


# Advanced Process Mining

Prof. Dr. Agnes Koschmider

## Lecture 6: Predictive Process Monitoring



- 0 Organization and Introduction
- I Process Discovery
- II Process Conformance
-  III Predictive Process Monitoring
- IV Event Log Preparation
- V Practical Tasks

## Definition of Business Process Monitoring

- **Business Process Monitoring** is the process-oriented monitoring of the most important or critical business processes of a company
- It includes the observation of all technical and business application-specific functions that are required for a smooth and reliable flow of the business processes

## Goals of Business Process Monitoring

- **To detect critical situations as early as possible**
- **To enable the customer's solution support organization to respond to and to solve problems as fast as possible**

- Proactive Monitoring: Activities are performed to prevent critical situations before they occur
- Reactive Monitoring: Once an error has occurred and detected, additional monitoring activities are required to analyse and solve the problem
- Automatic Monitoring: An automatic monitor triggers an alert immediately after an exceptional situation has occurred

**How likely is it that a running process will become “deviant”?**

**Will it end up in a negative outcome?**

**Will it fail to meet its SLAs in the next 24 hours?**

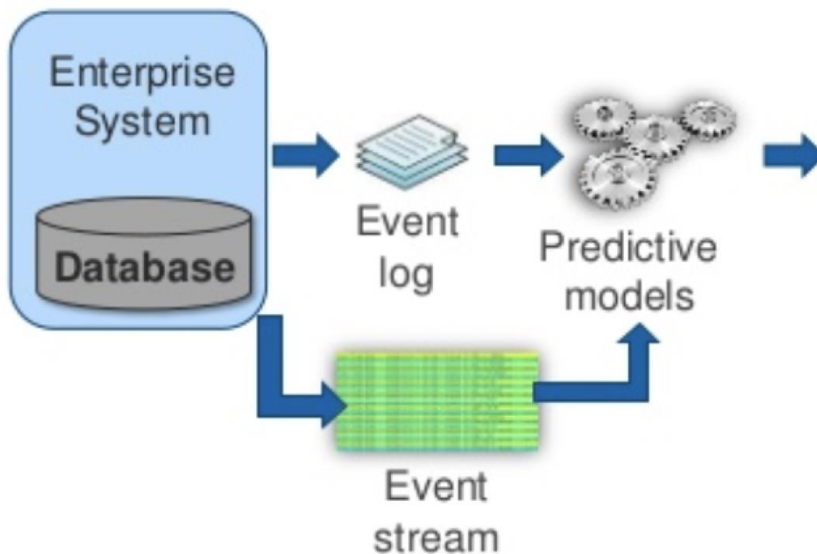
**Will it generate abnormal effort, costs or rework?**

# Business Process Monitoring



## C | A | U

Christian-Albrechts-Universität zu Kiel



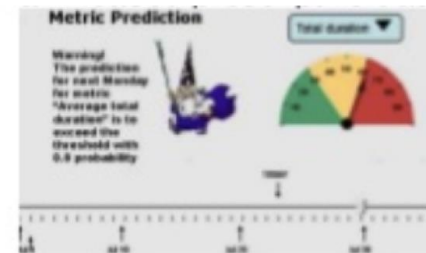
## Aggregate predictive dashboards



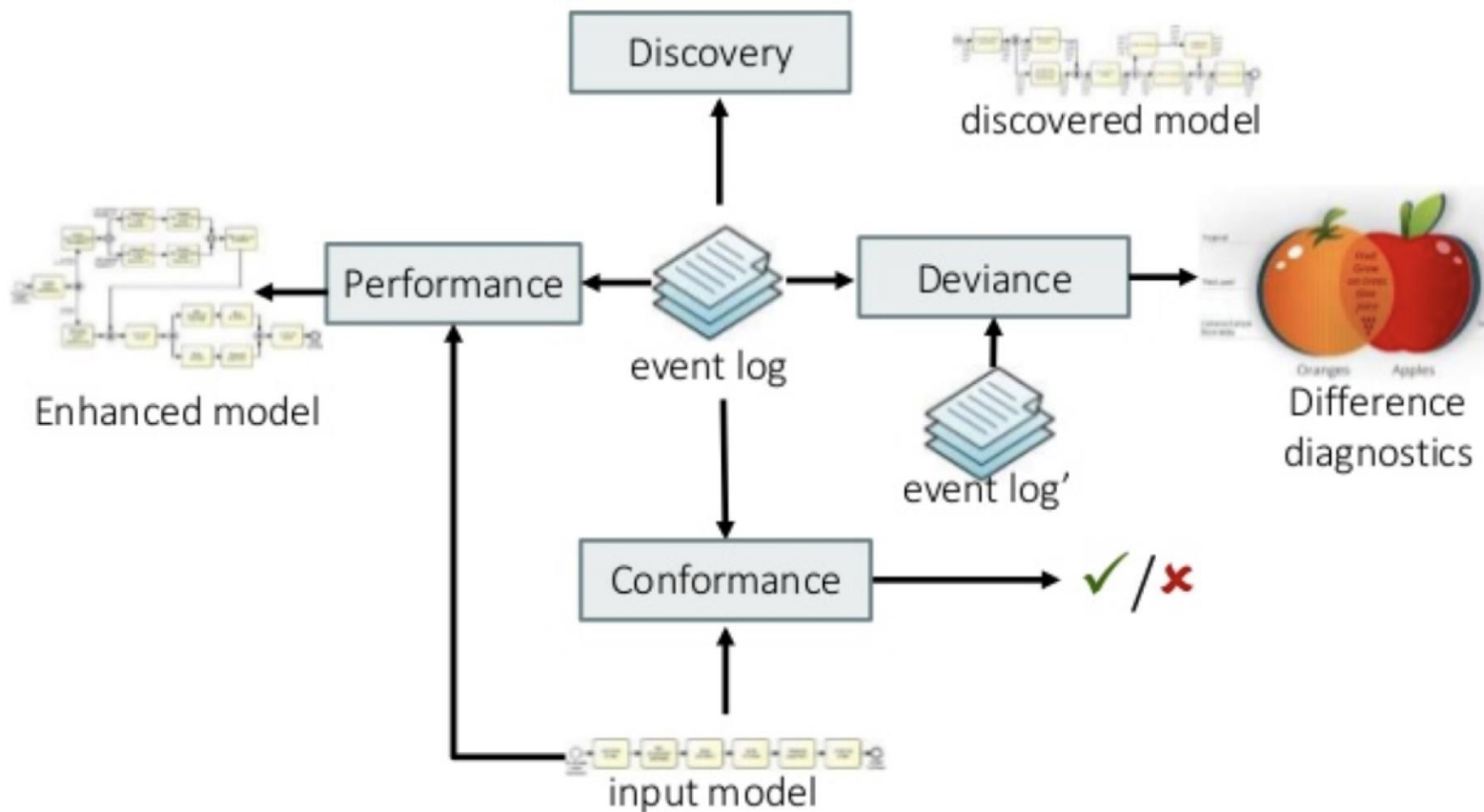
## Detailed predictive dashboard



## Alarm -based prescriptive dashboard



# Offline Process Mining



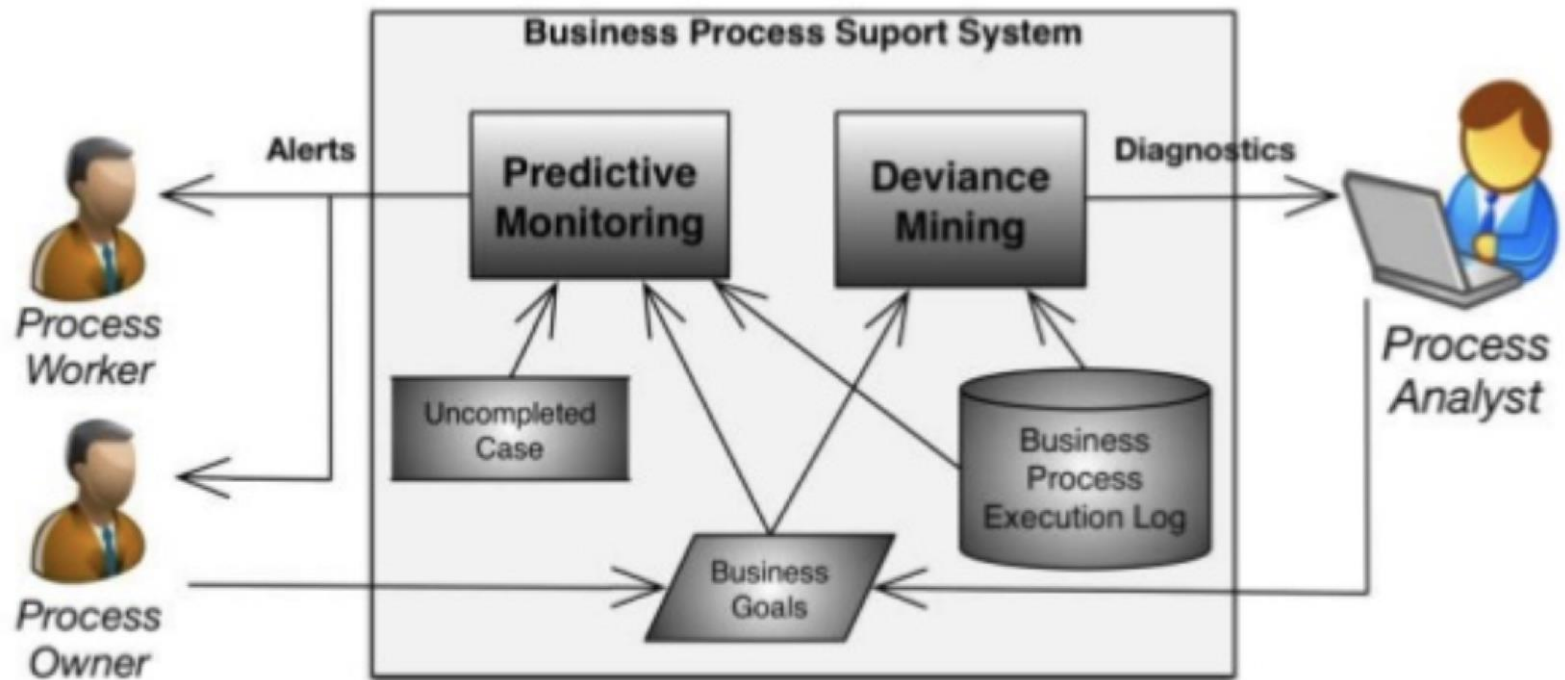


# Deviation Mining via Sequence Classification

- Apply discriminative sequence mining methods to extract features characteristic of one class
- Build classification models (e.g. decision trees)
- Extract difference diagnostics from classification model

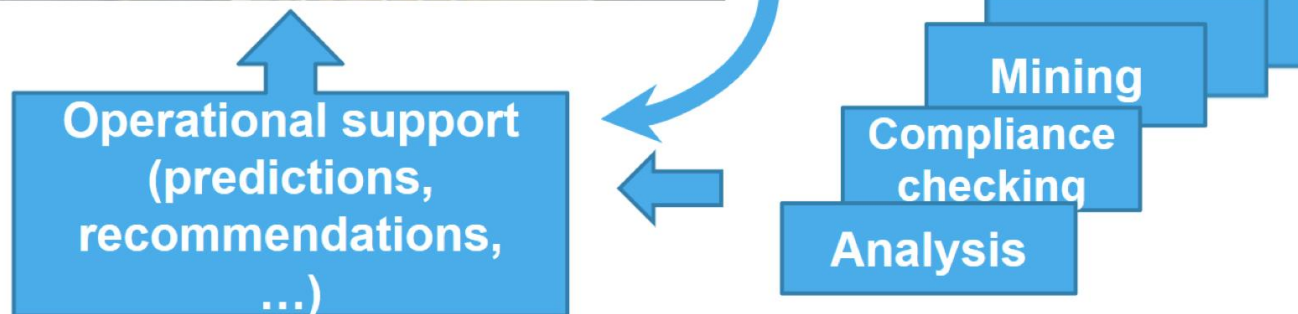
| Rank | Rule   |
|------|--|
| 1    | $\{(Open, 2):anomalous$                                    |
| 2    | $\{(Closed, 2), (Postponed, 0), (Finished, 0)\}:anomalous$ |
| 3    | $\{(Reopen, 2)\}:anomalous$                                |
| 4    | $\{(Closed, 1), (Rejected, 1), (Reopen, 0)\}:anomalous$    |
| 5    | $\{(Reopen, Closed, 1)\}:anomalous$                        |

# Deviance Mining and Predictive Monitoring



# Classical Setting for Business Process Monitoring(1)

**The classical setting for process mining:  
monitoring business processes**



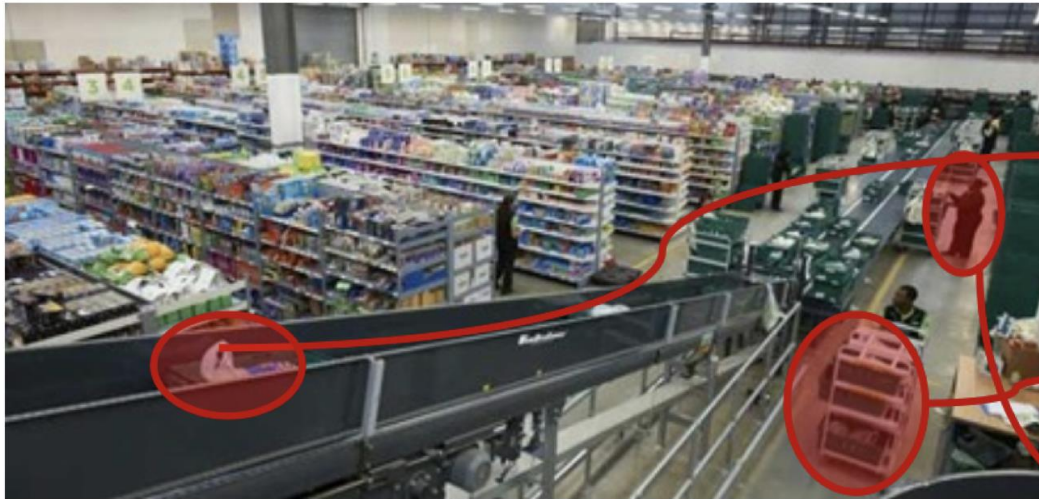
# Classical Setting for Business Process Monitoring(2)



We can tell at every moment:

- where each package is in the process,
- what is in each trolley,
- who is doing what

# Classical Setting for Business Process Monitoring(3)



We can tell at every moment:

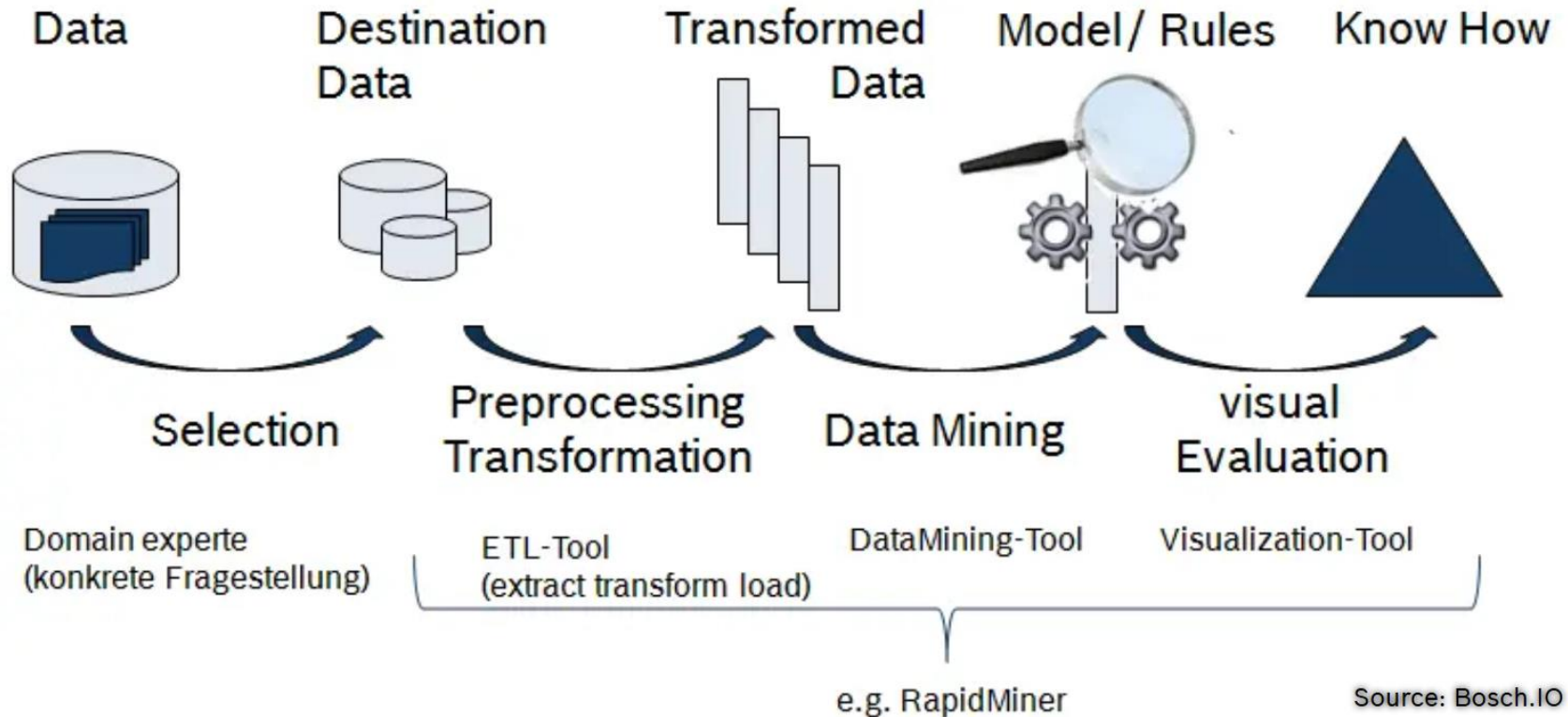
- where each package is in the process,
- what is in each trolley,
- who is doing what

We can apply process mining to

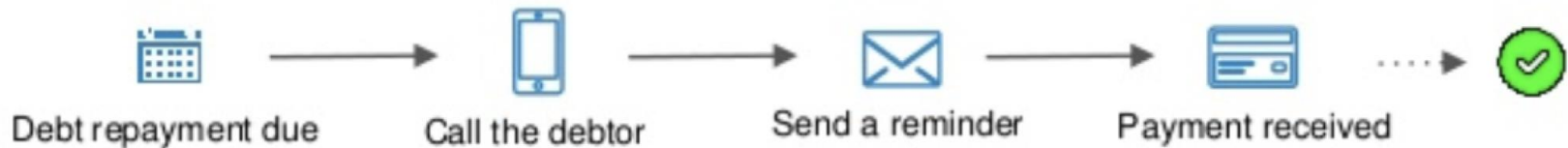
- understand the process,
- check compliance,
- generate predictions,
- ...



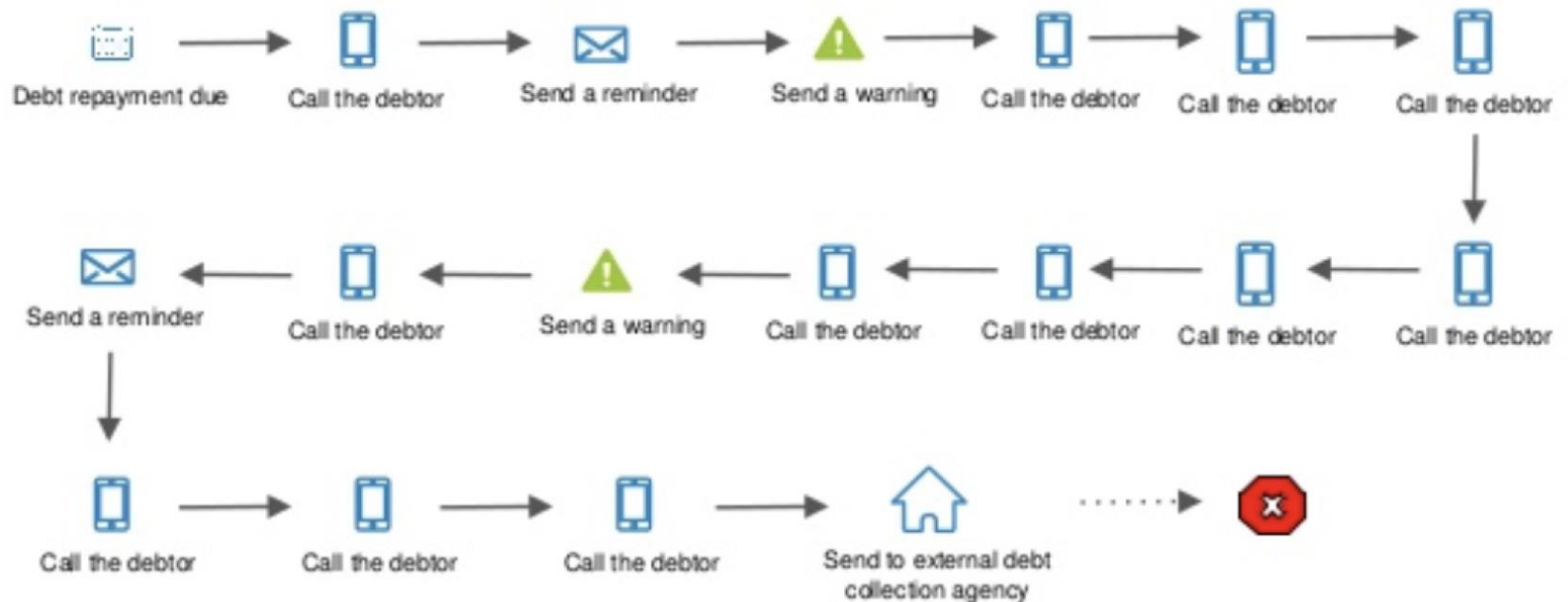
# Data Analysis and Predictive Monitoring



# Predictive Monitoring Example: Debt Recovery Process (1)



# Predictive Monitoring Example: Debt Recovery Process (2)



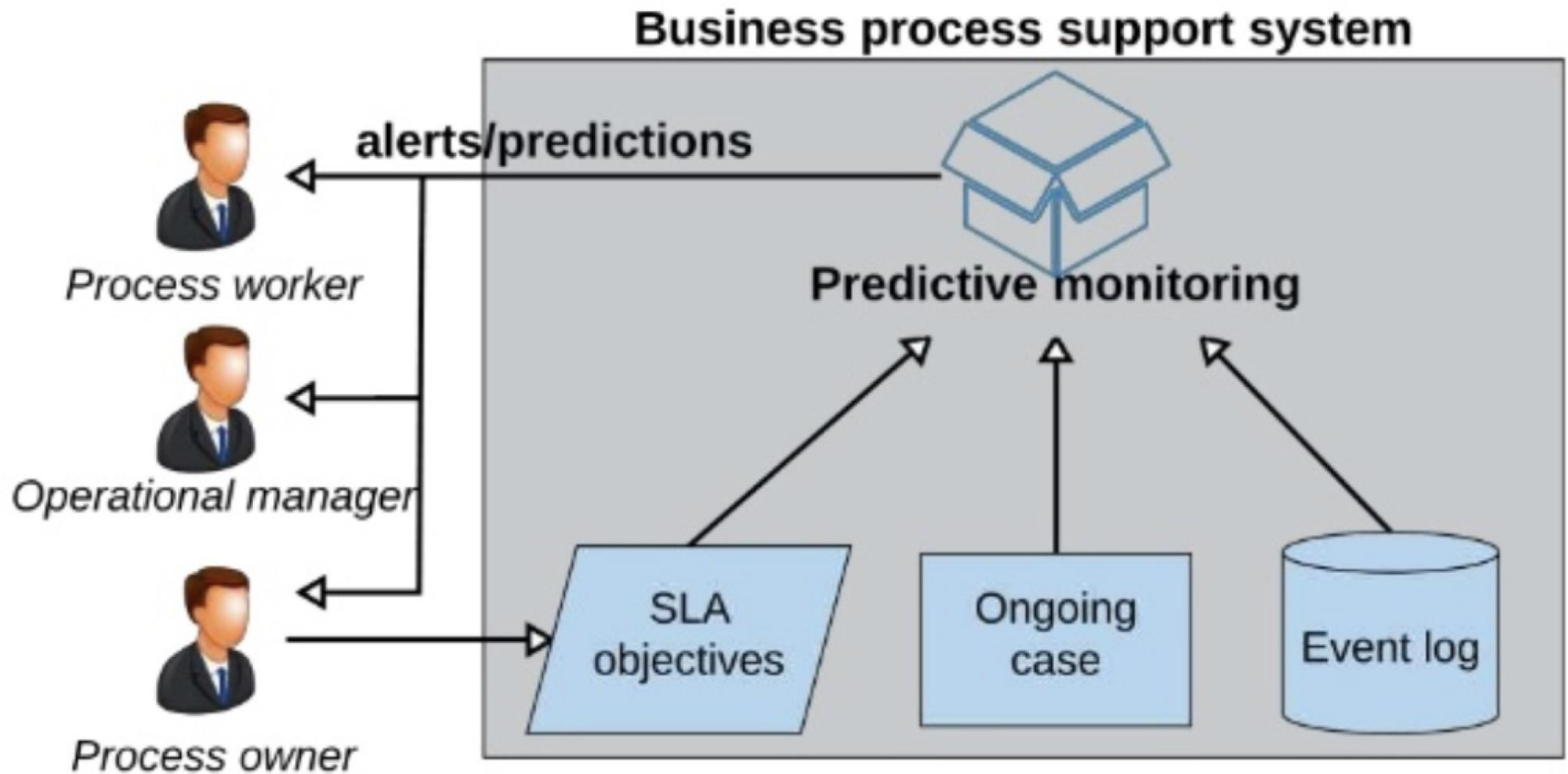


# Time is important

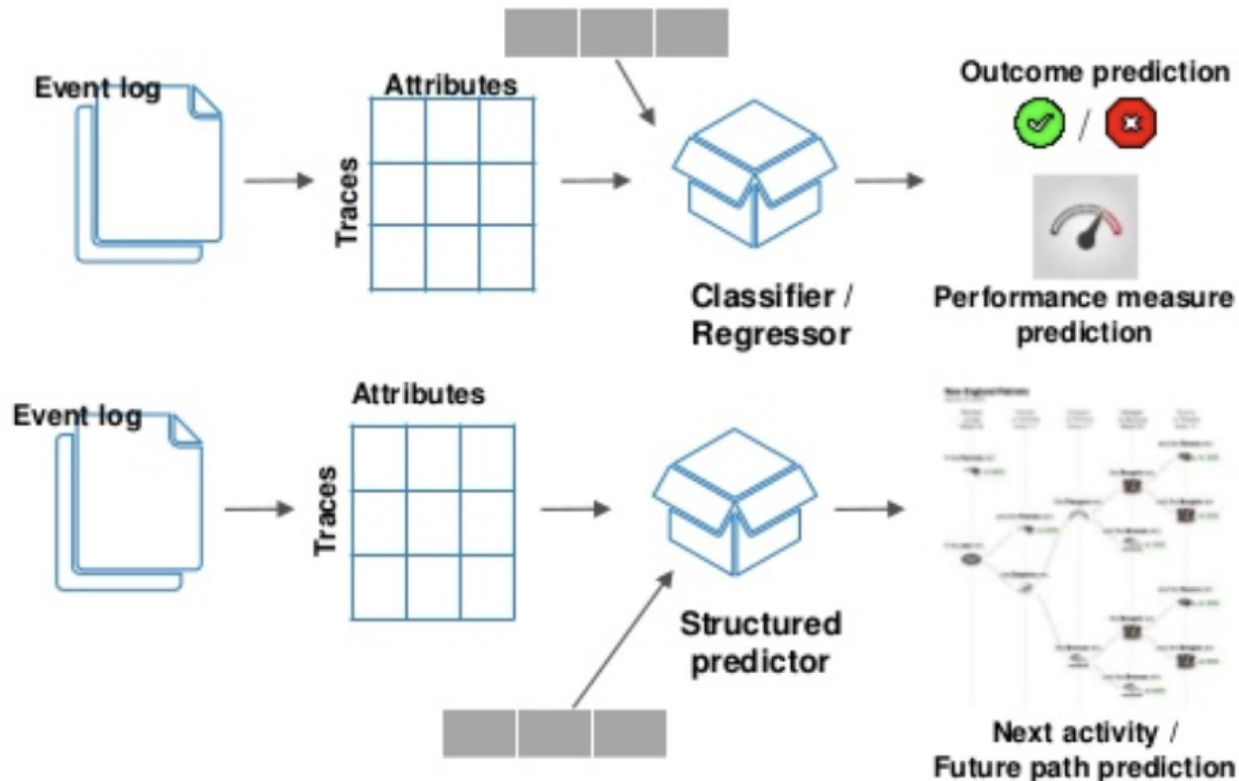


- **Absolute time and its abstractions:**
  - Between 9 pm and 6 am
  - On Sundays
  - ...
- **Relative time**
  - Time from the last feeding
  - Time from falling asleep
- **But also combinations of the two:**
  - Time from the last feeding till 6am

# Predictive Business Process Monitoring



## Predictive Process Monitoring: General Approach



# Predictive Monitoring Example

| Case id | Event id | Properties       |                    |          |      |     |
|---------|----------|------------------|--------------------|----------|------|-----|
|         |          | Timestamp        | Activity           | Resource | Cost | ... |
| 1       | 35654423 | 30-12-2010:11.02 | register request   | Pete     | 50   | ... |
|         | 35654424 | 31-12-2010:10.06 | examine thoroughly | Sue      | 400  | ... |
|         | 35654425 | 05-01-2011:15.12 | check ticket       | Mike     | 100  | ... |
|         | 35654426 | 06-01-2011:11.18 | decide             | Sara     | 200  | ... |
|         | 35654427 | 07-01-2011:14.24 | reject request     | Pete     | 200  | ... |
| 2       | 35654483 | 30-12-2010:11.32 | register request   | Mike     | 50   | ... |
|         | 35654485 | 30-12-2010:12.12 | check ticket       | Mike     | 100  | ... |
|         | 35654487 | 30-12-2010:14.16 | examine casually   | Pete     | 400  | ... |
|         | 35654488 | 05-01-2011:11.22 | decide             | Sara     | 200  | ... |
|         | 35654489 | 08-01-2011:12.05 | pay compensation   | Ellen    | 200  | ... |
| 3       | 35654521 | 30-12-2010:14.32 | register request   | Pete     | 50   | ... |
|         | 35654522 | 30-12-2010:15.06 | examine casually   | Mike     | 400  | ... |
|         | 35654524 | 30-12-2010:16.34 | check ticket       | Ellen    | 100  | ... |
|         | 35654525 | 06-01-2011:09.18 | decide             | Sara     | 200  | ... |
|         | 35654526 | 06-01-2011:12.18 | reinitiate request | Sara     | 200  | ... |
|         | 35654527 | 06-01-2011:13.06 | examine thoroughly | Sean     | 400  | ... |
|         | 35654530 | 08-01-2011:11.43 | check ticket       | Pete     | 100  | ... |
|         | 35654531 | 09-01-2011:09.55 | decide             | Sara     | 200  | ... |
|         | 35654533 | 15-01-2011:10.45 | pay compensation   | Ellen    | 200  | ... |

- What is the next activity for this case?

- When is this next activity going to take place?

- How long is this case still going to take until it is finished?

- What is the outcome of this case? Is the compensation going to be paid? Or rejected?

← Current situation

# Sequence encoding(1)

| Case ID | Timestamp  | Activity                | Resource | Loan goal | Requested amt | Offered amt |
|---------|------------|-------------------------|----------|-----------|---------------|-------------|
| C001    | 18-10-2016 | Check completeness      | Sue      | Mortgage  | 100 000       | -           |
| C001    | 19-10-2016 | Check credit history    | Sue      | Mortgage  | 100 000       | -           |
| C001    | 19-10-2016 | Calculate risk score    | Bob      | Mortgage  | 100 000       | -           |
| C001    | 20-10-2016 | Make offer              | Mike     | Mortgage  | 100 000       | 70 000      |
| C001    | 25-10-2016 | Make offer              | Mike     | Mortgage  | 100 000       | 80 000      |
| C002    | 20-10-2016 | Check completeness      | Sue      | Car       | 15 000        | -           |
| C002    | 20-10-2016 | Check credit history    | Sue      | Car       | 15 000        | -           |
| C002    | 22-10-2016 | Calculate risk score    | Elsa     | Car       | 15 000        | -           |
| C002    | 24-10-2016 | Reject application      | Elsa     | Car       | 15 000        | -           |
| C003    | 02-11-2016 | Check completeness      | Maria    | Mortgage  | 30 000        | -           |
| C003    | 04-11-2016 | Ask for additional data | Maria    | Mortgage  | 30 000        | -           |
| C003    | 10-11-2016 | Check credit history    | Maria    | Mortgage  | 30 000        | -           |
| ...     | ...        | ...                     | ...      | ...       | ...           | ...         |

Feature vector  $x$

Target  $y$

|       |   |   |   |     |   |   |   |     |    |   |   |   |     |   |
|-------|---|---|---|-----|---|---|---|-----|----|---|---|---|-----|---|
| $x^0$ | 1 | 0 | 0 | ... | 1 | 0 | 0 | ... | 14 | 0 | 0 | 0 | ... | 1 |
| $x^1$ | 1 | 0 | 0 | ... | 0 | 1 | 0 | ... | 15 | 1 | 0 | 0 | ... | 1 |
| $x^2$ | 0 | 1 | 0 | ... | 0 | 1 | 0 | ... | 18 | 0 | 0 | 0 | ... | 0 |
| $x^3$ | 0 | 0 | 1 | ... | 0 | 1 | 0 | ... | 10 | 0 | 0 | 0 | ... | 1 |
| $x^4$ | 0 | 0 | 1 | ... | 0 | 0 | 1 | ... | 17 | 0 | 1 | 0 | ... | 0 |

# Sequence encoding(2)

## ▷ Index-based encoding

| id | loan goal | activity_1         | ... | activity_m           | amount_1 | ... | amount_m | label |
|----|-----------|--------------------|-----|----------------------|----------|-----|----------|-------|
| 1  | mortgage  | check completeness | ... | make offer           | 0        | ... | 80 000   | 1     |
| 2  | car       | check completeness | ... | calculate risk score | 0        | ... | 0        | 0     |

## ▷ Aggregation encoding

| id | loan goal | count(check compl-ness) | ... | count(make offer) | min(amt) | ... | max(amt) | label |
|----|-----------|-------------------------|-----|-------------------|----------|-----|----------|-------|
| 1  | mortgage  | 1                       | ... | 1                 | 0        | ... | 80 000   | 1     |
| 2  | car       | 1                       | ... | 0                 | 0        | ... | 0        | 0     |

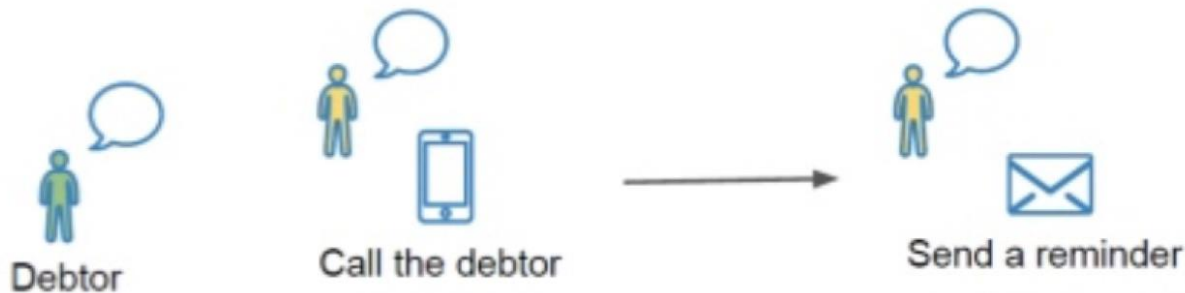
## ▷ LSTM

| id | event | loan goal | event                | amount | label |
|----|-------|-----------|----------------------|--------|-------|
| 1  | 1     | mortgage  | check completeness   | 0      | 1     |
|    | ...   | ...       | ...                  | ...    |       |
|    | m     | mortgage  | make offer           | 80 000 |       |
| 2  | 1     | car       | check completeness   | 0      | 0     |
|    | ...   | ...       | ...                  | ...    |       |
|    | m     | car       | calculate risk score | 0      |       |

| Case ID | Timestamp  | Activity                | Resource | Loan goal | Requested amt | Offered amt |
|---------|------------|-------------------------|----------|-----------|---------------|-------------|
| C001    | 18-10-2016 | Check completeness      | Sue      | Mortgage  | 100 000       | -           |
| C001    | 19-10-2016 | Check credit history    | Sue      | Mortgage  | 100 000       | -           |
| C001    | 19-10-2016 | Calculate risk score    | Bob      | Mortgage  | 100 000       | -           |
| C001    | 20-10-2016 | Make offer              | Mike     | Mortgage  | 100 000       | 70 000      |
| C001    | 25-10-2016 | Make offer              | Mike     | Mortgage  | 100 000       | 80 000      |
| C002    | 20-10-2016 | Check completeness      | Sue      | Car       | 15 000        | -           |
| C002    | 20-10-2016 | Check credit history    | Sue      | Car       | 15 000        | -           |
| C002    | 22-10-2016 | Calculate risk score    | Elsa     | Car       | 15 000        | -           |
| C002    | 24-10-2016 | Reject application      | Elsa     | Car       | 15 000        | -           |
| C003    | 02-11-2016 | Check completeness      | Maria    | Mortgage  | 30 000        | -           |
| C003    | 04-11-2016 | Ask for additional data | Maria    | Mortgage  | 30 000        | -           |
| C003    | 10-11-2016 | Check credit history    | Maria    | Mortgage  | 30 000        | -           |
| -       | -          | -                       | -        | -         | -             | -           |



# Predictive Monitoring with Unstructured Data



|        | Event1          | Event2          | Resource1 | Resource2 | Debtor | Summary1 | Summary2 |
|--------|-----------------|-----------------|-----------|-----------|--------|----------|----------|
| Trace1 | Call the debtor | Send a reminder | Sue       | Bob       | Mark   | ?        | ?        |

# Text Mining

The last ten years has seen a surge of interest in design science research in information systems, organizations, process modelling and software engineering. In this talk I present a framework for design science that shows how in design science research, we iterate over designing new artifacts for a context, and empirically investigating these artifacts in this context. To be relevant, the artifacts should potentially contribute to organizational goals, and to be empirically sound, research to validate new artifacts should provide insight into the effects of using these artifacts in an organizational context. The logic of both of these activities, design and empirical research, is that of rational decision making. I show how this logic can be used to structure our technical and empirical research goals and questions, as well as how to structure reports about our technical or empirical research. This gives us checklists for the design cycle used in technical research and for the empirical cycle used in empirical research. Finally, I will discuss in more detail what the role of theories in design science research is, and how we use theory to state research questions and to generalize the research results.

→ 0.2, 0.1, 0.8, 0.5, ..., 0.1

The tutorial first introduces the PPM including its activities: problem understanding, method finding, modeling, reconciliation, and validation.

→ 0.4, 0.8, 1.0, 0.2, ..., 0.4

What is a good business process model and how do you get value from it? We have for many years worked with SEQUAL, a general framework for understanding the quality of models and modelling languages, which covers all main aspects relative to quality of models. The framework has been widely cited since the first version was presented in the nineties, and the tutorial will focus on the most recent version of the framework (2016), specialised for quality of business process models, with a focus on how to achieve value through long-term usage of business process models in organizations.

→ 0.9, 0.0, 0.4, 0.5, ..., 0.2

Business process models have gained significant importance due to their critical role for managing business processes. Process models not only play a fundamental role for obtaining a common understanding of an organization's business processes, but are also important assets for improving business processes and to support the development of information systems. In this tutorial we will focus on the process of process modeling (PPM) and shed light on how process models are created.

→ 0.2, 0.3, 0.7, 0.6, ..., 0.6



# Text-Extended Index-Based Encoding



|        | Event1          | Event2          | Resource1 | Resource2 | Debtor | $t_{11}$ | ... | $t_{1n}$ | $t_{21}$ | ... | $t_{2n}$ |
|--------|-----------------|-----------------|-----------|-----------|--------|----------|-----|----------|----------|-----|----------|
| Trace1 | Call the debtor | Send a reminder | Sue       | Bob       | Mark   | 0.2      | ... | 0.1      | 0.4      | ... | 0.4      |

- Bag-of-N-grams
- Weighted bag-of-N-grams
- Latent Dirichlet Allocation (LDA)
- Paragraph Vector (PV)

# LSTM-Based Predictive Process Monitoring

