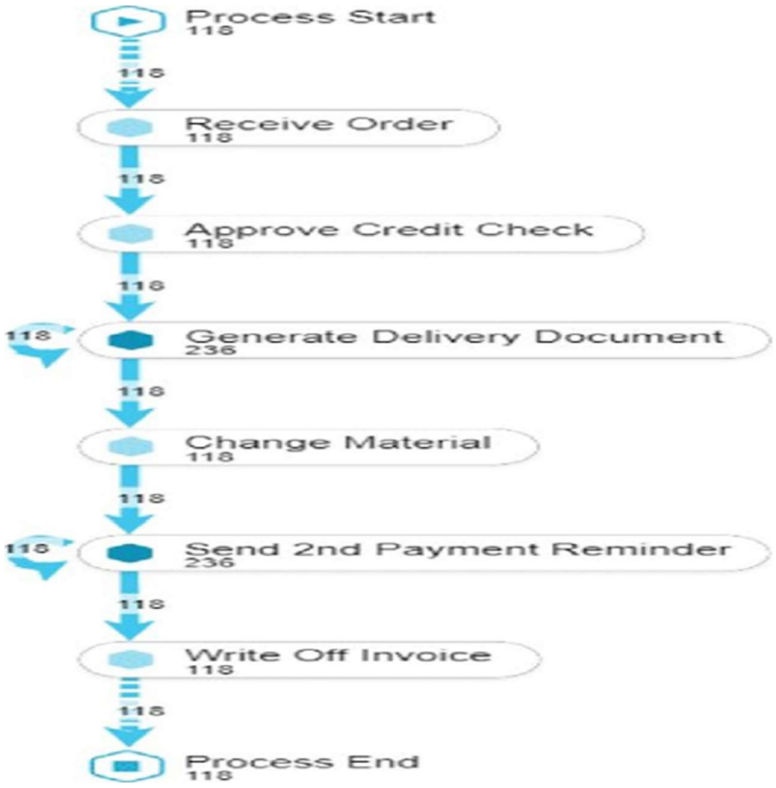


Fig. No. 4.4: Activity Frequency

4.2.2 Process Explorer:

The Process Explorer is another analysis tool to use when taking an exploratory approach. It's especially useful for quickly revealing activities beyond the most common ones. It also allows you to narrow your focus on a single activity, for example, an undesirable activity, to see which activities cases typically come from and which activities they're going to. You'll be nicely surprised by all that you can accomplish with it. When you display the Throughput Time RPI in the Process Explorer, you are looking at the time it took all of the cases in the analysis to go directly between the two presented activities.

In contrast, the time in the version Explorer reflects the cases in the version or variants selected. The individual who creates the analysis can change the measurements and KPIs. The automation rate is a typical custom RPI; it indicates the percentage of time that the task was accomplished automatically rather than manually.

4.2.3 Analysis Charts:

- A dimension is a category of attributes; for example, the dimension "customer name" is a category for individual customer names. Other examples of dimensions, depending on the nature of the process, can include vendor name, sales organization, region, and material group.
- Key Performance Indicators (KPIs) are used to calculate and add aggregated values; for example, case count, order value, invoice value, throughput time, and automation rate.

In case you're a bit fuzzy on these concepts, don't worry, it'll get clearer as we look at dimensions and KPIs in a column chart and an OLAP table.

4.3 Module 3: Technical Rising Star:

As businesses worldwide digitize, an increasing amount of log data is generated in their IT systems. This information is significant because it shows how business processes work within a company. Process Mining is a term that relates to data-driven methodologies for discovering, improving, and monitoring processes based on such data. Process Mining relies heavily on event logs.

4.3.1 Introduction to PQL and the Celonis Engine:

Executable Queries in Process Mining are used to gain valuable process insights, it is essential for Process Mining users to formalize their process questions as executable queries. For this purpose, we present the Celonis Process Query Language (Celonis PQL), which is a domain-specific language tailored towards a special process data model and designed for business users. It takes the input from the user and executes the queries in a custom-built query engine. PQL was designed to be easy-to-use, translate complex process questions into data queries, and make process mining accessible to every Celonis user.

- tailored towards a particular process data model and
- designed for business users.

It translates process-related business questions into queries and executes them on a custom-built query engine, the Celonis PQL Engine. The Process Query Language (PQL) is a domain-specific language tailored towards a special process data model and designed for business users. PQL enables the user to translate process-related business questions into queries, which are then executed by a custom-built query engine. For a PQL, the prospect's product usage data determines when the lead is ready to invest in your services.

4.3.2 Celonis Software Architecture:

The Celonis PQL Engine:

As you can observe in the graphic below, Celonis PQL is an integral component of the Celonis Software Architecture All Celonis applications use this language to query data from a data model. However, over the different components in the architecture overview to find out more. Click on "example" to explore a data model of a procurement process.

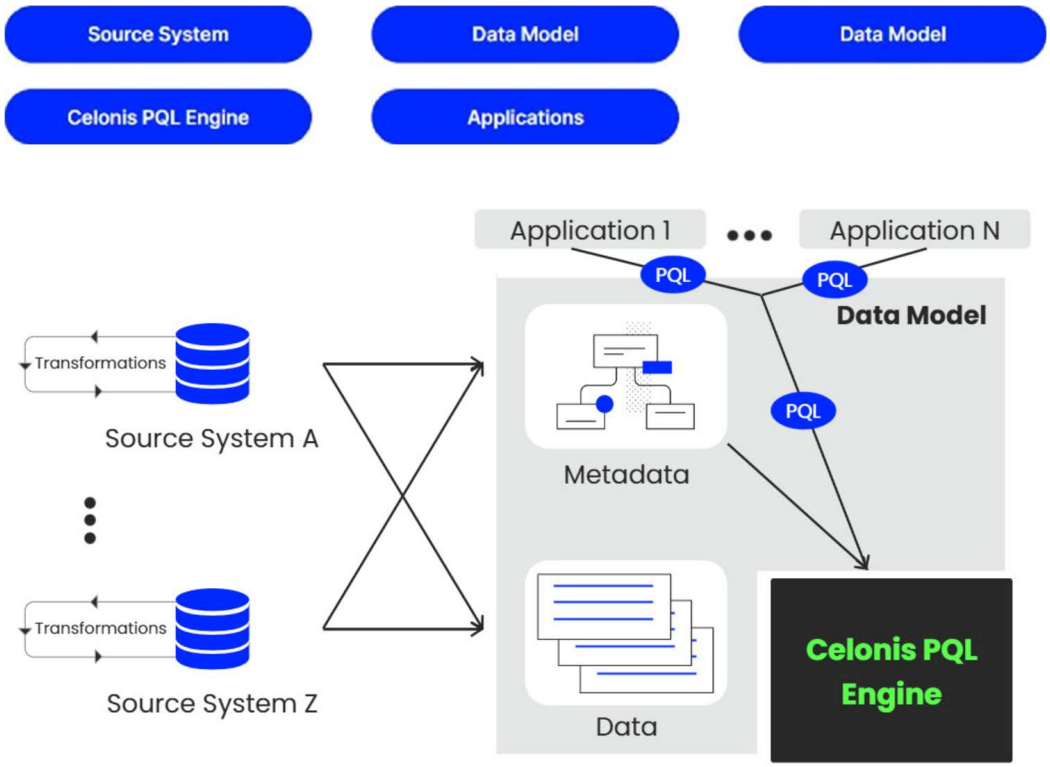


Fig. No 4.5: Celonis Software Architecture

Built on object-centric process mining (OCPM), the Celonis Object-Centric Data Model is an extensible data representation of an entire business. It acts as a single source of truth for all process intelligence and serves as the core of an organization's digital twin. The Celonis process mining software helps you find and capture value fast by improving the performance of your core business processes. The Celonis process mining software helps you find and capture value fast by improving the performance of your core business processes. In SQL, the marketing team qualifies based on their interaction with the brand.

4.3.3 PQL Queries:

PQL can be used in a variety of applications. It can be used in analyses, knowledge models, action flows, and other applications. However, you should not be concerned with imagery or design when constructing queries. You'd want to make a query, inspect the results, and then validate them to ensure they're correct. And here is where the Data Explorer comes in. PQL is a declarative language built on temporal logic. Temporal logic is an extension of classical propositional logic that includes operators that refer to system activity across time. These behavioral operators, known in PQL as predicted, give PQL a mathematically exact way of expressing features about the relationship between actions and occurrences in process instances.

- **Compactness:** PQL queries should enable intents to be captured in short, succinct programs that eliminate ungrounded code redundancy.
- **Decidability:** PQL queries should be algorithmically solvable across a wide range of inputs
- **Efficient:** PQL queries should necessitate reasonable and reachable computational resources.
- **Expressiveness:** PQL queries should allow for the description of many notions.
- **Intuitiveness:** PQL queries should be simple to read and understand.
- **Portability:** PQL searches should be able to run in a variety of contexts and data types.
- **Utility:** PQL queries should be able to perform a wide range of practical tasks.

4.3.4 Joins and Filters:

With Celonis PQL's basic aggregation capabilities, you can readily relate this logical example to what you can accomplish with your data. Numerous more common aggregations, such as summary statistics (min, max, median, quartiles, and standard deviation), are waiting for you in addition to counting, distinct process mining virtual internship counting, average, and total. In a data model, tables are linked together by particular relationships that link rows from one table to rows from another table A foreign key is used to accomplish this These connections can often be categorized as:

- One-to-many or I
- N or one-to-one e N:M, or many-to-many

Depending on the number of rows in one table that can be matched with a row in another table. In Celonis Data models, only one-to-many (l: N) relationships are supported. Every asset in Celonis has an underlying Data Model with multiple tables and I: N relationships. The joins between those tables are left-outer joins, where the N-side is on the left the tables are implicitly joined when writing a PQL query since all the relevant tables are initially gathered into a common table. This has effects, for instance, on RPI calculations that require many tables. The most N-sided table is usually the common table.

Dashboards provide the ability to build filters that can be used to change the data shown on every dashboard to take a deeper look at specific steps in the process, you can build subsets of your data using filters. The header bar and the Filters panel both allow for the creation of filters. The set of tables containing all the data is called the Data Model. In Celonis, every asset you build is based on an underlying Data Model. To obtain the desired information and interpret results or KPIs, it’s crucial that you know the structure of the Data Model you’re working with. The tables in a Data Model are connected via specific relationships to associate rows of one table with rows of another table. This is done using a foreign key.

4.3.5 Data Integration:

As a data engineer or analyst working in Data Integration (formerly known as Event Collection), you're responsible for bringing clean, real-time process data into the EMS In other words, you build the data pipeline. Process data is a collection of related actions with timestamps that follow a certain case or item. Every action is an "event," and it is your responsibility to compile and order all of these events. Well, that's precisely what Data Integration is for. It helps you connect to source systems, extract the relevant data, transform it to your needs, and load it into a polished Data Model. You can think of the Data Model as the fuel to all other work in your EMS. Once it's ready, your team picks it up and can get started on analyzing it.

When building the Data Model, you're also responsible for validation, scheduling, monitoring, as well as optimizing performance. In other words, in Data Integration, you lay the data pipeline foundation for Celonis implementations and are also responsible for its ongoing optimization and expansion. The more robust and performant your data foundation, the faster and more reliably your team can capitalize on the EMS' execution capabilities.

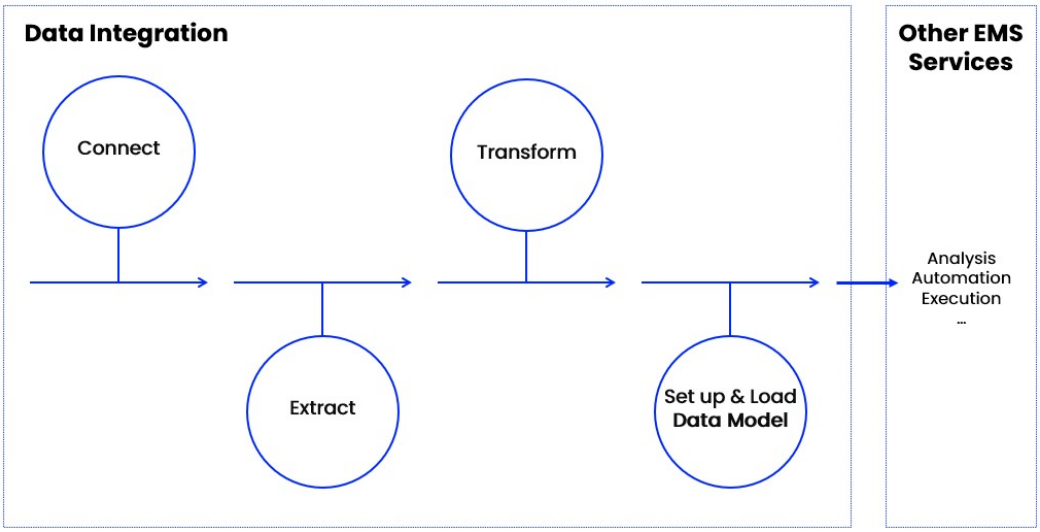


Fig. No. 4.6: Data Model

Celonis EMS is a powerful tool for integrating process data into its various systems- Data Integration helps connect to source systems, extract relevant data, transform it to your needs, and load it into a polished Data Model. This Data Model serves as the fuel for all other work in the EMS, allowing your team to analyze and act on it. The process of connecting to source systems involves using technologies like message queues, Restful APIs, Soap APIs, direct database access, or system-specific solutions.

Extracting data involves loading entire source system tables into Celonis, while transforming data into a Process Data Model involves creating the Activity table, which serves as the basis for the Data Model. The process of creating a visual representation of an information system communicates connections between data points and structures. Finally, an Execution Management System (EMS) is a specialized software platform that facilitates the execution of trades in financial markets, providing traders with tools to route orders, access market data, manage orders, and execute trades efficiently across multiple venues and asset classes.

Every component of the Celonis EMS is geared towards maximizing the execution capacity of your business. Celonis Execution Instruments analyze your processes, find bottlenecks and identify the gaps between how things should be and how they really are. Celonis Execution Applications provide best in class knowledge, actions and automations to systematically remove these execution gaps. The Celonis EMS comprises the following features:

- **Studio and PQL**: Transparency into how your process runs.
- **Data Integration**: Flexible development platform for data integration projects. New Celonis Connector Technology.
- **Marketplace**: Instead of weeks of IT setup, install and set up Marketplace Apps with a couple of clicks.

CHAPTER - 5

REAL TIME APPLICATIONS

Process mining is a data-driven methodology that helps organizations discover, monitor, and improve their business processes by analyzing event logs and transaction data. Real-time applications of process mining involve using this methodology to gain insights and make informed decisions as events occur. Here are some real-time applications of process mining:

- **Real-time Monitoring:** Process mining tools can continuously monitor processes as they happen, providing real-time insights into process performance. Alerts can be triggered when predefined conditions or deviations from the expected process occur, allowing organizations to take immediate corrective actions
- **Process Compliance:** Organizations can use process mining in real time to ensure that employees adhere to regulatory compliance and company policies. When a process deviation occurs, it can be detected and addressed promptly to prevent compliance issues.
- **Manufacturing Process Control:** In manufacturing, process mining can be applied to monitor and control production processes in real time. It helps identify quality issues, machine downtime, or production bottlenecks as they occur, allowing for immediate corrective actions.
- **Logistics and Transportation:** In the logistics and transportation industry, real-time process mining can optimize route planning, monitor vehicle and cargo conditions, and track deliveries in real time. It ensures that operations run smoothly and efficiently.
- **Energy and Utilities Management:** Process mining can be applied to monitor and optimize energy consumption and resource utilization in real time. This helps organizations reduce costs and minimize their environmental footprint.
- **Customer Support and Service Management**: Real-time process mining can be applied to track and optimize customer support processes. It helps organizations identify issues in real time, allocate support resources efficiently, and improve customer satisfaction by addressing problems promptly.

- **Human Resources Management**: In HR, real-time process mining can help organizations streamline recruitment, onboarding, and performance management processes. It allows HR teams to identify bottlenecks and make data-driven decisions to improve employee experiences.
- **Financial Transactions and Fraud Prevention**: Financial institutions can use real-time process mining to monitor and analyze financial transactions for fraud detection and prevention. It helps identify suspicious activities immediately, reducing financial losses.
- **Retail Operations:** Real-time process mining can optimize various retail operations, such as inventory management, cashier operations, and customer checkouts. It enables retailers to respond quickly to changing customer demand and market trends.
- **Event and Conference Management:** Event organizers can use real-time process mining to monitor and manage event logistics, attendee registration, and scheduling. This ensures that events run smoothly and adjustments can be made on the fly.
- **Emergency Response and Disaster Management:** In emergency situations, such as natural disasters or public health crises, real-time process mining can be used to coordinate response efforts, allocate resources efficiently, and monitor the execution of emergency plans.
- **Quality Control in Manufacturing**: Real-time process mining can be applied to ensure product quality in manufacturing. By continuously monitoring production processes and quality control checks, organizations can identify defects and take immediate corrective actions.
- **Call Center Operations:** Call centers can benefit from real-time process mining to monitor call volumes, agent performance, and customer satisfaction. It helps optimize call routing and staffing levels to meet service level agreements.

Real-time process mining is versatile and can be applied to various industries and domains where processes can benefit from continuous monitoring, optimization, and adaptability. It helps organizations stay agile, improve operational efficiency, and respond quickly to changing conditions and customer needs.

CHAPTER - 6

LEARNING OUTCOMES

After completing this program, we ought to be able to:

1. Recognize what Process Mining is and the fundamental principles underlying its operation.
2. Briefly describe what an event log is and why process mining requires it.
3. Determine the business applications of process mining.
4. Being able to recognize, evaluate, and enhance business processes using data-driven methodologies.
5. To draw conclusions from event logs, spot bottlenecks, inefficiencies, and optimization possibilities.
6. To extract and produce visual representations of processes to support efforts at process improvement and decision-making.
7. Develop expertise in interpreting results from process mining tools to improve organizational effectiveness and efficiency.

CONCLUSION

The process mining internship has been an enlightening experience, providing organizations with a powerful lens to scrutinize their operational intricacies, leading to refined efficiency, heightened compliance, and amplified performance. The internship has not only provided valuable learning opportunities but also practical skills in data manipulation and collaboration with diverse experts. The ability of process mining to uncover areas for improvement has been illuminating, as it has highlighted the vital role of data-driven insights in shaping optimized processes.

The internship has been an immersive education, honing data navigation and instilling the value of cross-disciplinary teamwork. The power of combining analytical acumen with domain experts' perspectives has enriched the overall understanding of process dynamics. The internship exemplifies the symbiotic relationship between learning and contribution, as it has made a tangible impact on refining and advancing operational excellence. The experience serves as a reminder that the pursuit of efficiency is a continuous journey that unfolds.

REFERENCES

[1] Celonis login: https://academy-login.celonis.com/s/login/

[2] Celonis process mining link: https://academy.celonis.com/learn/dashboard/

[3] Celonis rising star technical link:

https://academy.celonis.com/pages/academic-rising-star

[4] Process Mining link: https://www.processmining.org/