Process Oriented Data Science

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Campus d'Excel·lència Internacional

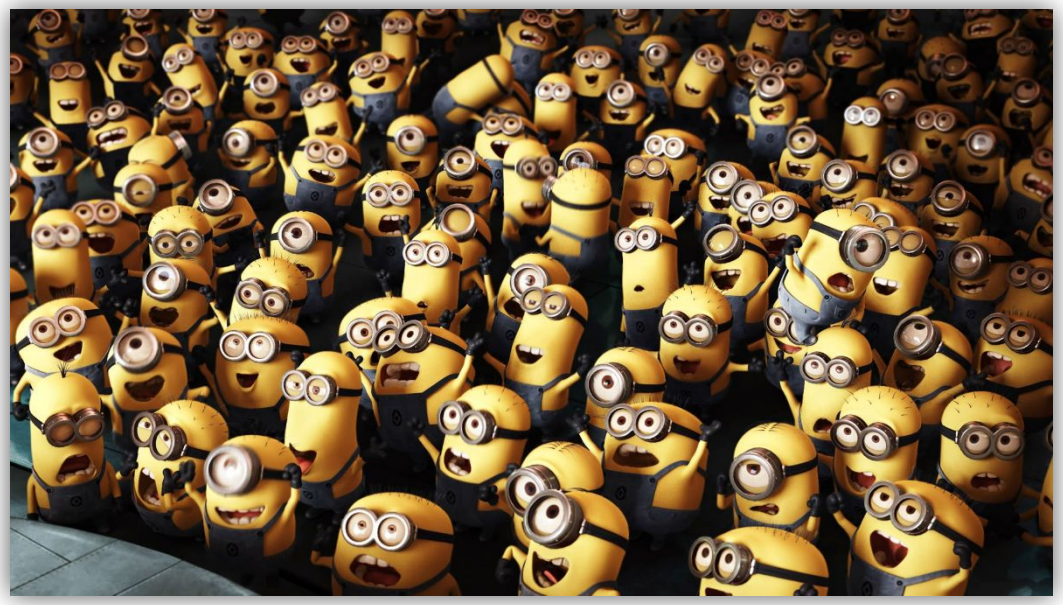
Josep Carmona
Computer Science Department

What is your background ?

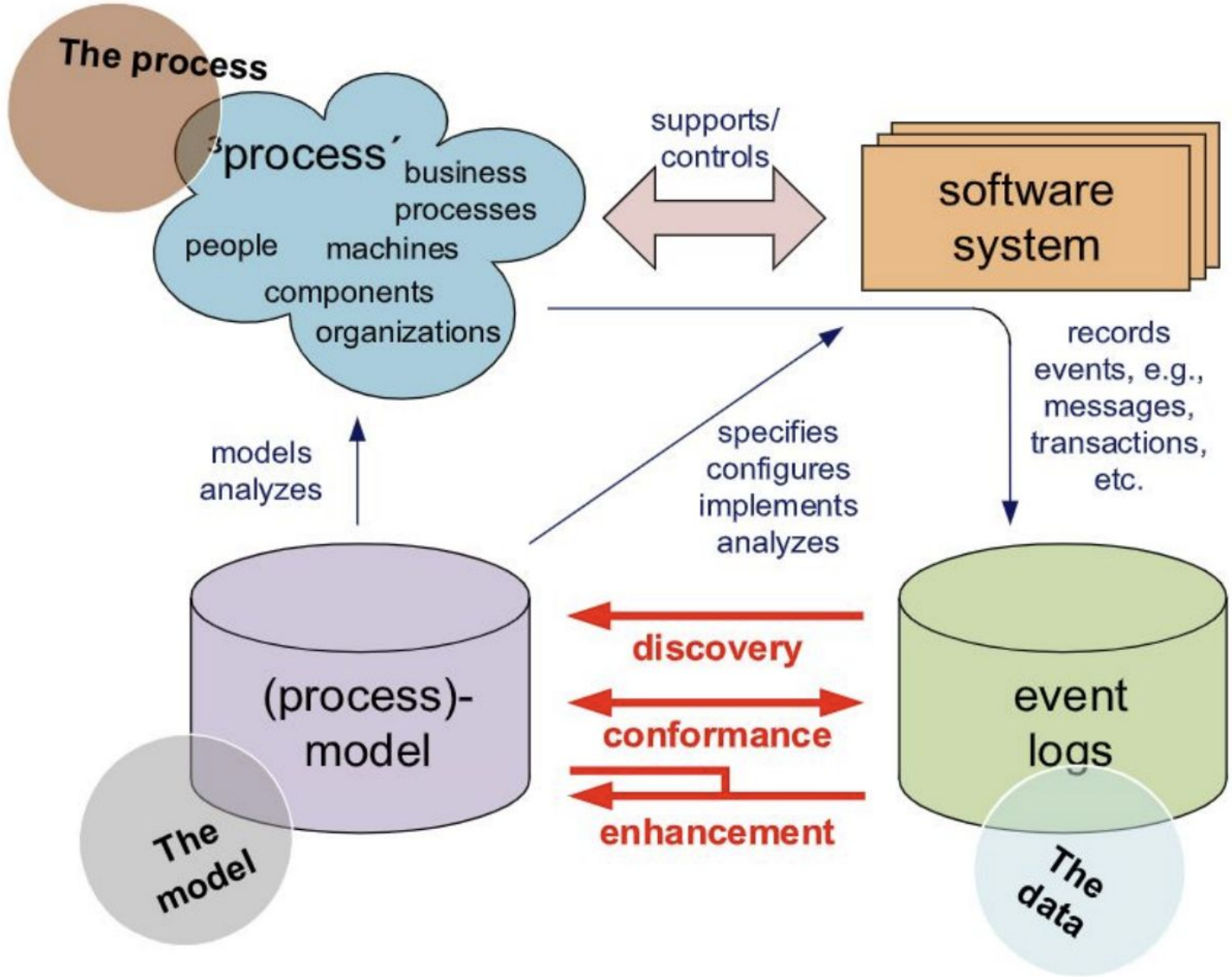
- CS vs something else?
- Bachelor at UPC or somewhere else?
- Other ?

What is your knowledge of

- Business processes?
- Behavioural formalisms, e.g., Petri nets?
- Data Mining?



Introduction



Course Goals

1. Understand the background, motivation, and potential usefulness of process mining
2. Master basic algorithms for discovery, conformance checking, and model enhancement
   - Understand their workings
   - Know their assumptions and limitations
   - Be able to adapt them to specific needs
3. Be able to start a process mining project in a real-world environment

Course Main Modules

- M1: Process Mining Overview, Positioning & Preliminaries (Event data & Process Models)
- M2: Process Discovery
- M3: Conformance Checking
- M4: Process Enhancement

Syllabus

- Process models and event data (2.5 weeks)
- Automatic process model discovery (2.5 weeks)
- Conformance checking of process models and event data (2.5 weeks)
- Evidence-based process enhancement grounded in event data (2 weeks)
- Assorted advanced techniques and applications (2 weeks)
- Methodology for PODS projects (1 week)

Theory/Execise Lectures

- Theory interleaved with exercises
- We use racó for collaborative reading of selected documents or videos. You can find there lectures and reading materials.
- Your participation in the lectures is key !
- In case of remote lecture (COVID19 comment):
  - Mute your mic
  - Questions on the chat or unmuting your mic

Lab

- Structure
  - First half of the semester: Hands-on, Modeling, Tooling (Bring your computer!)
  - Second half of the semester: course project
- Course project:
  - Groups of 3-5 students
  - List of projects to choose from
  - Submission of a report & presentation day
  - Types: comparison projects, repeatability, use-case, implementation
  - All projects require experimental work, ie., running different algorithms on certain data and measure results

Grading

- 60% final exam: similar to the exercises in class
- 40% course project, of which:
  - 20 % 2 Graded exercise in-lab (10% each)
  - 80% course Project.

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

Campus d'Excel·lència Internacional

Weekly Schedule

Week	Date	Lab (Tuesdays)	Date	Reading	Theory & Exercises (Thursdays)
1	10/9	Disco Hands-on (I) & Process Maps	12/9	V1	• PM Introduction & contextualization • Process models & event data: intro
2	17/9	Disco Hands-on (II)	19/9	R1	Process models & event data: basics
3	24/9		26/9	R2	Process models & event data: properties, algorithms & challenges
4	1/10	Disco Hands-on (III)	3/10	R3	• Quality dimensions for relating observed & modeled processes • Process Discovery: Alpha family
5	8/10	Process Modeling	10/10		Process Discovery: Advanced techniques
6	15/10	**Graded Exercise 1** ProM/Apromore (event data & discovery)	17/10	R4	Conformance checking: rule checking & token replay
7	22/10	Apromore (compliance & performance)	24/10		Methodology for PODS
8	29/10	Apromore (performance & variant)	31/10	R5	Midterm exams
10	5/11	Midterm exams	7/11		Midterm exams
11	12/11	Apromore (simulation)	14/11		Process enhancement techniques: basics techniques • Predictive process monitoring • Social network analysis
12	19/11	**Graded Exercise 2** Project work	21/11		Assorted advanced techniques I
13	26/11	Project work	28/11		Assorted advanced techniques II
14	3/12	Project work	5/12		*TBD: Celonis, DCR*
15	10/12	Project work	12/12		*Projects Presentations I,*
16	17/12	Project work **& Project Submission**	19/12		Project Presentations II

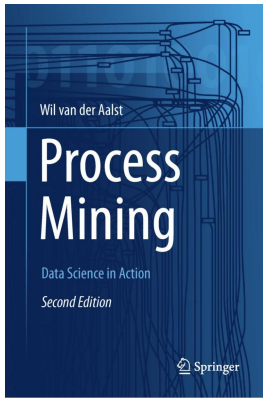
Bibliography

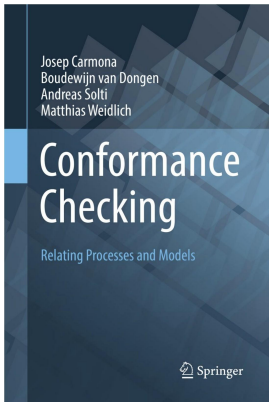
Basic

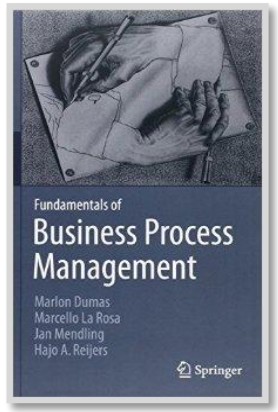
- Wil van der Aalst: Process Mining: Discovery, Conformance, and Enhancement of Business Processes, Springer 2016
- Josep Carmona, Boudewijn van Dongen, Andreas Solti, Matthias Weidlich: Conformance Checking, Springer 2018

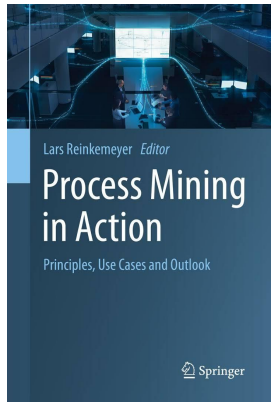
Additional (selected)

- Marlon Dumas et al.: Fundamentals of Business Process Management, Springer 2013
- Lars Reinkemeyer: Process Mining in Action: Principles, Use Cases and Outlook, 2020









Disclaimer

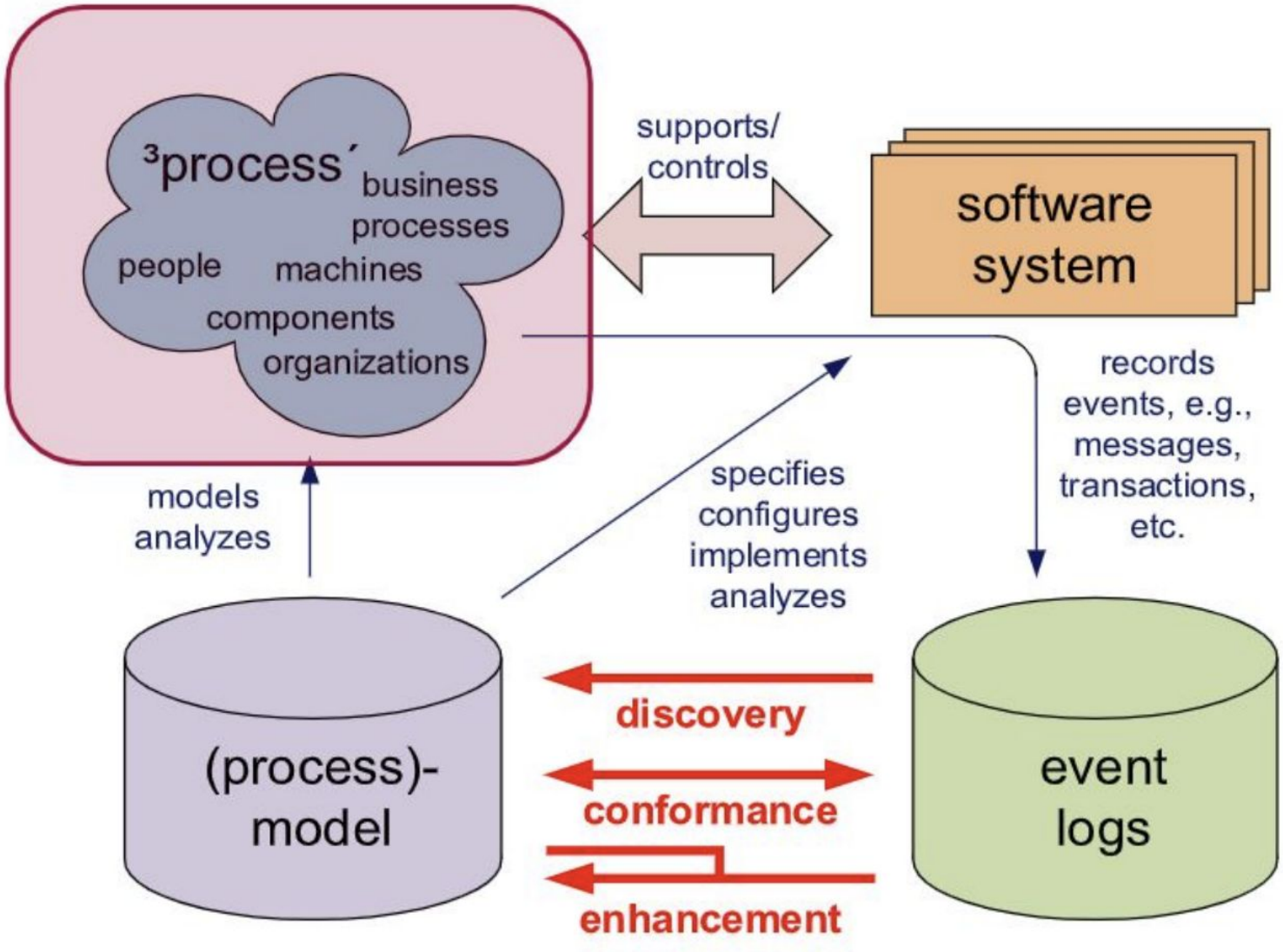
- Most of the material of this course is taken from my colleagues:
  - RWTH Aachen (Prof. Wil van der Aalst)
  - **Humboldt University zu Berlin (Prof. Matthias Weidlich)**
  - Technishe Universiteit Eindhoven (Prof. Boudewijn van Dongen)
  - University of Tartu (Prof. Marlon Dumas)
  - University of Melbourne (Prof. Marcello La Rosa)
  - Technical University of Denmark (Prof. Andrea Burattin)
- Hence, this material is only provided for your learning, please do not share nor publish

Questions ?

Outline

- **M1: Process Mining Overview, Positioning & Preliminaries (Event data & Process Models)**
- M2: Process Discovery
- M3: Conformance Checking
- M4: Process Enhancement

Context


'process'
business
processes
people
machines
components
organizations
supports/
controls
software
system
records
events, e.g.,
messages,
transactions,
etc.
models
analyzes
specifies
configures
implements
analyzes
(process)-
model
discovery
conformance
enhancement
event
logs

Relevance of Business Processes

Business processes are everywhere

- Products and services are provided by activities
- Execution of activities requires coordination
- Success of this coordination influences costs, time, and quality of products and services

"a collection of activities that take one or more kinds of input and create an output that is of value to the customer" [Hammer & Champy 1993]

"a set of logically related tasks performed to achieve a defined business outcome for a particular customer or market" [Davenport 1992]

Question

What business processes
did you face in the last two weeks?

Have you been satisfied
with how they have been run?

Scenario: Insurance Claim Handling

Record claim

Check coverage

Request proof of loss

Do field check

Take decision

Inform claimant

Compensation payment

Archive claim

Scenario: Online Reselling

Submit order

Check credit history

Charge credit card

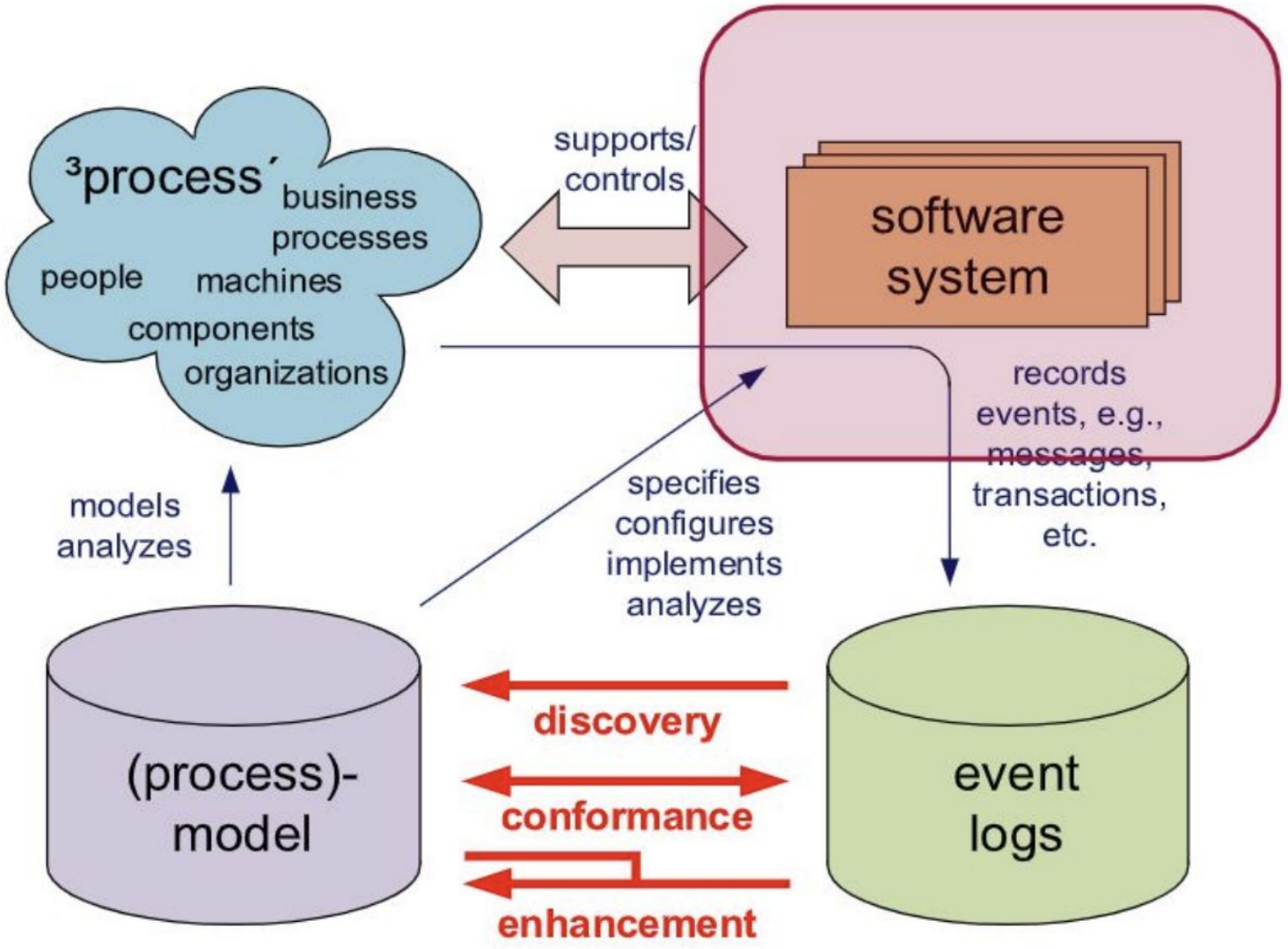
Check availability

Plan shipments

Aggregate shipments

Last mile delivery

Context


'process'
business
processes
people
machines
components
organizations
supports/
controls
software
system
records
events, e.g.,
messages,
transactions,
etc.
models
analyzes
specifies
configures
implements
analyzes
(process)-
model
discovery
conformance
enhancement
event
logs

Process-oriented Information System

Process-oriented Information System (POIS)

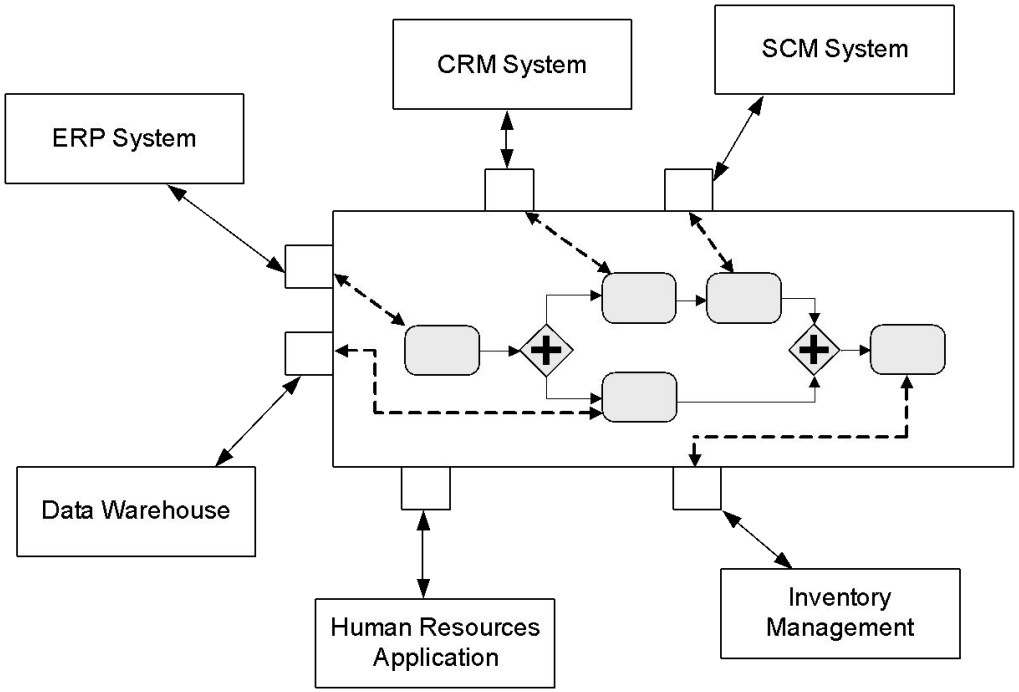
- ”a generic software system that is driven by explicit process representations to coordinate the enactment of business processes” [Weske 2007]

Process-orchestration

- ”a system acts as a central agent that controls the execution of the process activities, very similar to a conductor centrally controlling the musicians in an orchestra”

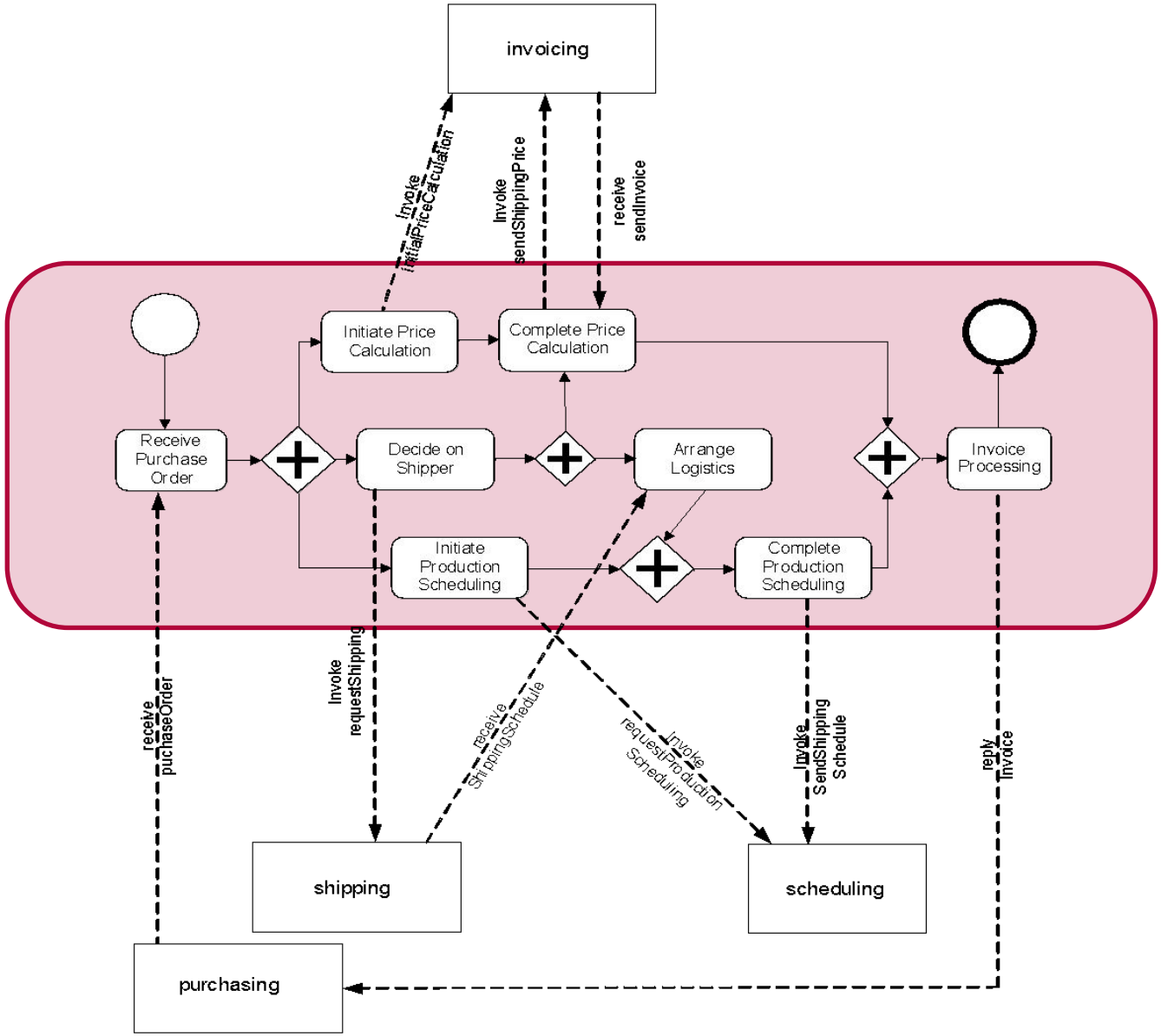
Process-based Integration

Integration logic is encoded in process model
Workflow engine executes the integration process
System activities vs. human activities



from M. Weske: Business Process Management, © Springer-Verlag Berlin Heidelberg 2007

Service Composition Process



Beyond System Workflows

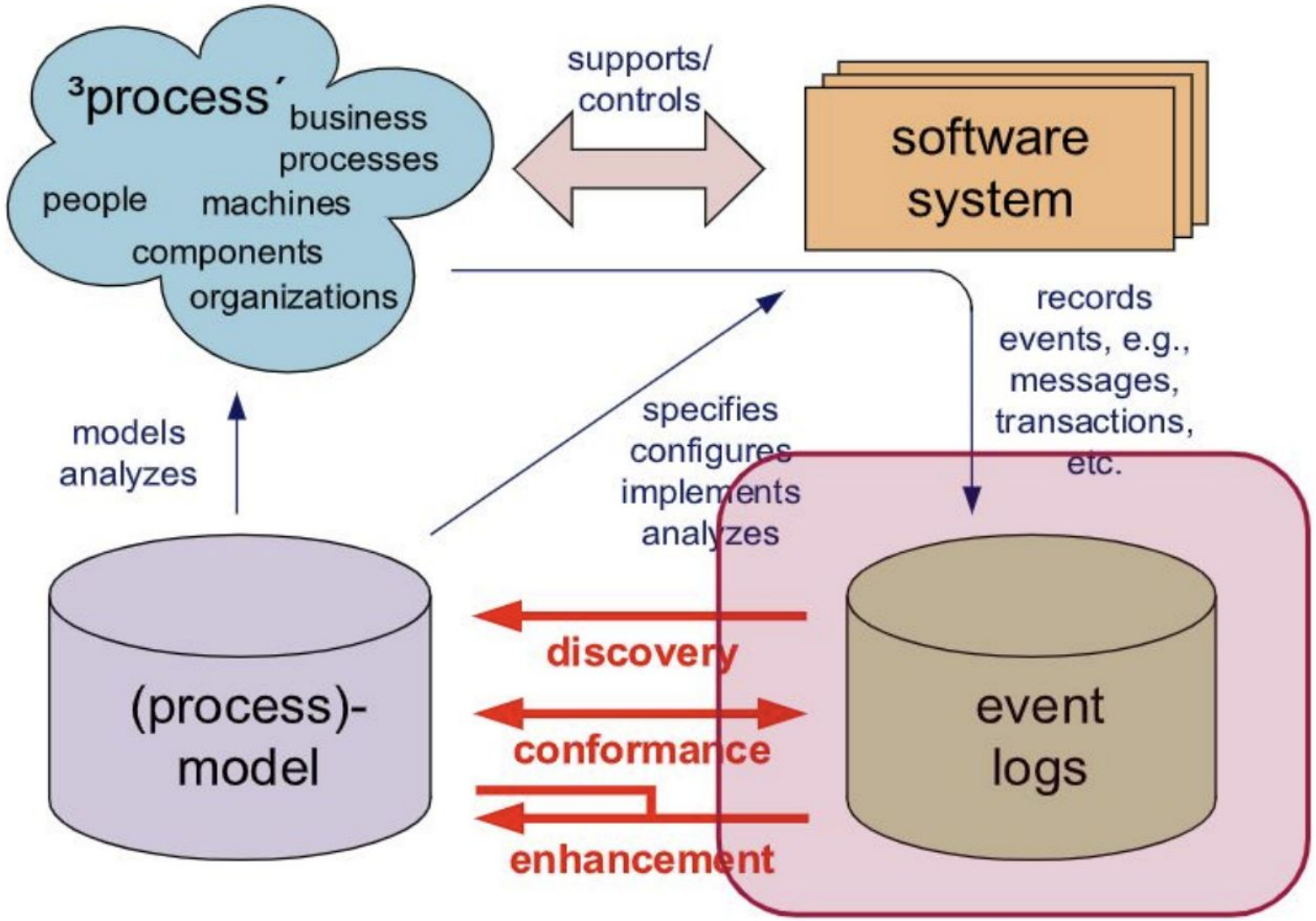
Human Interaction Workflows

- User interaction during process execution
- Combination of manual and fully automated activities
- Active control of process by interaction with process participants

Human workflow systems typically also include:

- Modelling and integration of process participants (roles, capabilities)
- Provisioning of specific interfaces (work lists)

Context


'process'
business
processes
people
machines
components
organizations
supports/
controls
software
system
records
events, e.g.,
messages,
transactions,
etc.
models
analyzes
specifies
configures
implements
analyzes
(process)-
model
discovery
conformance
enhancement
event
logs

Events in Everyday life

Events in Process Mining

Event – *happening of interest*

- Have timestamp: occurrence time, arrival time, …
- Carry data
  - Typical relational model based on attributes
  - Payload is modelled as key-value pairs

Event type – type for events of similar structure and semantics

- Events are instances of event type
- Defines the set of attributes of the events

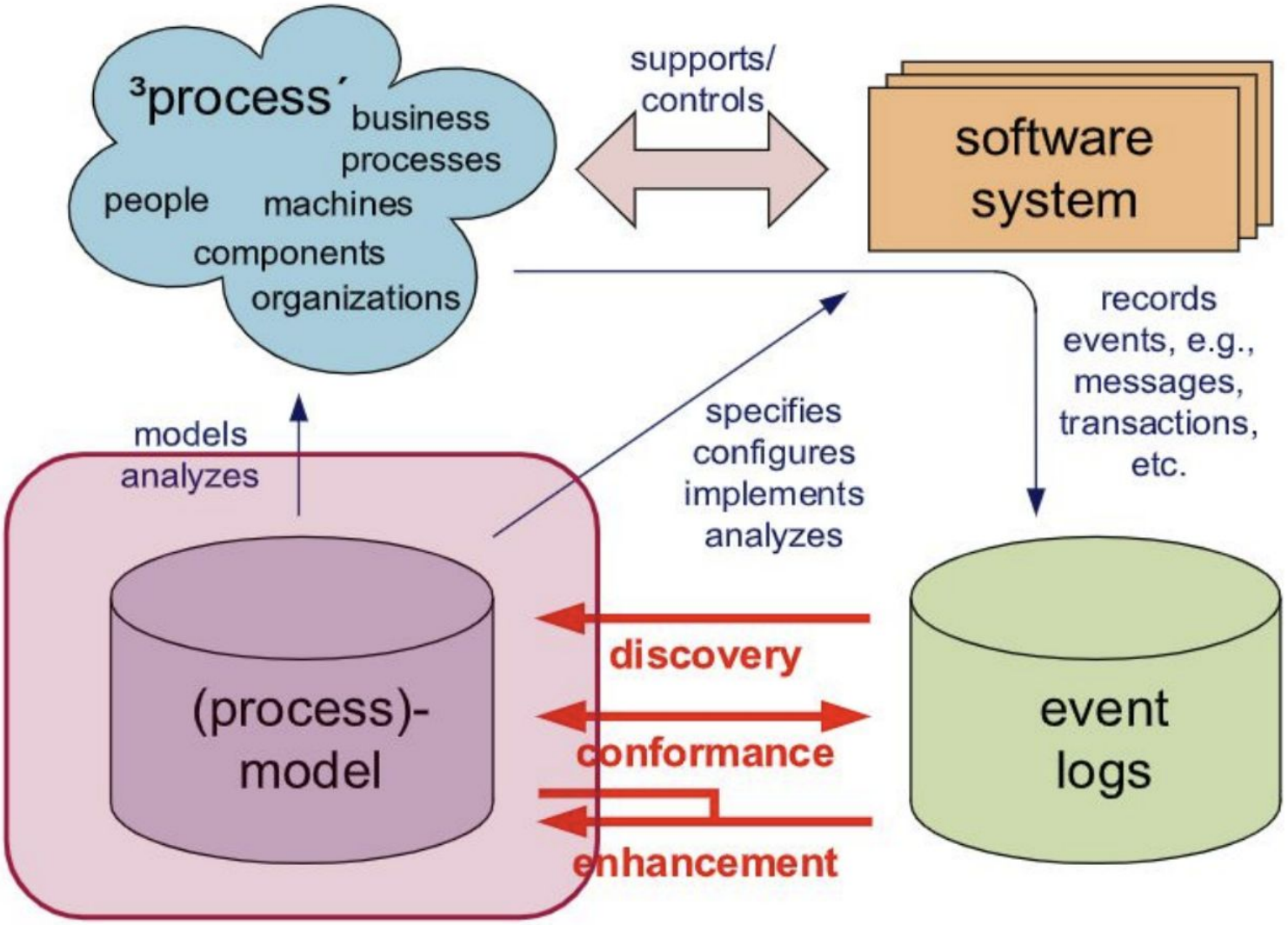
Event Logs

Process context

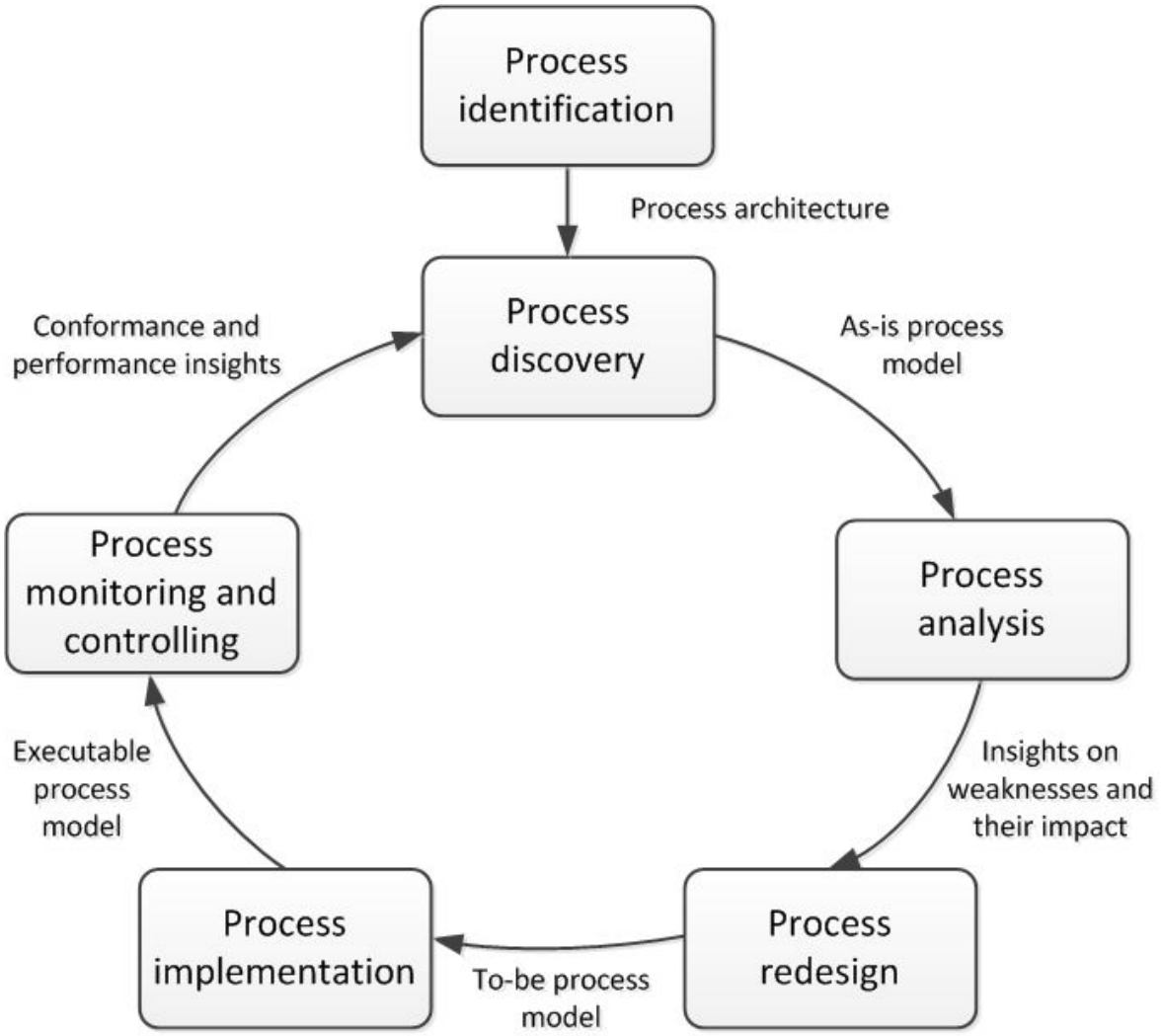
- Activity – what has been executed?
- Time – when has it been executed?
- Case – for which process instance has it been executed?

Trans. ID	Activity	Start Time	End Time	Resource
8287	Enter customer data	08:34:15	08:37:44	User jsmith
8287	Check credit	08:37:52	08:38:05	Equifax service call
1399	Enter customer data	08:37:59	08:44:40	User sjones
8287	Enter order	08:38:09	08:38:39	ERP system call
1399	Check credit	08:44:58	08:45:06	Equifax service call
4283	Enter order	08:45:01	08:45:35	ERP system call
1399	Enter order	08:45:18	08:45:38	ERP system call

Context


process
business
processes
people
machines
components
organizations
supports/
controls
software
system
records
events, e.g.,
messages,
transactions,
etc.
models
analyzes
specifies
configures
implements
analyzes
(process)-
model
discovery
conformance
enhancement
event
logs

BPM Lifecycle and Process Models



Purposes of Modeling

Large variety of modelling purposes

- Business purposes
- Information systems purposes

Business purposes

- Documentation, guidelines, work instructions
- Process redesign, from as-is to to-be
- Staff planning, often using statistical annotations
- Quality certification

Purposes of Modeling

Information systems purposes

- Enterprise Resource Planning (ERP) system selection
  - ERP systems provide business functionality
  - System selection based on delta-analysis of own processes and implemented process
- Software development
  - Process models as requirement documents
- Process implementation
  - Workflow system supports execution of cases
  - Different degrees of automation of activities

Essence of Modeling

Original

© Stiftung Deutsches Technikmuseum Berlin

Abstraction

Model of

Abstraction

Model of

Model Level

Process models answer questions

- **What is done?**
- **Who is responsible?**
- **What are the decisions taken?**
- **How long does it take to finish the process?**
- **Who is affected by a change in the process?**

Mapping Business Processes

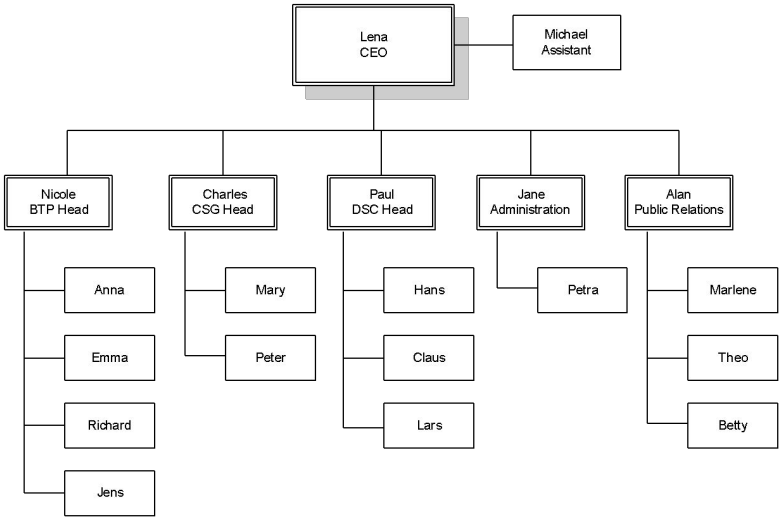
What is mapped to a process model?

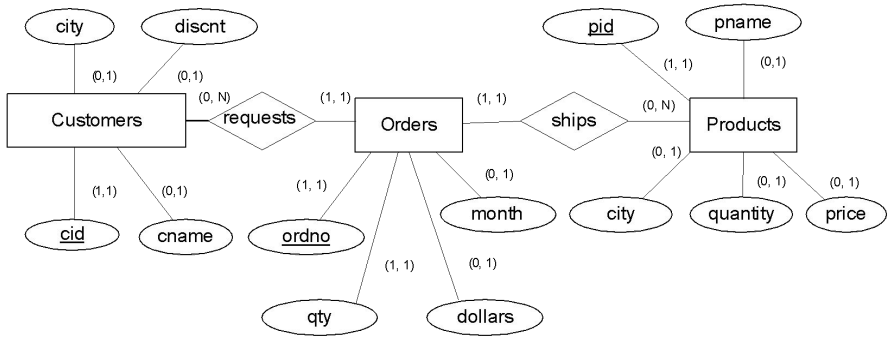
- Activities
  Building blocks that describe elementary pieces of work
- Routing conditions
  Describe temporal and logical constraints on the execution of activities
- Inputs, Outputs
  Informational or physical artefacts processed by activities
- Events
  How time, messages, exception influence the execution
- Resources

Perspectives

Perspectives of process modelling

- Control flow
- Organisational structures
- Data structures
- IT landscapes





from M. Weske: Business Process Management, © Springer-Verlag Berlin Heidelberg 2007

Take Away

Process Mining links data analysis with process management

Event data is used to discover process models, assess their conformance, or enhancement them with additional details