# Process Mining

## Automatic process discovery for monitoring and steering towards business objectives

*Process mining techniques can extract knowledge from event logs commonly available in today's information systems.*

*These techniques provide the means to automatically discover process flows in combination with the related business performance and conformance metrics.*

*By referencing the recorded event chains and using predictive indicators related to, for example, cost or duration, near real-time steering of the processes execution can be enabled.*

*The effects of the decisions taken are tracked for their effectiveness to the business objectives set.*

Process mining is widely available for adoption in operational support systems. Process mining covers a range of special techniques and algorithms. Process mining can be positioned between computational intelligence and data mining on the one hand and process modelling and analysis on the other hand.

The promise of process mining starts with the *discovery of real processes execution* (i.e., not assumed processes) and this up to monitoring of the effects after the introduction of changes. The approach starts with *extracting knowledge from event logs*.

Knowledge related to:

(a) *performance*: total duration, activity duration, waiting time;

(b) *performance inhibitors*: bottlenecks, loops, ping-pong;

(c) *conformance*: i.e., analyse deviations by comparing model and log;

(d) *resources*: *Resource Mining* or tracking hand-over of work between resources;

(e) processing cost: Process execution-based cost tracking to steer the execution toward the targeted cost objectives.

BDO

# 1. PROCESS MINING BETWEEN DATA MINING AND BPM

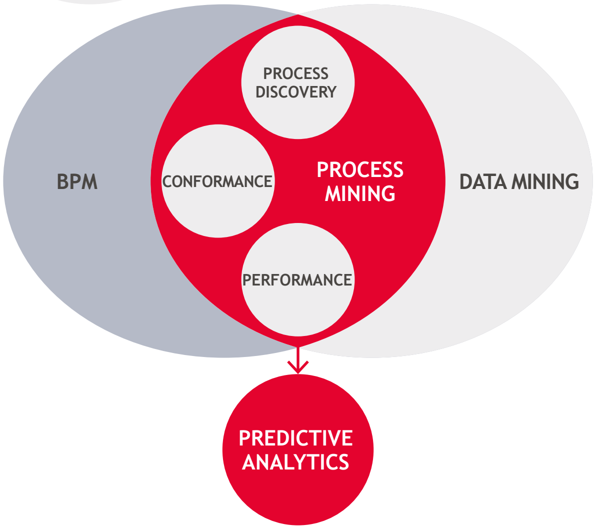



*Figure 1: Process mining interacting and complementing data mining and Business Process Management initiatives.*

*Process mining provides an important bridge between data mining and business process modelling and analysis*. Under the *Business Intelligence* (BI) umbrella, many buzzwords have been introduced to refer to reporting and dashboard tools. *Business Activity Monitoring* (BAM) refers to technologies enabling the real-time monitoring of business processes. The set-up has proven to be impressive let alone the maintenance.

*Complex Event Processing* (CEP) refers to technologies to process large amounts of events, utilising them to monitor, steer and optimise the operations in real-time. *Corporate Performance Management* (CPM) is focused primarily on measuring the financial performance of a process organisation. Also related are management approaches such as *Continuous Process Improvement* (CPI), Business *Process Improvement* (BPI), *Total Quality Management* (TQM), and *Six Sigma*. These approaches have in common that processes are "put under a microscope" to see whether further improvements are possible.

Process mining is an enabling technology for CPM, BPI, TQM, and Six Sigma projects with the possibility to be embedded in operational management platforms. BI tools and management approaches such as *Six Sigma* and TQM aim to improve operational performance, e.g., reducing flow time and defects. ***Process Mining provides the "hard" data for these initiatives***. Organisations are putting more emphasis on GRC (*governance, risks, compliance*). Legislations such as the Sarbanes-Oxley Act (SOX) and the Basel II require the focus on compliance.

Process mining techniques offer factual intelligence of the actual compliance and ascertain the validity and reliability of the findings. Management trends related to process improvement (e.g., Six Sigma, TQM, CPI, and CPM) and compliance (SOX, BAM, etc.) are benefitting from process mining.

Process mining algorithms are embedded in dedicated exploration tooling. We provide integration services to embed Process Intelligent (PI) algorithms in your operational systems.

## Bridging Traditional Data Analytics and Process Analytics

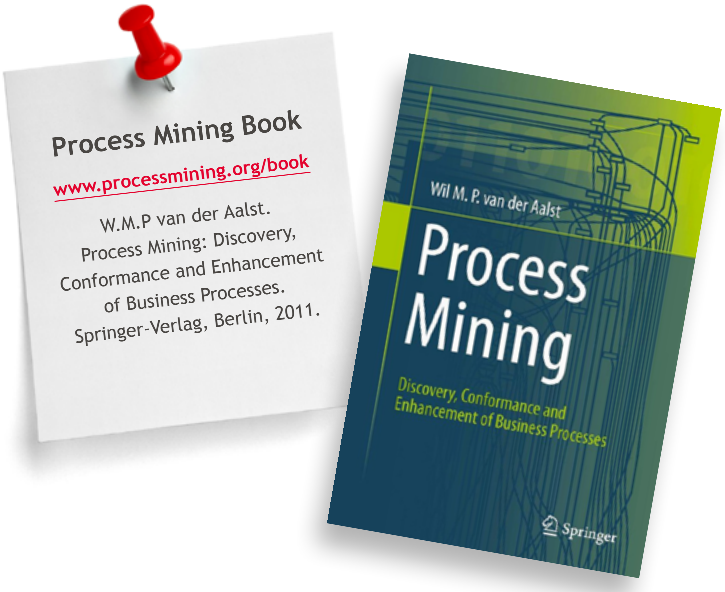
| Characteristic | Data Analyses | Process Analyses |
|---|---|---|
| Data Structure | Flat<br>(Table of records) | Hierarchical<br>(Cases with events) |
| Central Object | Events | Cases |
| **Analyses** | **(Traditional Data Analytics Support)** | |
| Process Discover<br>(Extracting visual process maps) | Partial/Not | Easy[1] |
| Conformance Testing<br>(End to end process flow review) | Partial/Not | Easy[1] |
| Statistical Analysis<br>(Extract distributions, correlations, relations) | Easy | Hard[1] |
| Filtering | Easy at event level | Easy[1]<br>for some filters |
| Querying | Easy at event level | Hard[1] |

*Source: Seppe vanden Broucke et al. (KU-Leuven).*
[1] Hence the need/opportunity for process analyses tooling.

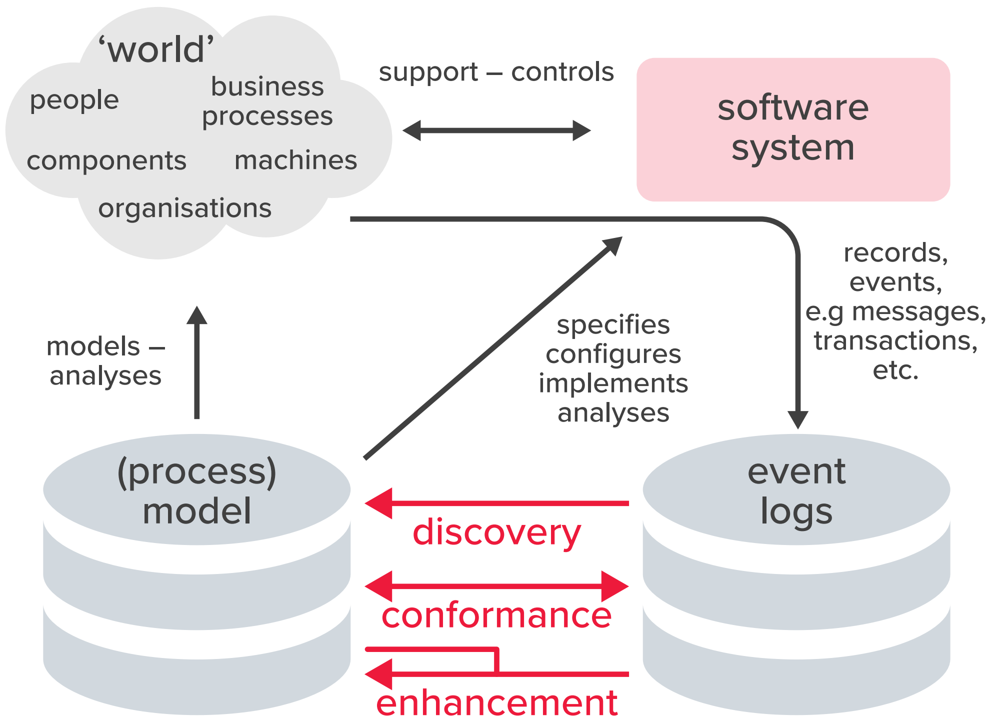
*Table 1: Bridging traditional Data Analytics and Process Analytics.*

# 2. PROCESS MINING

The expanding capabilities of information systems depend on computing techniques and power. The challenge is to exploit event data to provide insights, identify bottlenecks, anticipate problems, record policy violations, recommend countermeasures, and streamline start with *event logs*.




Process mining techniques start with detecting the event hierarchy in the ledger of records. Each event refers to an *activity* (i.e., a well-defined step in some process), that is related to *a case* (the full chain of activities i.e., a customer service order, incident ticket). Event logs hold relevant information about the sequence of activities. For each: (a) the *resource* (i.e., person or device) executing or initiating the activity, (b) the *timestamps* of the activity (Start and End time of an activity), (c) other *data elements* recorded with the event. (e.g., the cost of an activity, a document reference, the client, the supplier, …).




Source: Will van der Aalst, Process Mining.

*Figure 2: Positioning of the three main types of process mining: (a) discovery, (b) conformance checking, (c) enhancement.*

BDO

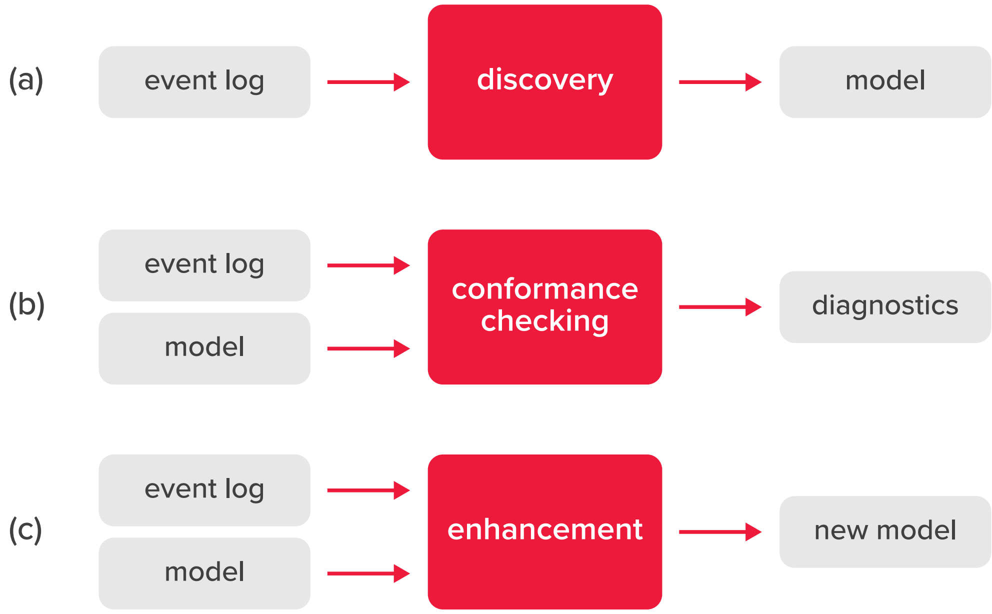



*Figure 3: The three basic types of process mining explained in terms of input and output: (a) discovery and process performance measurement, (b) conformance checking, and (c) enhancement.*

As shown in Fig. 2, event logs can be used to conduct three types of process mining. The first type of process mining is *discovery*. A discovery technique takes an event log and produces a model without using any a-priori information. Process discovery is the most prominent process mining technique. For many organisations, it is surprising to see that existing techniques are indeed able to discover real processes merely based on example executions in event logs. The second type of process mining is *conformance*. Here, an existing process model is compared with an event log of the same process. Conformance checking can be used to check if reality, as recorded in the log, conforms to the model and vice versa. Note that different types of models can be considered: conformance checking can be applied to procedural models, organisational models, declarative process models, business rules/policies, laws, etc. The third type of process mining is *enhancement*. Here, the idea is to extend or improve an existing process model using information about the actual process recorded in some event log. Whereas conformance checking measures the alignment between model and reality, this third type of process mining aims at changing or extending the a-priori model. For instance, by using timestamps in the event log one can extend the model to show bottlenecks, service levels, throughput times, and frequencies.

Figure 3 describes the three basic forms of process mining.

**(a) Discovery**. Turn the event log in process event chains, 1 for every case. Event chains are mapped into a process model of choice. (Petri net, BPMN, EPC, or UML activity diagram). The model may also describe other perspectives (e.g., a resource network). Conformance checking techniques need an event log and a model as input. The output consists of diagnostic information showing differences and commonalities between the model and log. Techniques for model enhancement (repair or extension) also need an event log and a model as input. The output is an improved or extended model.

Process mining may cover different perspectives. The *control-flow* perspective focuses on the control-flow, i.e., the ordering of activities. The goal of mining this perspective is to find a good characterisation of all possible paths. The result is typically expressed in terms of a Petri net or some other process notation (e.g., EPCs, BPMN, or UML activity diagrams). The *organisational perspective* focuses on information about resources hidden in the log, i.e., which actors (e.g., people, systems, roles, or departments) are involved and how are they related. The goal is to either structure the organisation by classifying people in terms of roles and organisational units or to show the resource network. The *case perspective* focuses on the properties of cases. A case can be characterised by its path in or by the actors working on it.

## Process Mining Characteristics:

### 1. Process mining goes beyond control-flow discovery.

Control-flow discovery is often seen as the most exciting part of process mining. However, process mining is not limited to control flow discovery. On the one hand, discovery is one of the three basic forms of process mining (discovery, conformance, and enhancement). On the other hand, the case and time perspectives next to the organisational perspective also provide important business opportunities.

### 2. Process mining is not just a specific type of data mining.

Process mining covers the "missing link" between data mining and traditional model driven BPM. Most data mining techniques are not process centric at all. Process models potentially exhibiting concurrency are incomparable to simple data mining structures such as decision trees and association rules. Therefore, completely new types of representations and algorithms are used.

### 3. Process mining is not limited to offline analysis.

Process mining techniques extract knowledge from historical event data. Although "post mortem" data is used, the results can be applied to running cases. For example, the completion time of a partially handled customer order can be predicted using a discovered process model. The resulting operational process intelligent logs (case, event-x and variant) are to be considered as learning ledgers and reference case stores.

Cases can also be characterised by the values of the corresponding data elements. For example, if a case represents an insurance claim, it is interesting to know the claim officer, the experts consulted, and the amount paid to the expert. The *time perspective* is concerned with the timing and frequency of events. As events bear time-stamps (Start and/or End of an activity) it is possible to discover bottlenecks, measure service levels, monitor the utilisation of resources, up to predicting the remaining processing time of running cases.

#### Process mining as a supporting technique

The Business Process Management (BPM) life-cycle is shown in Fig. 4. The BPM life-cycle shows seven phases of a business process and its corresponding information system(s). In the (re)design phase a new process model is created or an existing process model is adapted. In the analysis phase, a candidate model and its alternatives are analysed. After the (re)design phase, the model is implemented (implementation phase) or an existing system is (re)configured (reconfiguration phase). In the execution phase, the designed model is enacted. During the execution phase, the process is monitored. Moreover, smaller adjustments may be made without redesigning the process (adjustment phase). In the diagnosis phase, the enacted process is analysed and the output of this phase may trigger a new process redesign phase. Process mining is a supporting tool for most of the phases shown in Fig. 4. Obviously, the diagnosis phase can benefit from process mining.

#### Process mining and *operational support*

Predictions and recommendations based on the model(s) learned can be used to steer running cases. Decision support cockpits are equipped with reference data and metrics to adjust processes and guide process execution. Whereas Fig. 4 shows the overall BPM life-cycle, Fig. 5 (page 7) provides insight into how to approach an operational service process intelligent (OSPI) initiative.

Figure 5 describes the stages in such a set-up. Any OSPI project starts with a planning and a justification for this planning (Stage 0: Set-Up). After initiating the project, event data, models, objectives, and questions need to be extracted from systems, domain experts, and management.

### Stage 1: Data Preparation
This requires an understanding of the available data ("What data are available for the analysis?") and an understanding of the domain ("What are the business questions?") and results in the focus area's shown in Fig. 5 (i.e., event log data ledger, inventory of handmade models, business objectives, and research questions).

### Stage 2: Data Exploration
The OSPI API generates 3 process analytic views. The Case Log, The extended Event log and The Process Variant Logs. The Case log holds the computed metrics for every case like duration (Min, Max, Avg, Median), the number of activities performed, processing cost, % usage by resource skill or qualification, …

The Extended Event Log holds the original events and reference indexes to the variant log.

The Variant Logs hold the *unique* event chains from an activity flow or resource flow perspective. Knowledge is acquired from these 3 Process Intelligent files using advanced data and process visualisation techniques. These insights are used for business management steering like suggestions for process adjustments, instructions for more diligent data entry procedures, (extra) validation steps for more granular steering of the process execution.

### Stage 3: Operational Analytics
The acquired knowledge is used to prepare the initial "rules" or "boundaries" for operational steering of the process execution towards the business targets set. (example: thresholds of expert time dedicated to a case category, external cost thresholds, maximum waiting time before a customer touchpoint).

### Stage 4: Operational Intelligence
Design, implementation and ongoing support of the OSPI Cockpit.

## Guiding Principles

**GP1:** Event Data are to be Treated in a Record to Analyse manner.

**GP2:** Events from source file extraction and Event Log construction are driven by business questions.

**GP3:** Control-flow constructs are supported.

**GP4:** Events should be related to model Elements.

**GP5:** Models should be treated as purposeful abstractions of reality.

**GP6:** Process Mining is a Continuous Process.

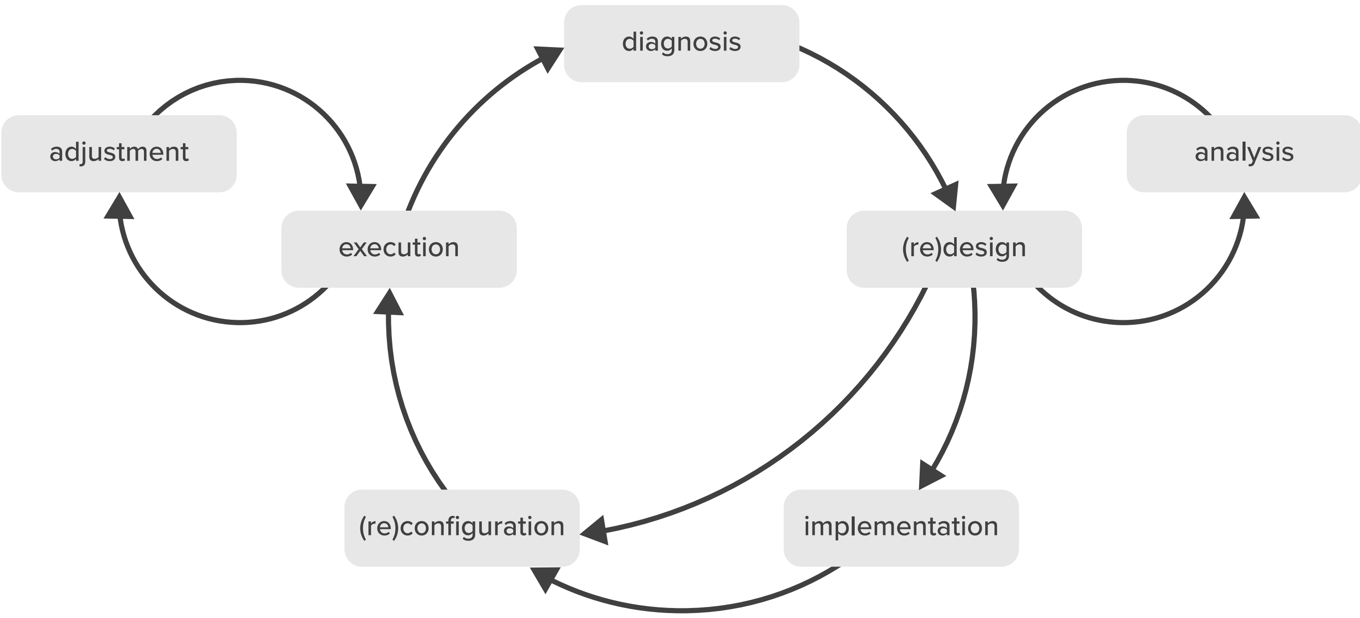



*Figure 4: The BPM life-cycle identifying the various phases of a business process and its corresponding information system(s); process mining (potentially) plays a role in all phases (except for the implementation phase).*

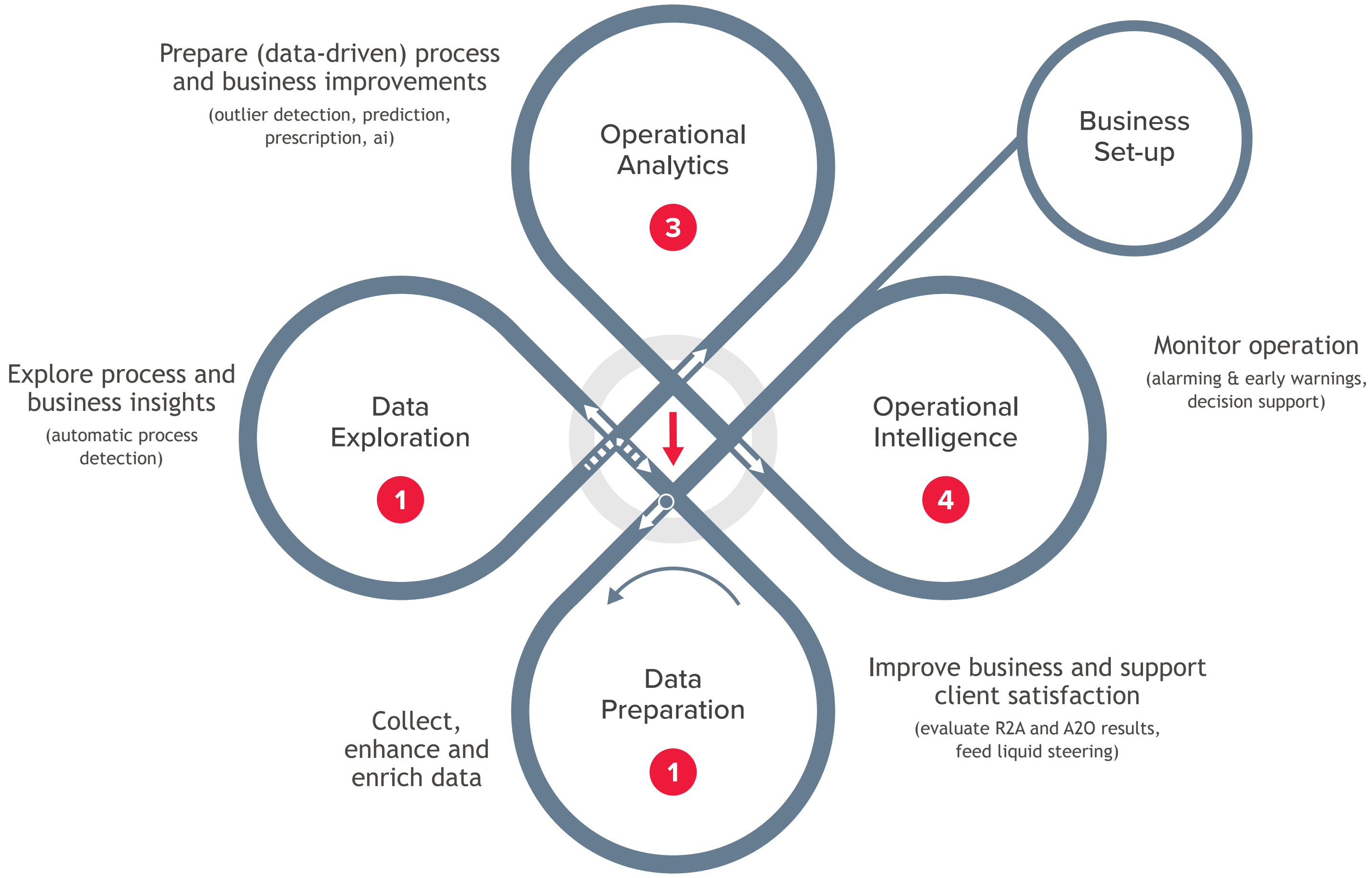



*Figure 5: The BDO Operational Service Process Intelligence approach.*

# 3. OPPORTUNITIES & CHALLENGES

Process mining becomes a unique opportunity for cost and user-satisfaction-sensitive organisations that need to manage non-trivial operational processes. For instance, in the case of service outsourcing (BPO). On the one hand, there is an incredible growth of event data although sometimes in dispersed systems. On the other hand, process execution and operational information need to be actionable to meet requirements related to efficiency, customer satisfaction, compliance, and complaint handling. With the use of event data as a starting point stated opportunities will have challenges that need to be addressed. Challenges ahead might be:

## C1: Finding, Merging, and Cleaning event data

It can take special efforts to extract event data and prepare them for process mining.

- Data may be *distributed* over a variety of data sources. These data need to be merged into a single ledger. Finding the identifiers to link the different event data is the initial task.
  For example, one system uses a name and birthdate to identify a person whereas another system uses the person's social security number. Time stamps are coming from different systems without the systems being on a "World-Clock" synchronisation. (Different time zones for instance).

- Event data are often "object-centric" rather than "process-centric". For example, individual products, pallets, and containers may have RFID tags and recorded events refer to these tags. However, to monitor a customer order such object-centric events need to be merged and pre-processed.

- Event data may be *incomplete*. A common problem is that events do not explicitly point to process instances. Often it is possible to derive this information, but this may take some effort. Also, time information may be missing for some events. One may need to interpolate timestamps to still use the timing information available. Different data/ time formats often occur between different log files.

- An event log may contain *outliers*, i.e., exceptional behaviour also referred to as noise. How to define outliers? How to detect such outliers?
  These effects and occurrences need to be identified to be addressed such as cleaning.

- Logs may contain events at d*ifferent levels of granularity*. In the event log of a claim management information system events may refer to contract or polis and maybe not a case or claims. Multiple claims can occur in parallel. Also, timestamps may have different levels of granularity ranging from milliseconds precision (28-9-2016h11m28s32ms342) to coarse date information (28-9-2016).

- Events occur in a context (Epidemic, workload, social condition, day of the week, etc.). This context may explain certain phenomena, e.g., the response time is longer than usual because of work-in-progress or holidays. For analysis, it is desirable to incorporate this context. This implies the merging of event data with contextual data. Here the "curse of dimensionality" kicks in as analysis becomes intractable when adding many variables.

As indicated earlier, organisations need to treat event logs in a Record to analyse manner rather than some by-product.

## C2: Dealing with complex event logs having diverse characteristics

Event logs may have very different characteristics. Some event logs may be extremely large making them more difficult to pre-process whereas other event logs need to be constructed since the activities are stored within a single record in subsequent fields.

As event logs contain only sample behaviour, they should not be assumed to be complete. Process mining techniques need to deal with incompleteness by using an "open world assumption": the fact that something did not happen does not mean that it cannot happen.

The use of tools allowing for quick data analyses and process mining feasibility tests is advisable.

The test can indicate potential performance problems and warn for logs that are far from complete or too detailed.

Some logs contain events at a very low abstraction level. These logs tend to be extremely large and the individual low-level events are of little interest to the stakeholders. Therefore, one will need powerful process attribute filters or activity aggregator functionality.

## Opportunities in the Challenges

**C1**: Finding, Merging, and Cleaning event data.

**C2**: Dealing with complex event logs having diverse characteristics.

**C3**: Representative business cases.

**C4**: Dealing with Concept Drift.

**C5**: Improving the representational bias used for process discovery.

**C6**: Balancing between quality criteria such as Fitness, Simplicity, Precision, and Generalisation.

**C7**: Cross-organisational mining.

**C8**: Providing operational support.

**C9**: Combining Process Mining with other types of analysis.

### C3: Representative business cases are emerging

Operational Process mining is for most companies new at the date this document is published. Also, companies don’t want to publish to fast their initiative(s) in this field.

Representative Business Cases are growing pointing to best practices in approach and use of solid techniques.

Most of the commercial tooling in the market focus on Exploration for Insight leaving the "user" to determine the actions and subsequent follow-up of the effect of the actions.

The most effective approach to date adheres to the "Glover with 4 leaves approach" (Fig. 5) to reap the business benefits of all stages and satisfy all stakeholders. "Don’t find other/new issues, please solve a handful" is a common request.

### C4: Dealing with “Concept Drift”

The term *concept drift* refers to the situation in which the process is changing by the measures taken to rationalise or adjust it while being analysed.

For instance, at the beginning of the measurements two activities may be concurrent (2 resources) whereas later, after resource reduction initiatives, the log will indicate that this job (the activities) is executed by the same resource (more experienced hence more expensive). Processes may change due to periodic/seasonal changes (e.g., "in December there is more demand" or "on Friday afternoon there are fewer employees available") or due to changing conditions (e.g., "a more competitive market ") or organisational change.

Induced changes impact processes and it is vital to detect, analyse, and operationally frame them. Concept drift in a process can be discovered by splitting the event log into smaller logs and analysing the "footprints" of the smaller logs. Such "second-order" analysis requires much more event data. Nevertheless, few processes are in a steady state, and understanding concept drift is of prime importance for the management of processes. That’s why additional algorithms support is used to detect concept drift and the effect on the models (prediction).

### C5: Improving the representational bias used for process discovery

Process discovery techniques produce a model which is a process model deducting from the event data. Different algorithms produce different representation models. In practice most used algorithms are (a) Fuzzy Miner, (B) Heuristic Miner and (C) Multiphase Miner (D) Fusion Miner.

A fuzzy-mining algorithm uses significance/ correlation metrics to interactively simplify the process model at the desired level of abstraction and hence are good for the exploration of cases with many activities.

It allows for visualising very complex process flows dynamically from simplified to complex "spaghetti" representations. The fuzzy miner visualisations do not allow good monitoring over various periods or compliance checks.
The other miners allow a more structural representation. A heuristic net can be converted to other types of process models, such as a Petri net. The Multi-phase miner was the first algorithm to explicitly use the OR split/join semantics, as found in EPCs.

### C6: Improving the representational bias used for process discovery

A process discovery technique produces a model using a language (e.g., BPMN or Petri nets). However, it is important to separate the visualisation of the result from the representation used during the actual discovery process. The selection of a target language often encompasses several implicit assumptions. It limits the search space; processes that cannot be represented by the target language cannot be discovered. This so-called "representational bias" used during the discovery process should be a conscious choice and should not be (only) driven by the preferred graphical representation.

### C7: Balancing between quality criteria such as fitness, simplicity, precision, and generalisation

Event logs are often not complete, i.e., only example behaviour is given. Process models typically allow for an exponential or even infinite number of different traces (in the case of loops). Moreover, some traces may have a much lower probability than others. That's why it is not realistic to assume that every possible trace is present in the event log.

A Complication might be that some alternatives are less frequent than others. These may be considered "noise". It is not possible to build a reasonable model for such noisy behaviours. The discovered model needs to be abstracted from this; it is better to investigate low-frequency behaviour using conformance checking.

### C8: Cross-organisational mining opportunity

Traditionally, process mining is applied within a single organisation. However, as outsourcing, supply-chain integration, and cloud computing become more widespread, there are scenarios where the event logs of multiple organisations are available for analysis. In principle, there are two settings for *cross-organisational process mining*.

The collaborative set-up is where different organisations work together to handle process instances. One can think of such a cross-organisational process as a "jigsaw puzzle", i.e., the overall process is cut into parts and distributed over organisations that need to cooperate to successfully complete cases. Analysing the event log within one of these organisations involved is insufficient. To discover end-to-end processes, the event logs of different organisations need to be merged. This is a non-trivial task as events need to be correlated across organisational boundaries. Some larger organisations, therefore, provide "data windows" to their service providers while keeping the data at their end.

We may also consider the setting where different organisations are essentially executing the same process while sharing experiences, knowledge, and potentially a common infrastructure. Consider these organisations share an infrastructure (processes, databases, etc.). On the other hand, they are not forced to follow a strict process modelas the system can be configured to support variants of the same process. Process mining is used in best practice analyses, besand t in class analyses.

### C9: Providing operational support opportunity.

Initially, the focus of process mining was on the analysis of historical data. Today, many data sources are updated in (near) real-time and sufficient computing power is available to analyse events when they occur. That's why process mining is now possible in online operational support. Three operational support activities can be identified: ***detect***, ***predict***, and ***recommend***. The moment a case deviates from the predefined process, this can be detected and the system can generate an alert. Often one would like to generate such notifications immediately (to still

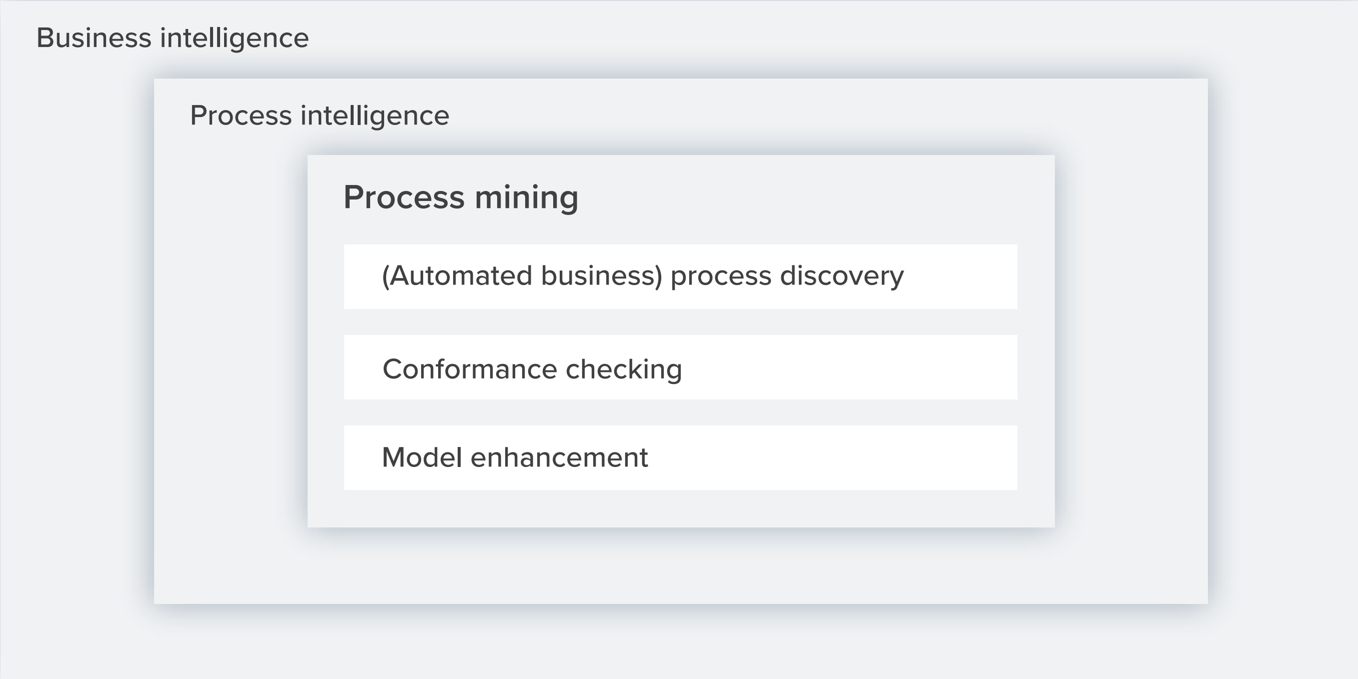



*Figure 6: Relating the different terms.*

be able to influence the next action to be taken) and not in an offline fashion. Historical data can be used to build predictive models and learning data stores. These can be used to guide running process instances. For example, it is possible to predict the remaining processing time and cost of a case leaving the stakeholder options to proceed, or not, to the next activity.

Based on such alarms and predictions, one can foresee recommendations to be made that propose actions. Actions that would reduce costs or shorten the flow time. Applying process mining techniques in such a real-time setting requires a structured approach and appropriate computational infrastructure. (Liquid Algorithm Set-Up) The need for data quality monitoring and logging of the decisions taken, for contextual follow-up of the process executions (explanations for concept drift), is advised.

Practice learns that storing the analysed and tagged historic event chains suits service claim handling, allowing root-cause analyses.

## C10: Combining process mining with other types of analysis

Operations management, and operations research, are branches of management science that heavily rely on modelling.

A variety of mathematical models ranging from linear programming and project planning to queueing models, Markov chains, and simulation are used. Data mining can be defined as "the analysis of (often large) data sets to find unsuspected relationships between attributes (dependent and independent variables) and to combine and summarise the data in novel ways that are both understandable and useful to the data owner".

A wide variety of techniques have been developed: classification (e.g., decision tree learning), regression, clustering (e.g., k- means clustering) and pattern discovery (e.g., association rule learning).

Both fields (operations management and data mining) provide valuable analysis techniques. The opportunity is to combine the techniques with process mining.

## Terminology

The following terms are used in the process mining space: workflow mining, (business) process mining, automated (business) process discovery, and (business) process intelligence. Different organisations seem to use different terms for overlapping concepts.

For example, Gartner is promoting the term "Automated Business Process Discovery" (ABPD). The term "workflow mining" seems less suitable as the creation of workflow models is just one of the many possible applications of process mining.

Similarly, the addition of the term "business" narrows the scope to certain applications of process mining. There are numerous applications of process mining (e.g., analysing the use of high-tech systems or analysing websites) where this addition seems to be inappropriate. Figure 7 relates some of the terms just mentioned.

All technologies and methods that aim at providing actionable information that can be used to support decision-making can be positioned under the umbrella of Business Intelligence (BI). (Business) process intelligence is the combination of BI and BPM, i.e., BI techniques are used to analyse and improve processes and their management. Process mining is a concretisation of process intelligence taking event logs as a starting point. (Automated business) process discovery is just one of the three basic types of process mining. Figure 7 may be a bit misleading in the sense that most BI tools do not provide process mining functionality as described in this document. The term BI is often conveniently skewed towards a tool or method covering only a small part of the broader BI spectrum.

There may be commercial reasons for using alternative terms. Some vendors may also want to emphasise an aspect (e.g., discovery or intelligence). However, to avoid confusion, it is better to use the term "process mining" for the discipline covered by this document.

## Glossary

**Activity**: a well-defined step in the process. Events may refer to the start, completion, cancellation, etc. of an activity for a specific process instance.

**Automated Business Process Discovery**: see Process Discovery.

**Business Intelligence** (BI): broad collection of tools and methods that use data to support decision-making.

**Business Process Intelligence**: see Process Intelligence.

**Business Process Management** (BPM): the discipline that combines knowledge from information technology and knowledge from management sciences and applies both to operational business processes.

**Case**: see Process Instance.

**Concept Drift**: the phenomenon that processes often change over time. The observed process may gradually (or suddenly) change due to seasonal changes or increased competition, thus complicating analysis.

**Conformance Checking**: analysing whether reality, as recorded in a log, conforms to the model and vice versa. The goal is to detect discrepancies and measure their severity. Conformance checking is one of the three basic types of process mining.

**Cross-Organisational Process Mining**: the application of process mining techniques to event logs originating from different organisations.

**Data Mining**: the analysis of (often large) data sets to find unexpected relationships and summarise the data in ways that provide new insights.

**Event**: an action recorded in the log, e.g., the start, completion, or cancellation of an activity for a process instance.

**Event Log**: a collection of events used as input for process mining. Events do not need to be stored in a separate log file (e.g., events may be scattered over different database tables).

**Fitness**: a measure determining how well a given model allows for the behaviour seen in the event log. A model has a perfect fitness if all traces in the log can be replayed by the model from beginning to end.

**Generalisation**: a measure determining how well the model can allow for unseen behaviour. An "overfitting" model is not able to generalise enough.

**Model Enhancement**: one of the three basic types of process mining. A process model is extended or improved using information extracted from some log. For example, bottlenecks can be identified by replaying an event log on a process model while examining the timestamps.

**MXML**: an XML-based format for exchanging event logs. XES replaces MXML as the new tool-independent process mining format.

**Operational Support**: online analysis of event data to monitor and influence running process instances. Three operational support activities can be identified: *detect* (generate an alert if the observed behaviour deviates from the modelled behaviour), *predict* (predict future behaviour based on past behaviour, e.g., predict the remaining processing time), and *recommend* (suggest appropriate actions to realise a goal, e.g., to minimise costs).

**Precision**: measure determining whether the model prohibits behaviour very different from the behaviour seen in the event log. A model with low precision is "underfitting".

**Model Enhancement**: one of the three basic types of process mining. A process model is extended or improved using information extracted from some log. For example, bottlenecks can be identified by replaying an event log on a process model while examining the timestamps.

**Process Discovery**: one of the three basic types of process mining. Based on an event log a process model is learned. For example, the α algorithm can discover a Petri net by identifying process patterns in collections of events.

**Process Instance**: the entity being handled by the process that is analysed. Events refer to process instances. Examples of process instances are customer orders, insurance claims, loan applications, etc.

**Process Intelligence**: a branch of Business Intelligence focusing on Business Process Management.

**Precision**: measure determining whether the model prohibits behaviour very different from the behaviour seen in the event log. A model with low precision is "underfitting".

**Process Mining**: techniques, tools, and methods to discover, monitor and improve real processes (i.e., not assumed processes) by extracting knowledge from event logs commonly available in today's (information) systems.

**Representational Bias**: the selected target language for presenting and constructing process mining results.

**Simplicity**: a measure operationalising Occam's Razor, i.e., the simplest model that can explain the behaviour seen in the log, is the best model. Simplicity can be quantified in various ways, e.g., the number of nodes and arcs in the model.

**XES**: is an XML-based standard for event logs. The standard has been adopted by the IEEE Task Force on Process Mining as the default interchange format for event logs (cf. www.xes-standard.org).

## Interested?

Please contact:

**WALTER VANHERLE**
**Partner**
BDO Advisory | Digital
Process & Data Intelligence
Analytics & RPAutomation
**E-mail:** walter.vanherle@bdo.be
**Tél.:** +32 (0)497 59 62 26

▶ Follow us

▶ www.bdo.be