Christian-Albrechts-Universität zu Kiel

# Advanced Process Mining

Prof. Dr. Agnes Koschmider

**Lecture 1: Introduction**

# Ad Persona

Christian-Albrechts-Universität zu Kiel

Prof. Dr. Agnes Koschmider

Since 05/2019:
Professor of Information Systems (Process Analytics)

Education:
Habilitation, Applied Informatics, KIT
Promotion, Applied Informatics, KIT

TU/e

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

ALPEN-ADRIA UNIVERSITÄT
KLAGENFURT I WIEN GRAZ

Almaden Institute
IBM

# Organization

Christian-Albrechts-Universität zu Kiel

- **Lecture**
  - Thursday, 10:15 – 11:45 h, HRS3 - R.218b
  - will be offered as online course
- **Exercise**

  weekly, start April 16.


**Dominik Janssen**

| | |
|---|---|
| room | Hermann-Rodewald-Str. 3, R. 505 |
| consultation hours | Tue, 10:00 - 11:00 h |
| eMail | dominik.janssen@informatik.uni-kiel.de |

# Organization

Christian-Albrechts-Universität zu Kiel

- **Lecture recording**
  - There will be a video recording for the lecture

- **Exam:** to be announced

# Organization

Christian-Albrechts-Universität zu Kiel

- **Consultation hours – Prof. Dr. Agnes Koschmider**
  - after the lecture
  - WED, 13:00 – 14:00 h

  (please make an appointment in advance)

  Hermann-Rodewald-Str. 3, R. 504a

- **eMail** ak@informatik.uni-kiel.de
- **Phone** 0431 880-6387

# Teaching Program

Christian-Albrechts-Universität zu Kiel

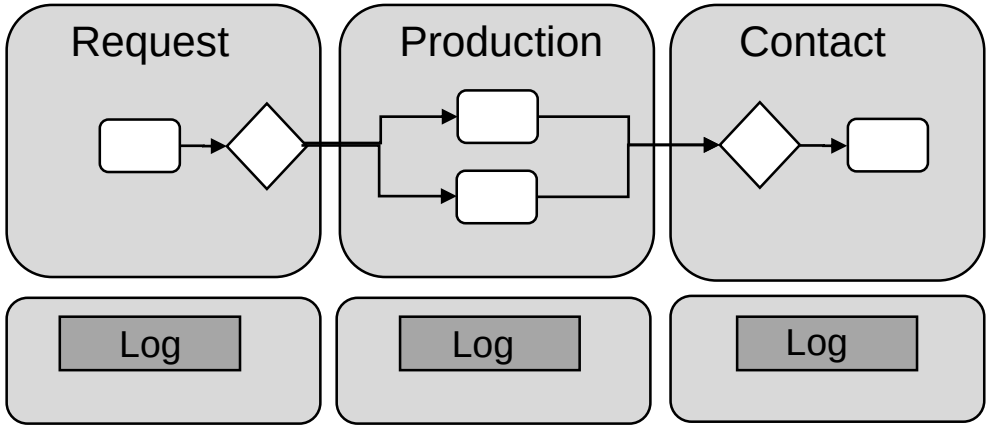
Blockchain &
Robotic Process Automation

Application Domain

Entwicklung digitaler
Plattformen

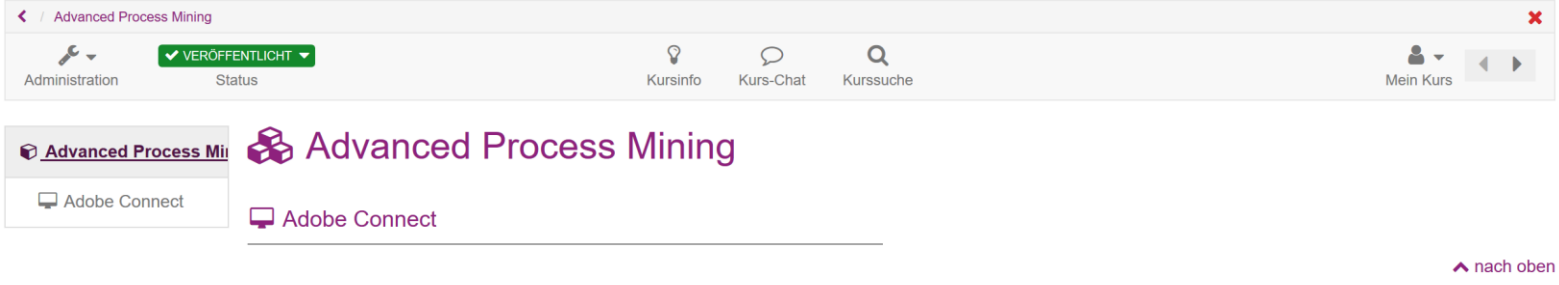
Data Processing

Process Mining 1
&
Process Mining 2



# Services in OLAT

Christian-Albrechts-Universität zu Kiel




- In OLAT you will find all information and announcements for the lecture and exercise
- The forum primarily serves as mutual exchange of students. Assistants will also be happy to answer questions in individual cases and get involved in the discussions

# Lecture Overview

Christian-Albrechts-Universität zu Kiel

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

# Literatur

Christian-Albrechts-Universität zu Kiel


- A. Drescher, A. Koschmider, A. Oberweis: Modellierung und Analyse von Geschäftsprozessen, Grundlagen und Übungsaufgaben mit Lösungen, De Gruyter, 2017. Link
- W. van der Aalst: Process Mining: Data Science in Action. Springer, 2016. Link
- J. Carmona, B. van Dongen, A. Solti, M. Weidlich: Conformance Checking: Relating Processes and Models. Springer, 2018. Link

# Learning Objective Chapter 0

Christian-Albrechts-Universität zu Kiel

- recap of the lecture process mining 1
- understand what process mining is and what it is not
- understand what a business process is
- spot on use cases for process mining

# Processes in Science

Christian-Albrechts-Universität zu Kiel



## Process (science)

From Wikipedia, the free encyclopedia

*For other uses, see Process (disambiguation).*

In science, a **process** is every sequence of changes of a real object/body which is observable using scientific method. Therefore, all sciences analyze and model *processes*.

Processes are always properties of dynamic systems, they are characterized by such system attributes as variables and parameters. Every process model has distinguished input and output variables, it can be autonomous or controlled.

The recognition of a process is an arbitrary subjective mental operation/event because it depends on different circumstances, observer's goal, perception and conceptualization tools.

There are numerous **taxonomies** of processes, roughly speaking, they are divided on: continuous and discrete, stable and not stable, convergent or not convergent, cyclic and not cyclic, linear and not linear, as well as they are grouped according to the name of the domain where they are analyzed.

**Some examples of physical, technological and biological processes**:

combustion, crystallization, centrifugation, diffraction, diffusion, dispersion, distillation, electrolysis, electrophoresis, emulsification, evaporation, hydrolysis, nuclear fission, nuclear fusion, oxidation, phosphorescence, pyrolysis, reduction, reflection, refraction, scattering, sedimentation, sublimation, birth, cell division, fermentation, fertilization, germination, growth, geotropism, heliotropism, hybridization, metamorphosis, photosynthesis, transpiration

# Other Processes



- ***Biological Process***: a process of a living organism
- ***Chemical Process***: a means of changing one or more chemicals or chemical compounds
- ***Mental process***: functions or processes done with the mind
- ***Process in computing***: a computer program or an instance of a program running concurrently with other programs
- …

# What is a Business Process?

Christian-Albrechts-Universität zu Kiel

***Collection of activities***, which are executed

- manually, semi-automatically or automatically,
- according to certain rules

***to reach a business goal.***

# Characteristics of Business Processes

Christian-Albrechts-Universität zu Kiel

- activities are connected to each other via acting persons, machines, document flows, resources, etc.
- activities are executed by persons or by machines in a specific order to perform certain tasks.
- a business process might have *structured*, *weakly structured* and *unstructured* components (*sub-processes*).

# Characteristics of Business Processes

Christian-Albrechts-Universität zu Kiel

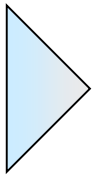
- a business process creates something which is of value for a customer.
- customers might be internal or external to an organization
- a collaborative business process is characterized by the fact that at least two organizations cooperatively execute its activities.
  - **intra-organizational** business process
  - **inter-organizational** business process

# Process Model

Christian-Albrechts-Universität zu Kiel

***Process model***
***(or process definition, process schema)***

- describes the structure of a real business process
- specifies all possible paths along a business process
- specifies the rules for choosing a path
- specifies all activities that must be executed

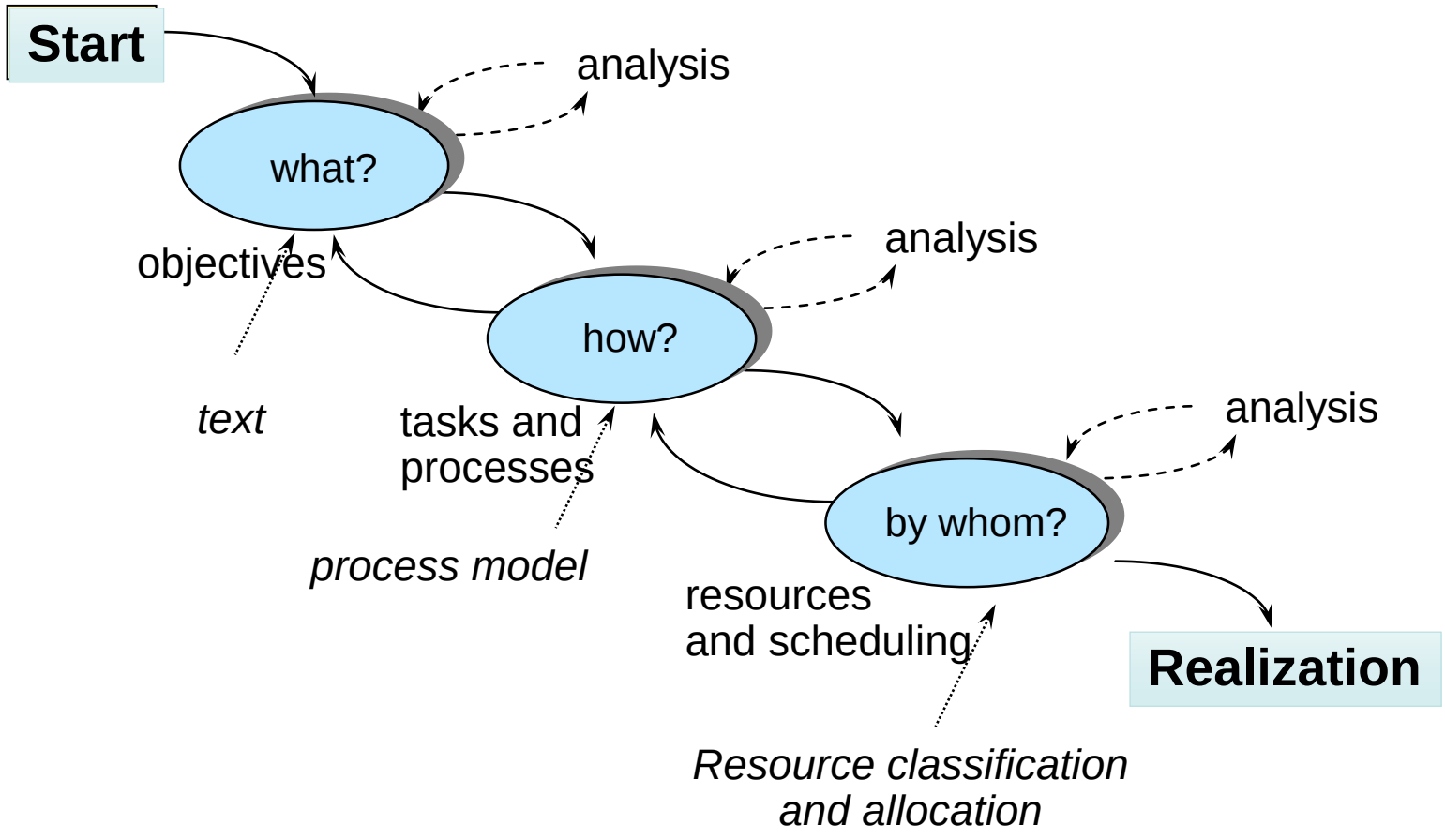


*A process model is a template.*
*Starting from there all process instances are initiated.*

# Process Design

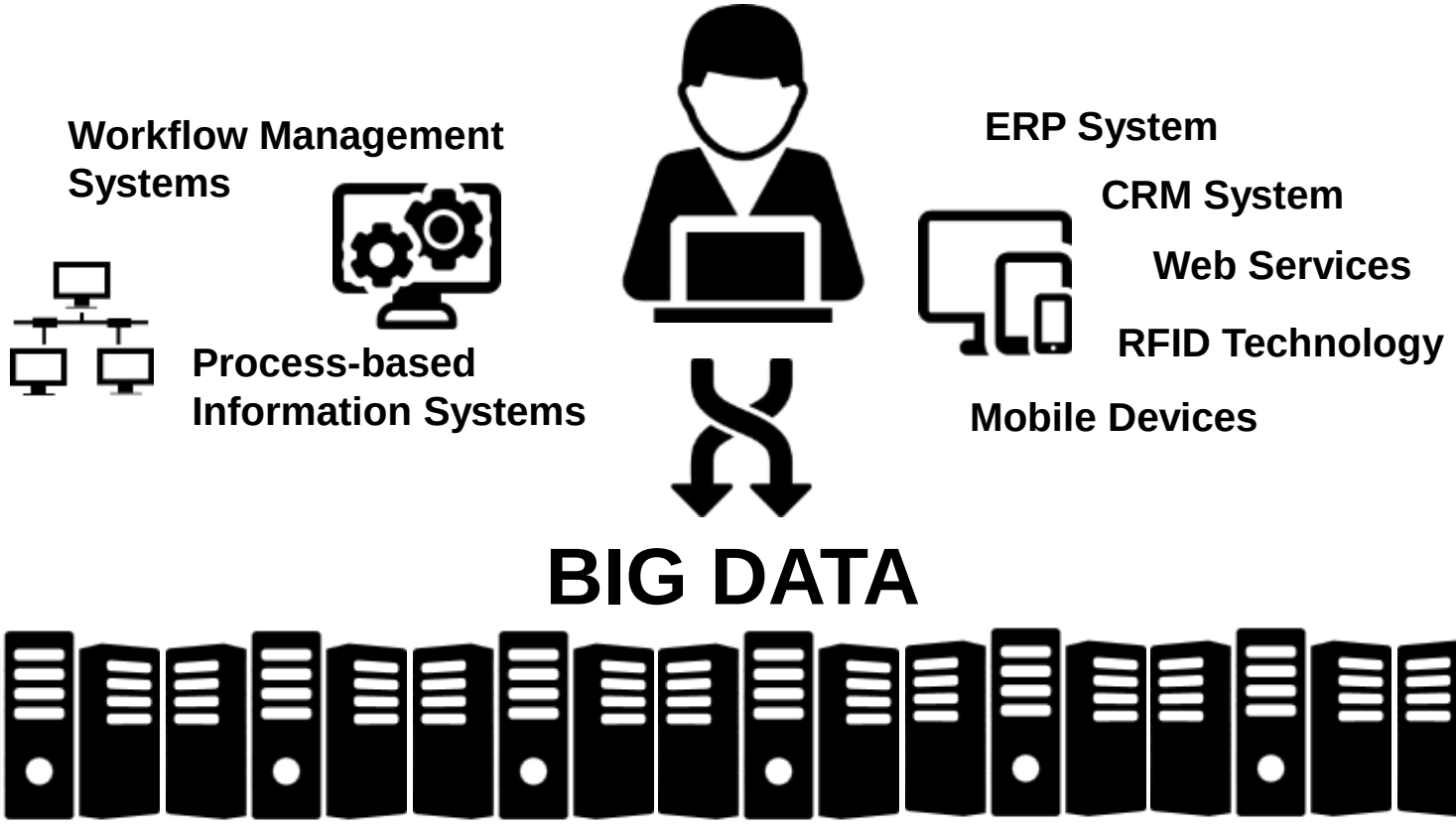
Christian-Albrechts-Universität zu Kiel

**Steps:**



# Process Mining – For What?

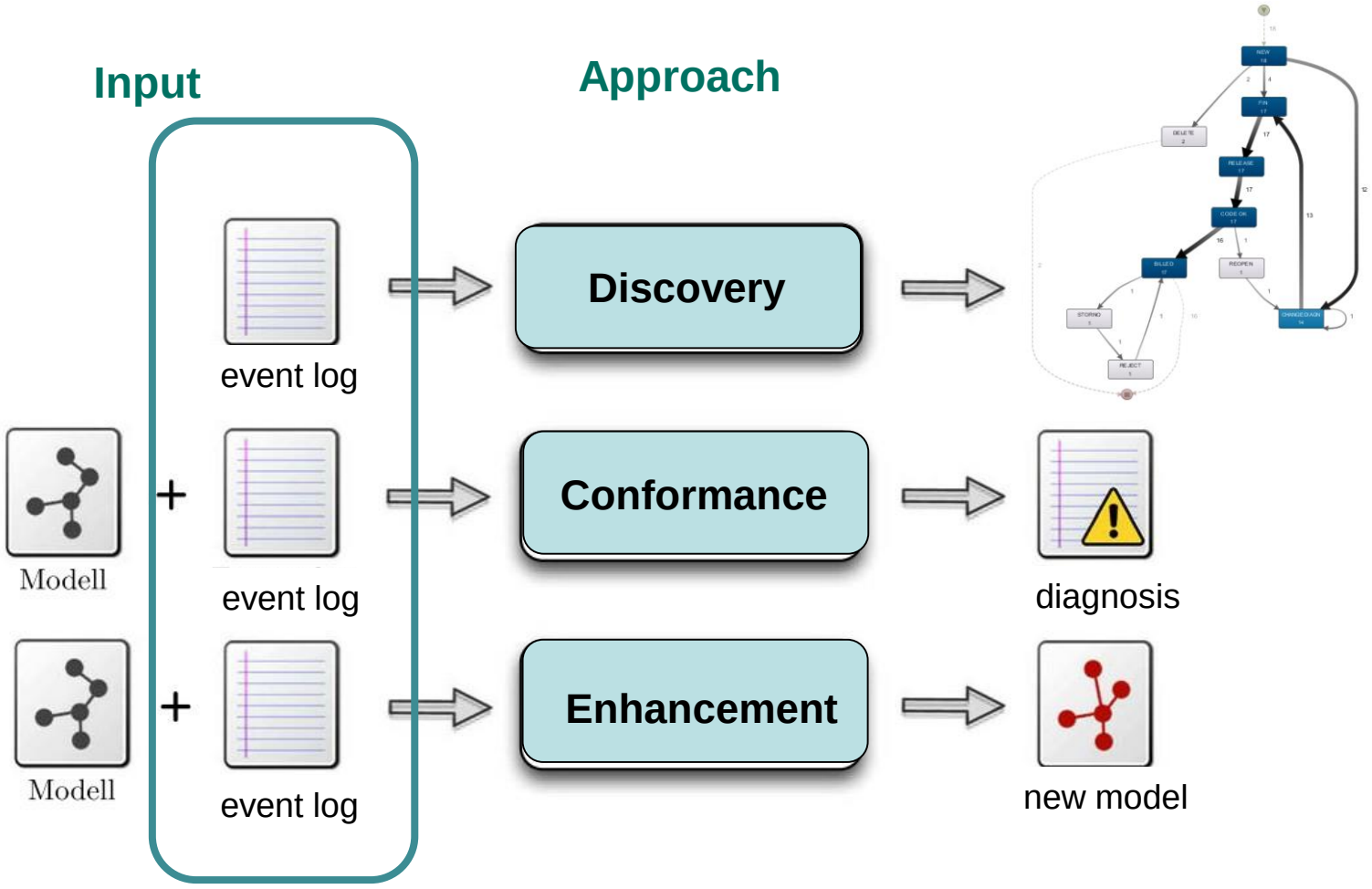
Christian-Albrechts-Universität zu Kiel




- Identify process models from log files

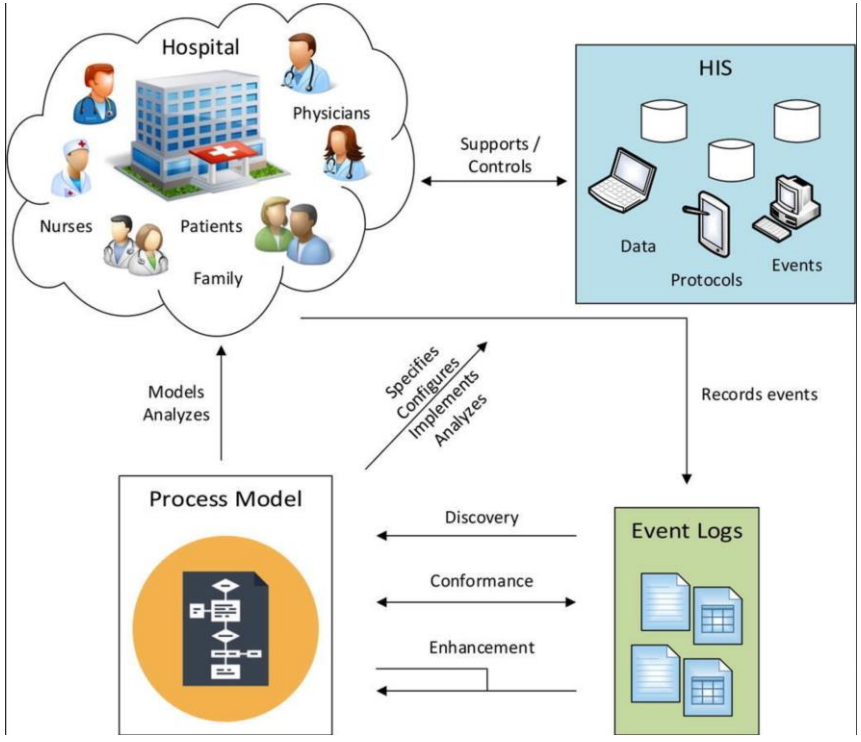
# Applications of Process Mining

Christian-Albrechts-Universität zu Kiel



# Process mining in healthcare

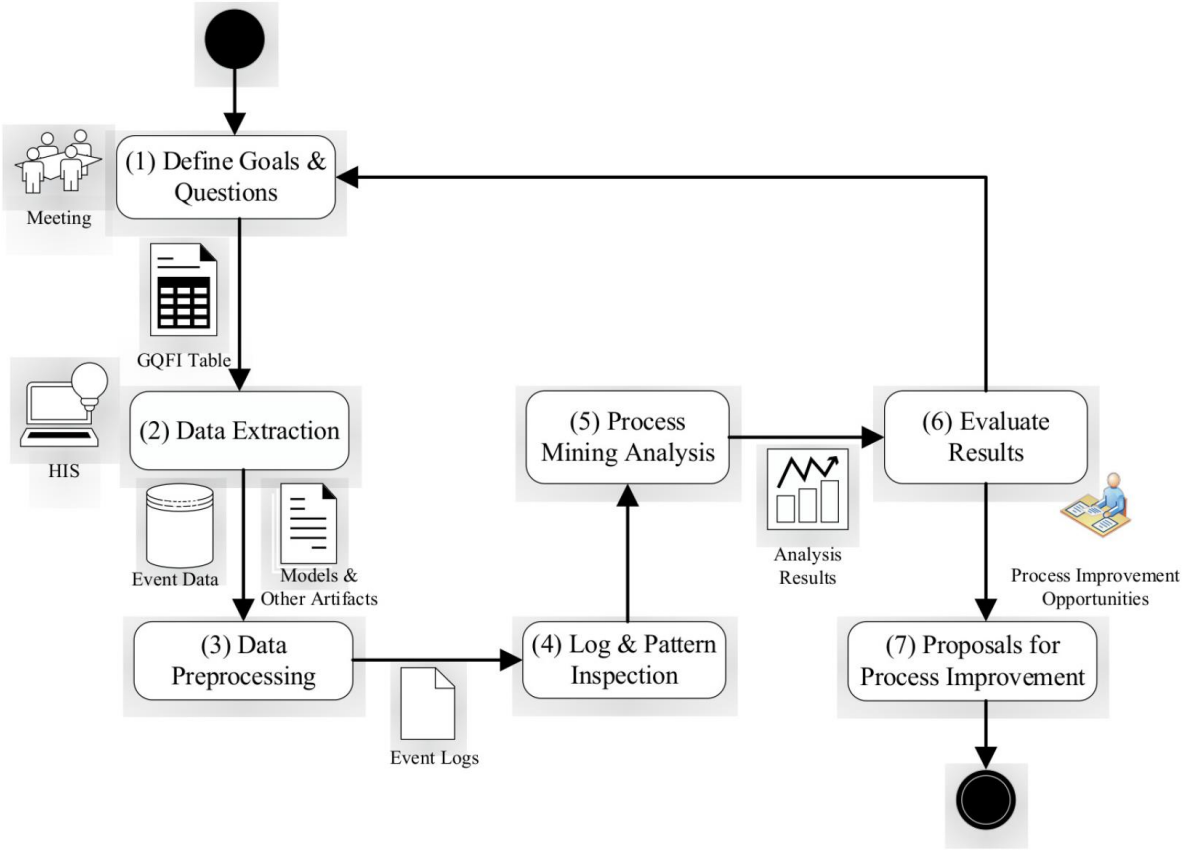
Christian-Albrechts-Universität zu Kiel




source: https://www.sciencedirect.com/science/article/pii/S1532046416300296

# Process mining in healthcare

Christian-Albrechts-Universität zu Kiel




**A Goal-Driven Evaluation Method Based On Process Mining for Healthcare Processes**

# Applications of Process Mining

Christian-Albrechts-Universität zu Kiel

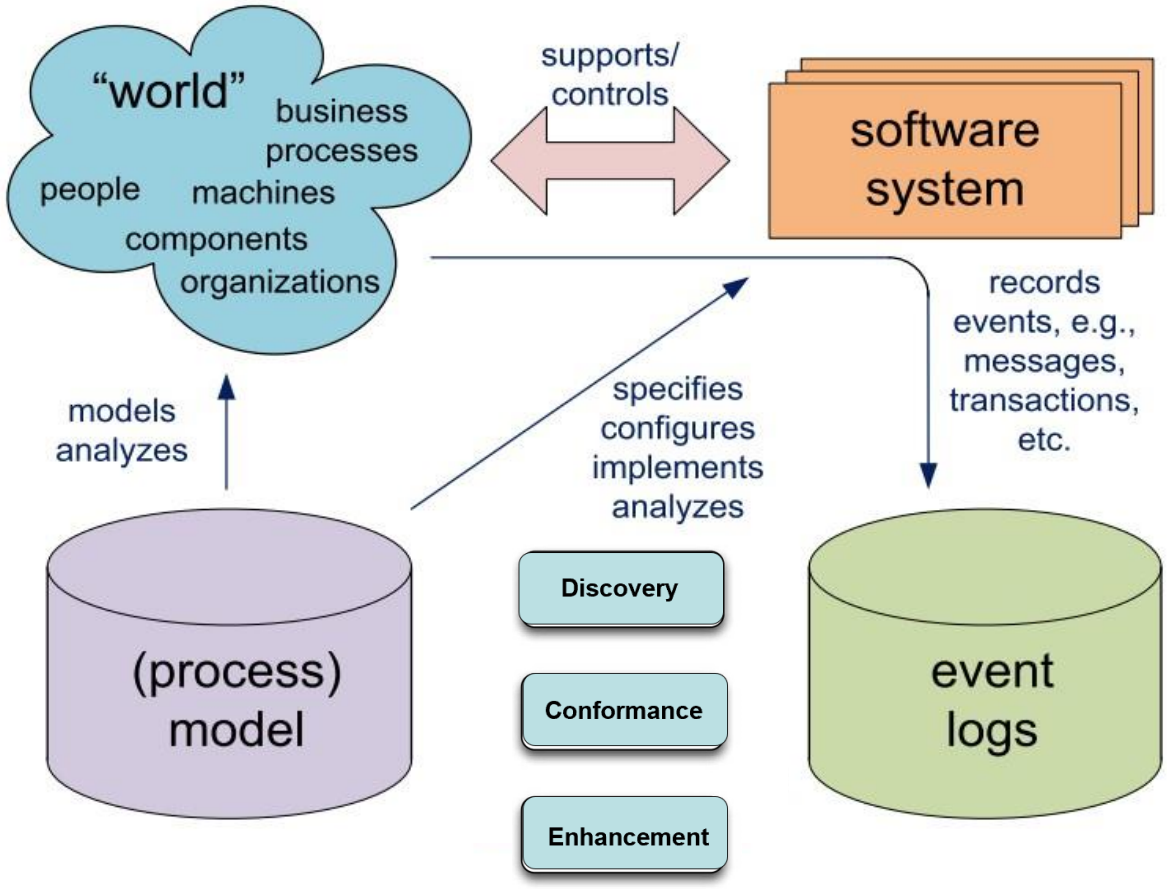



Bild: Van der Aalst, Wil M. P. (2016): *Process Mining – Data Science in Action* (S. 32). Springer, Berlin, Heidelberg.

# What Process Mining Is Not

Christian-Albrechts-Universität zu Kiel

- BI or Reporting Tool
  - Process mining is an analysis tool while BI-dashboards are for monitoring and reporting
  - A process mining analysis *can* result in a new KPI that then should be monitored, but it can also lead to a process change
- Process Modeling Tool
- New Improvement Methodology
- IT Project
- Data Mining, AI, or Machine Learning Tool
- Simulation Tool
- Just for Some Processes or IT Systems
- Magic Bullet

https://fluxicon.com/book/read/positioning/#bi-or-reporting-tool

# Definitions

Christian-Albrechts-Universität zu Kiel

**Event, attribute**

Let Ɛ be the set of all possible events. Let A be a set of attribute names. For each event e ∈ Ɛ and each attribute name a ∈ A ist $\#_a$(e) the value of attribute a for each event e. We denote $\#_a$(e) =⊥ if attribute $a$ for $e$ is not defined

**Event log, case, trace**

- Process mining assumes the existence of an event log where each event refers to a **case**, an activity, and a point in time. An event log can be seen as a **collection of cases** and a case can be seen as a trace/**sequence of events**.

# Definitions

Christian-Albrechts-Universität zu Kiel

- $\#_{Activity}$(e): describes the activity observed in event $e$ in the form of the activity name
- $\#_{Time}$(e): describes the time stamp, which is the time of the observation
- $\#_{Resource}$(e): describes by which user, which machine, or resource (resource), the activity was executed

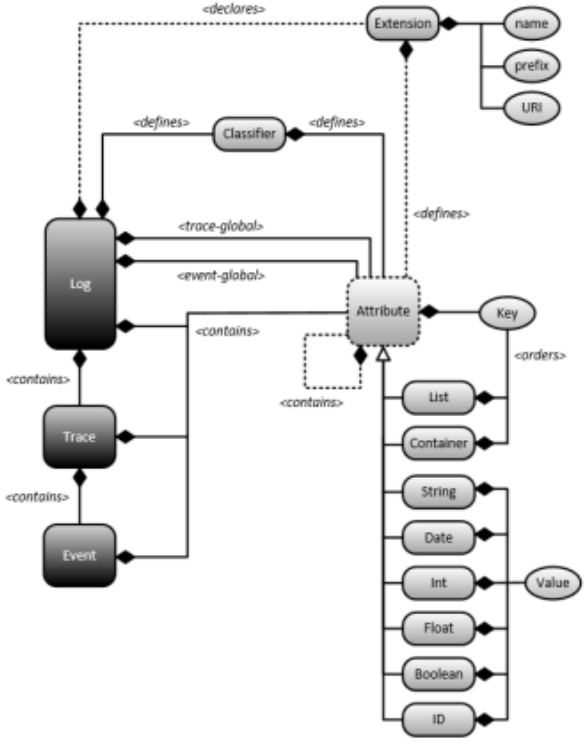
| $e$ | $\pi_a(e)$ | $\pi_t(e)$ |
|---|---|---|
| e1 | register | 1559819400 |
| e2 | register | 1559819460 |
| e3 | select | 1559819590 |
| e4 | select | 1559825300 |
| e5 | pay | 1559825350 |
| e6 | cancel | 1559893020 |

# Event Log – Exchange Format

Christian-Albrechts-Universität zu Kiel

- Format to exchange event logs
  - MXML (Mining eXtensible Markup Language)
  - XES (eXtensible Event Stream) - de facto standard

XES
Extensible Event Stream




- **Log:** represents a process
  - is assigned to
- **Trace:** represents a case
  - is assigned to
- **Event:** represents an activity

Souce: XES Standard Definition, version 2.0, http://www.xes-standard.org/_media/xes/xesstandarddefinition-2.0.pdf

# Event Log – Beispiel im XES-Format

Christian-Albrechts-Universität zu Kiel

```
<log xes.version="1.0" xmlns="http://code.deckfour.org/xes">
  <trace>
    <event>
      <date key="time:timestamp" value="2010-12-30T11:02:00.000+01:00"/>
      <string key="Activity" value="register request"/>
      <string key="Resource" value="Pete"/>
      <string key="Costs" value="50"/>
    </event>
    <event>
      <date key="time:timestamp" value="2010-12-31T10:06:00.000+01:00"/>
      <string key="Activity" value="examine thoroughly"/>
      <string key="Resource" value="Sue"/>
      <string key="Costs" value="400"/>
    </event>
    <event>
      <date key="time:timestamp" value="2011-01-05T15:12:00.000+01:00"/>
      <string key="Activity" value="check ticket"/>
      <string key="Resource" value="Mike"/>
      <string key="Costs" value="100"/>
    </event>
  </trace>
  <trace>
    <event>
      <date key="time:timestamp" value="2011-01-06T15:02:00.000+01:00"/>
      <string key="Activity" value="register request"/>
      <string key="Resource" value="Pete"/>
      <string key="Costs" value="50"/>
    </event>
    <event>
      <date key="time:timestamp" value="2011-01-07T12:06:00.000+01:00"/>
      <string key="Activity" value="check ticket"/>
      <string key="Resource" value="Mike"/>
      <string key="Costs" value="100"/>
    </event>
    <event>
      <date key="time:timestamp" value="2011-01-08T14:43:00.000+01:00"/>
      <string key="Activity" value="examine thoroughly"/>
      <string key="Resource" value="Sean"/>
      <string key="Costs" value="400"/>
    </event>
  </trace>
</log>
```

Log

Trace

Event

Trace

Event

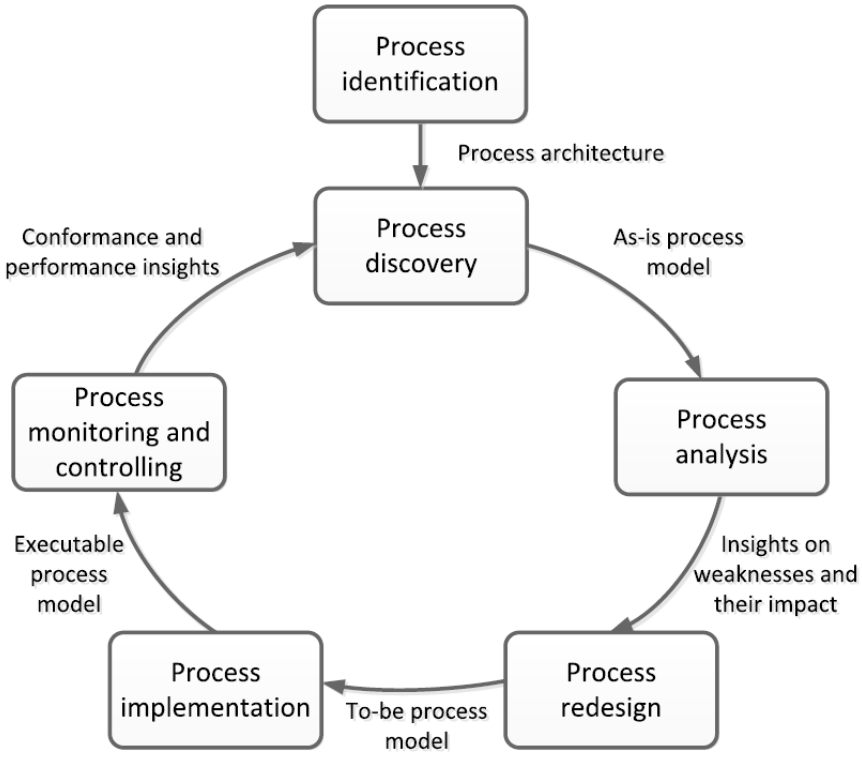
# Simplified Event Log

Christian-Albrechts-Universität zu Kiel

| Patient | Activity | Time | Staff | Further Attributes |
|---|---|---|---|---|
| P1 | A | 0 | S1 | … |
| P1 | B | 5 | S2 | … |
| P1 | C | 5 | S3 | … |
| P2 | A | 15 | S2 | … |
| P2 | B | 7 | S3 | … |
| P2 | C | 10 | S2 | … |
| P3 | A | 15 | S2 | … |
| P3 | B | 7 | S3 | … |
| P4 | A | 10 | S1 | … |
| P4 | D | 10 | S3 | … |
| P5 | A | 0 | S4 | … |

# Business Process Life Cyle (1)

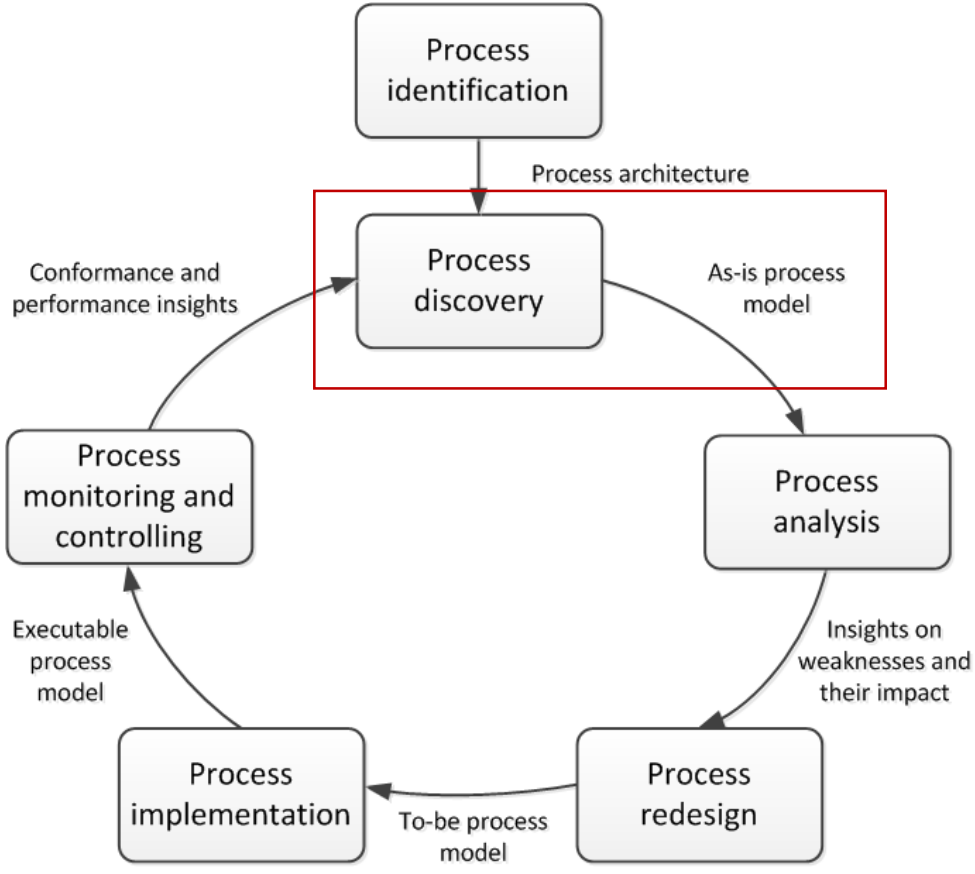
Christian-Albrechts-Universität zu Kiel




M. Dumas M. La Rosa, J. Mendling H.A. Reijers: Fundamentals of Business Process Management, 2013

# Business Process Life Cyle (2)

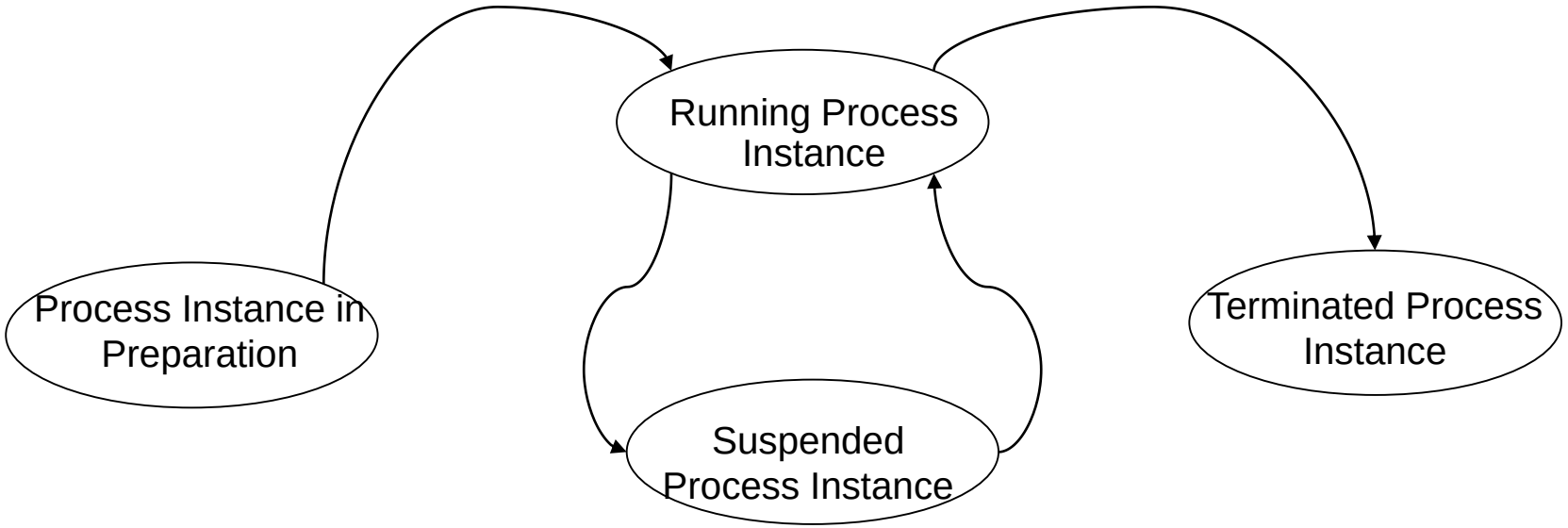




# Business Process Life Cyle (3)

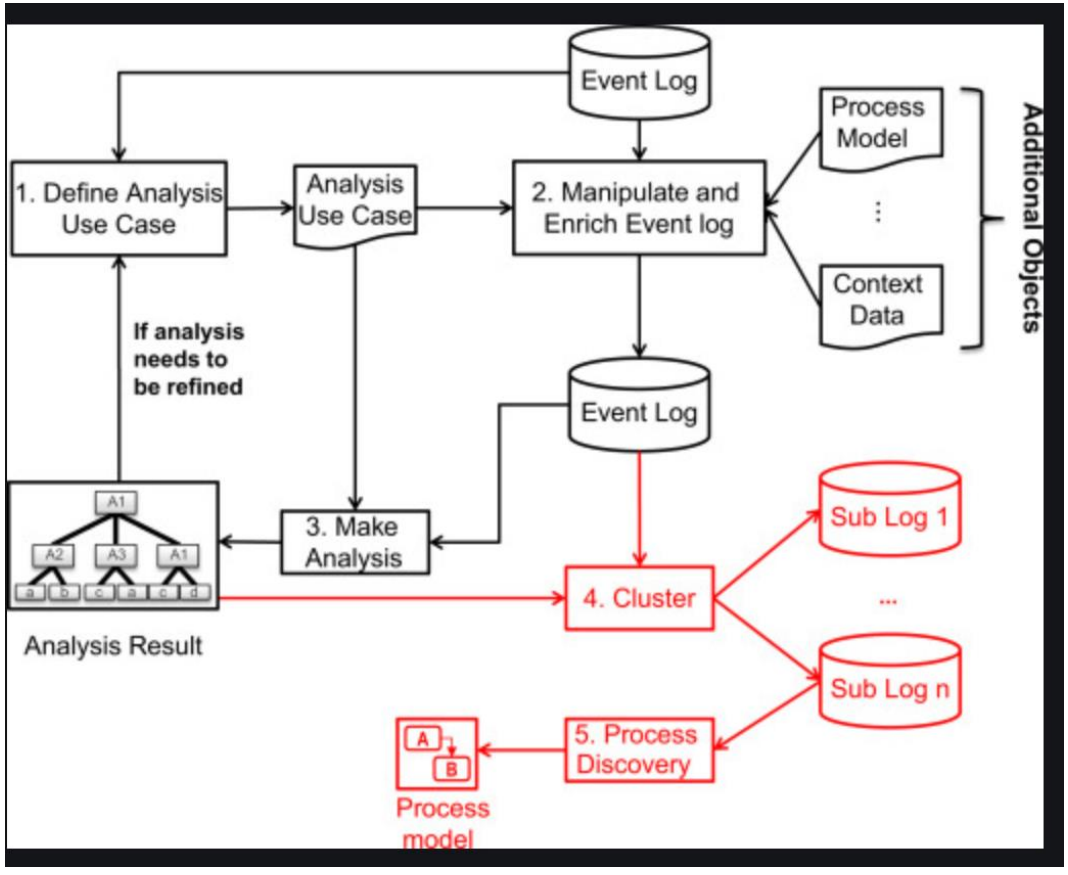
Christian-Albrechts-Universität zu Kiel

**Process Instance**



# Event Log Prepration

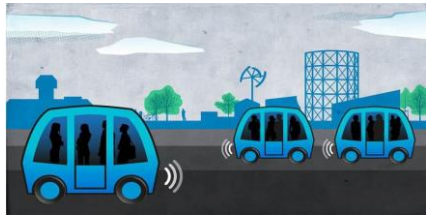
Christian-Albrechts-Universität zu Kiel



# Use Cases

Christian-Albrechts-Universität zu Kiel

Can we predict any changes in the consumer behavior due to new conditions?

Can we foresee any changes in the fleet?

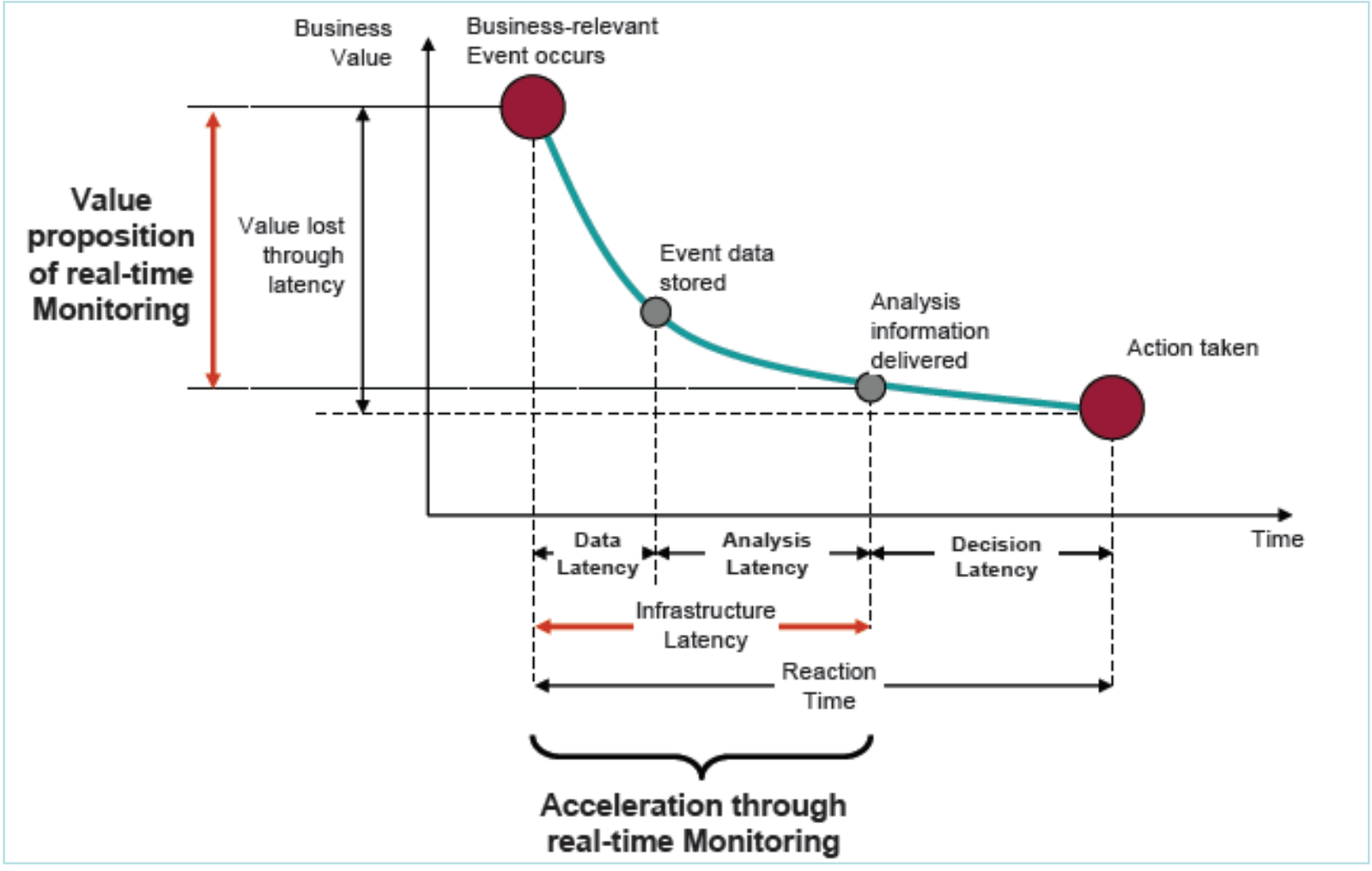
```
6/24/2008 20:20:26 M16-1
6/24/2008 20:20:26 M16-1
6/24/2008 20:20:31 M16-0
6/24/2008 20:20:52 M16-1
6/24/2008 20:20:52 M16-1
6/24/2008 20:20:55 M16-0
6/24/2008 20:20:55 M16-1
6/24/2008 20:20:58 M09-1
6/24/2008 20:20:58 M15-1
6/24/2008 20:20:58 M09-1
6/24/2008 20:20:58 M08-1
6/24/2008 20:20:58 M07-1
6/24/2008 20:20:58 M07-1
6/24/2008 20:21:01 M15-0
6/24/2008 20:21:01 M16-0
6/24/2008 20:21:04 M07-0
6/24/2008 20:21:04 M08-0
6/24/2008 20:21:04 M09-0
6/24/2008 20:21:07 M06-1
6/24/2008 20:21:07 M09-1
6/24/2008 20:21:07 M09-0
6/24/2008 20:21:08 M06-0
6/24/2008 20:21:14 M09-1
6/24/2008 20:21:16 M09-0
6/24/2008 20:21:16 M09-0
6/24/2008 20:21:23 M09-1
6/24/2008 20:21:28 M09-0
6/24/2008 20:21:31 M09-1
6/24/2008 20:21:31 M09-1
6/24/2008 20:21:37 M09-0
6/24/2008 20:21:37 M09-0
6/24/2008 20:21:38 M09-1
6/24/2008 20:21:40 M09-0
6/24/2008 20:21:46 M09-1
6/24/2008 20:21:49 M09-0
```

Can we improve the re-routing of ships due to changes detection in data?


| CaseID | Date | Time | Activity | Role |
|---|---|---|---|---|
| 1231 | 30-12-2018 | 11:02 | price updated | Pete |
| | 31-12-2018 | 10:06 | block released | Sue |
| | 05-01-2017 | 15:12 | billing released | Mike |
| | 06-01-2018 | 11:18 | changed released | Sara |
| | 07-01-2018 | 14:24 | date changed | Pete |
| 1254 | 30-12-2017 | 11:32 | price updated | Mike |
| | 30-12-2017 | 12:12 | billing released | Mike |
| | 30-12-2017 | 14:16 | block released | Pete |
| | 05-01-2018 | 11:22 | changed released | Sara |
| | 08-01-2018 | 12:05 | date changed | Ellen |
| 1698 | 30-12-2017 | 14:32 | price updated | Pete |
| | 30-12-2017 | 15:06 | date changed | Mike |
| | 30-12-2017 | 16:34 | billing released | Ellen |
| | 06-01-2018 | 09:18 | hanged released | Sara |
| | 06-01-2018 | 12:18 | order fullfilled | Sara |
| | 06-01-2018 | 13:06 | block released | Sean |
| | 08-01-2018 | 11:43 | billing released | Pete |
| | 09-01-2018 | 09:55 | hanged released | Sara |
| | 15-01-2018 | 10:45 | payment done | Ellen |

# Monitoring Business Events

Christian-Albrechts-Universität zu Kiel



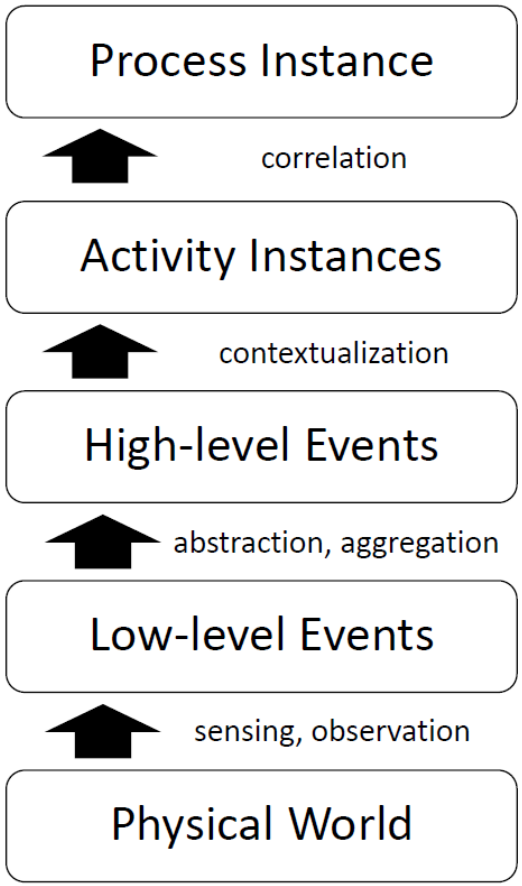
# What are the challenges ?

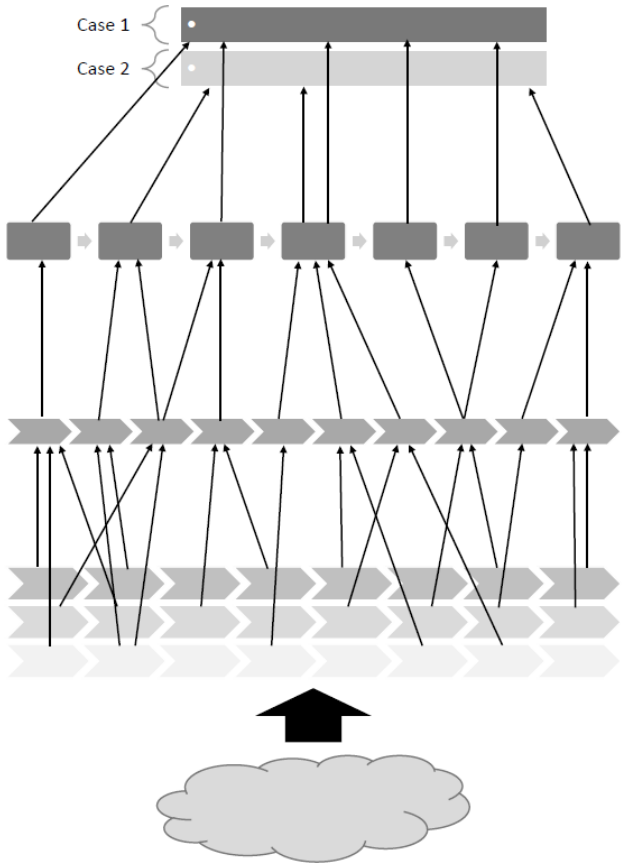
Christian-Albrechts-Universität zu Kiel

without consideration of sensor data changes

correlation challenge

no entity-individualization




A. Koschmider, F. Mannhardt, T. Heuser: On the Contextualization of Event-Activity Mappings, Proceedings of BPM 2018 Workshops, Springer.

S. J. van Zelst·F. Mannhardt,·M. de Leoni, A. Koschmider: Event Abstraction in Process Mining - LiteratureReview and Taxonomy, to appear

# Chapter 0