**Process Mining：**事件序列（Time Series）还原成流程图（Graph Relations），计算节点与路径的关系。

**Extracting, Filtering, and Cleaning Event Data - 数据处理**

事件日志：

* 文件格式：XES / CSV / OCEL
* 中间结果：pandas DataFrame / event log object / ocel
* 数据模型：CaseID（实例） / ObjectID（对象）、Activity（活动）、Timestamp（时间戳）、alternative attributes（可选属性）

**Object-Centric Event Logs:** Traditional event logs, used by mainstream process mining techniques, require the events to be related to a **case**. A case is a set of events for a particular purpose. A **case notion** is a criteria to assign a case to the events.However, in real processes this leads to two problems:

* If we consider the Order-to-Cash process, an order could be related to many different deliveries. If we consider the delivery as case notion, the same event of Create Order needs to be replicated in different cases (all the deliveries involving the order). This is called the **convergence** problem.
* If we consider the Order-to-Cash process, an order could contain different order items, each one with a different lifecycle. If we consider the order as case notion, several instances of the activities for the single items may be contained in the case, and this make the frequency/performance annotation of the process problematic. This is called the **divergence** problem.

**Object-centric event logs** relax the assumption that an event is related to exactly one case. Indeed, an event can be related to different **objects** of different **object types**.

**Process discovery - 流程镜像**

* 发现算法：Alpha Miner、Inductive Miner、Heuristic Miner、Correlation Miner
* 流程结果：Business Process Modeling Notation (BPMN)、Directly Follows Graph (DFG)、Petri Net、Process Tree

**Filter - 筛选**

* Performance(耗时)

1. 发生时间段：events, traces\_contained, traces\_intersecting
2. 持续时间长度

* 活动节点：起点、终点
* 路径模式(variants)：A variant is a set of cases that share the same control-flow perspective, so a set of cases that share the same classified events (activities) in the same order.

1. Top-N
2. Coverage Percentage
3. 路径长度、路径耗时
4. directly follows relation：直接指向
5. eventually follows relation：最终指向
6. prefixes、suffixes：前置/后缀路径
7. activity occurs at least min\_occurrences times.

* 可选属性：按case或event维度筛选

**Conformance checking - 遵从性诊断**

* 推演算法：

1. Token-based replay（不推荐）
2. Alignments
3. Log skeleton
4. Temporal profile

**Simulation - 模拟预测**

* 模型：Monte Carlo Simulation

**推荐书籍**

1. 《Process Mining: Data Science in Action》——Wil van der Aalst