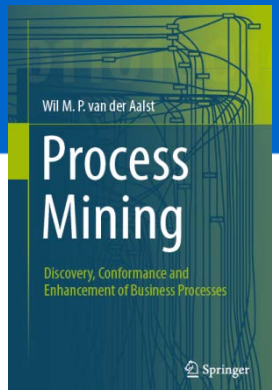
Process Mining: Data Science in Action

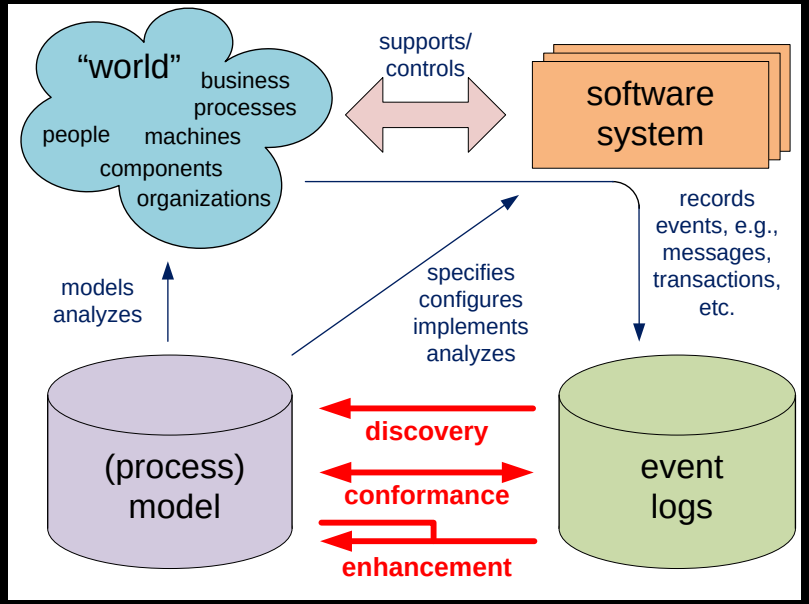
Refined Process Mining Framework



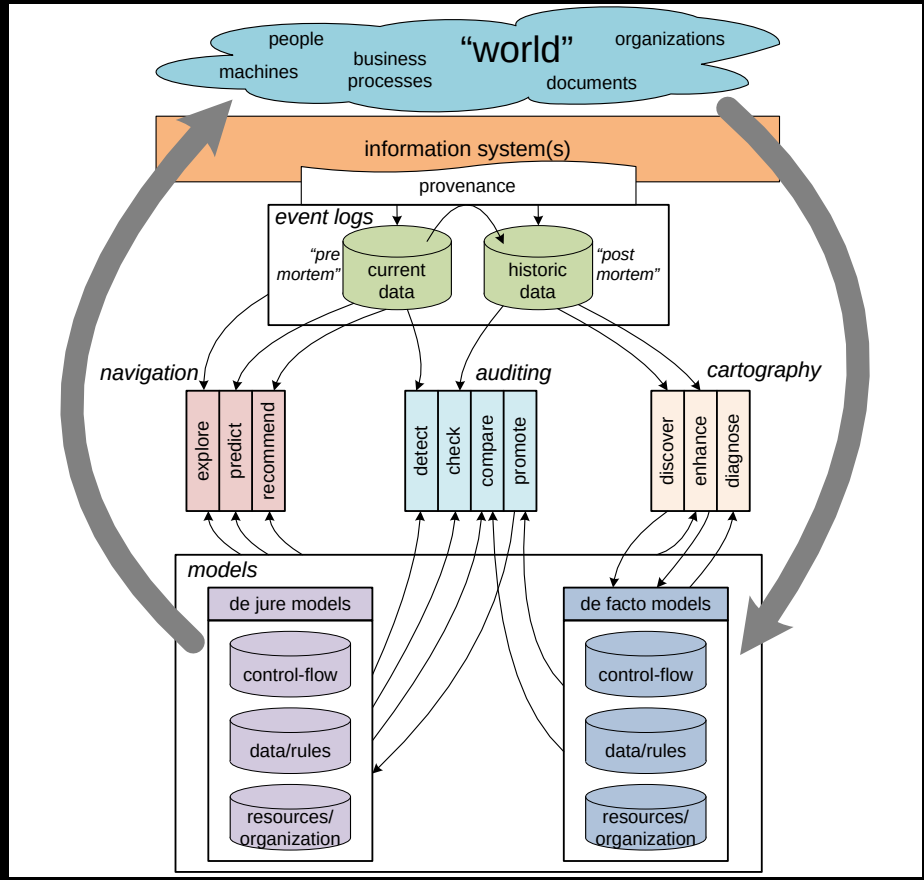
prof.dr.ir. Wil van der Aalst
www.processmining.org

Technische Universiteit Eindhoven University of Technology

Where innovation starts



refined process mining framework



process discovery
(alpha miner, heuristic miner, region-based miners, etc.)

offline

conformance checking
(token-based, footprints, alignments, etc.)

organizational mining

decision point mining

bottleneck mining

---

operational support

data quality

concept drift analysis

queue mining

prediction

recommendation

declarative mining

reference model mining

…

process configuration

mining of event streams

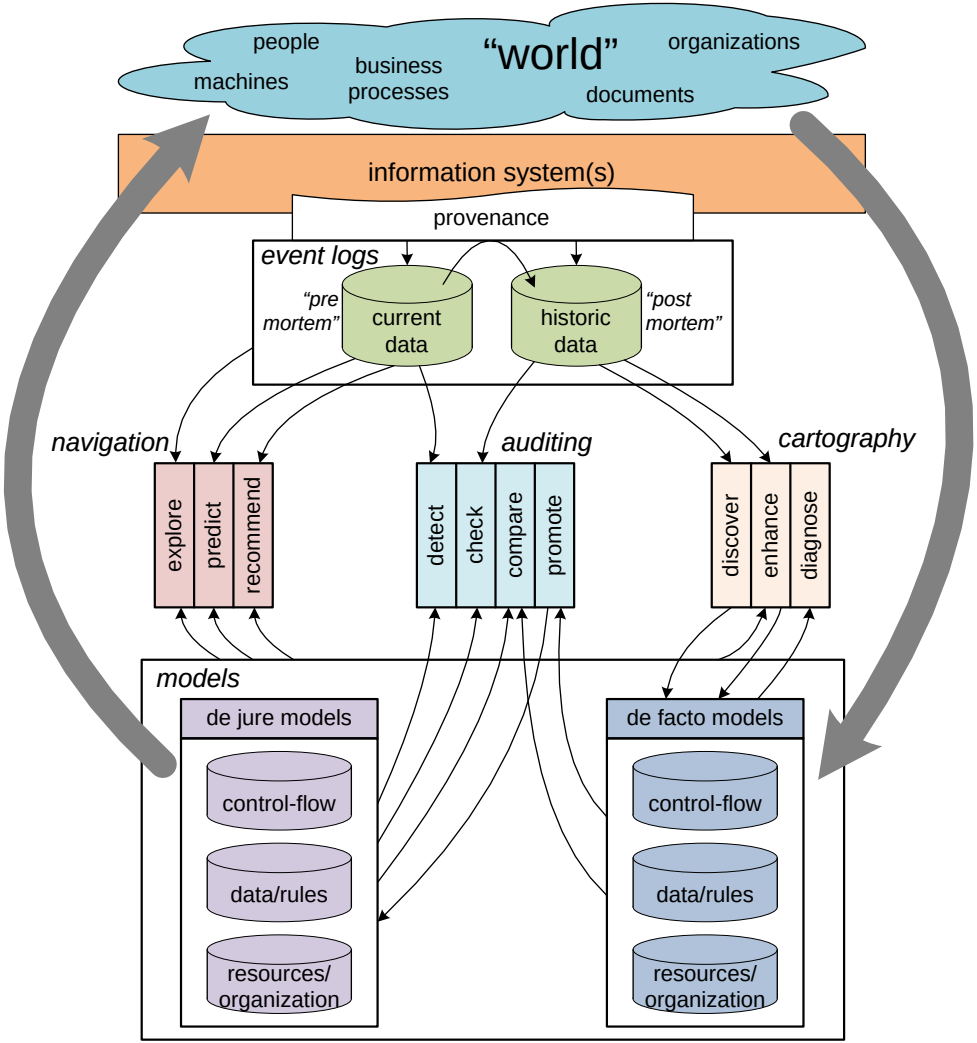
artifact-centric mining

distributed process mining

model repair

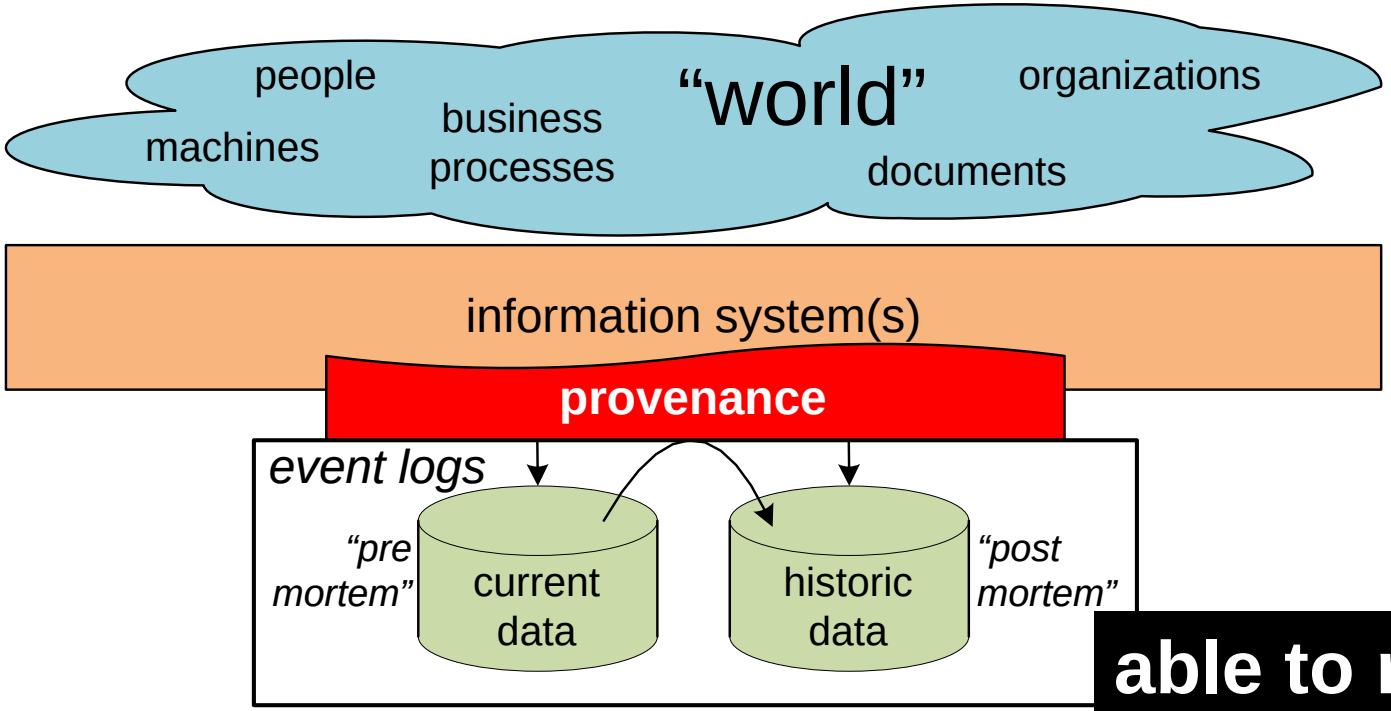
mining on partially ordered event data

also online


people
machines
business processes
"world"
organizations
documents
information system(s)
provenance
event logs
"pre mortem"
current data
historic data
"post mortem"
navigation
auditing
cartography
explore
predict
recommend
detect
check
compare
promote
discover
enhance
diagnose
models
de jure models
control-flow
data/rules
resources/ organization
de facto models
control-flow
data/rules
resources/ organization


refined process mining framework

(Business) process provenance

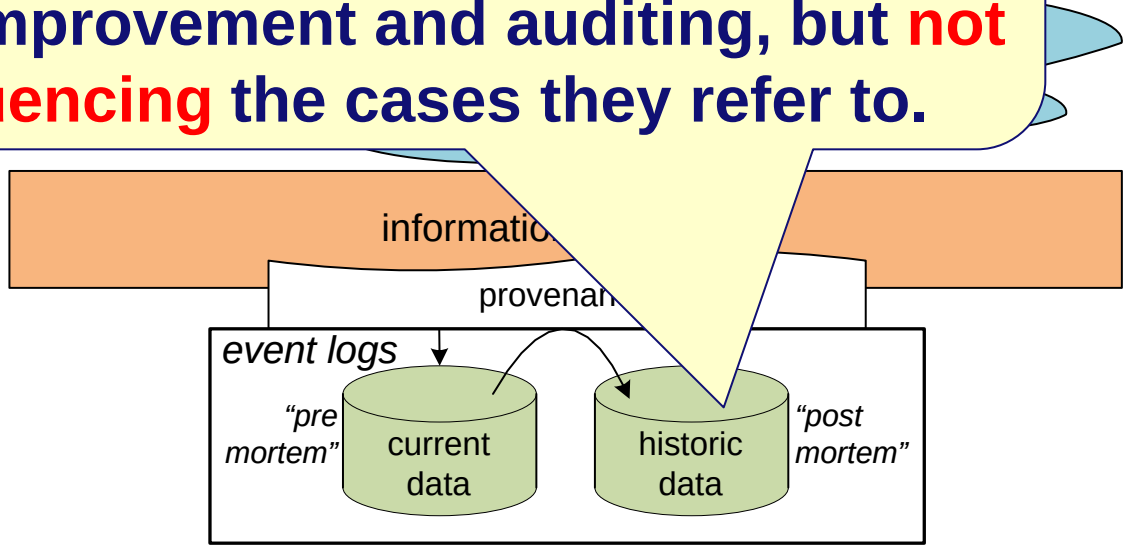


able to replay history

TU/e

Pre mortem and post mortem event data

“Post mortem” event data refer to information about cases that have completed, i.e., these data can be used for process improvement and auditing, but not for influencing the cases they refer to.

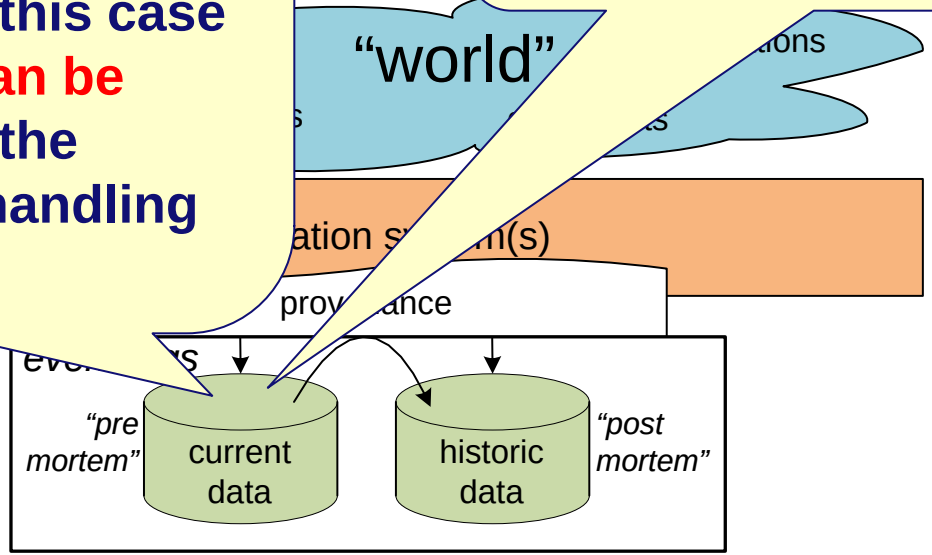


TU/e

Pre mortem and post mortem event data

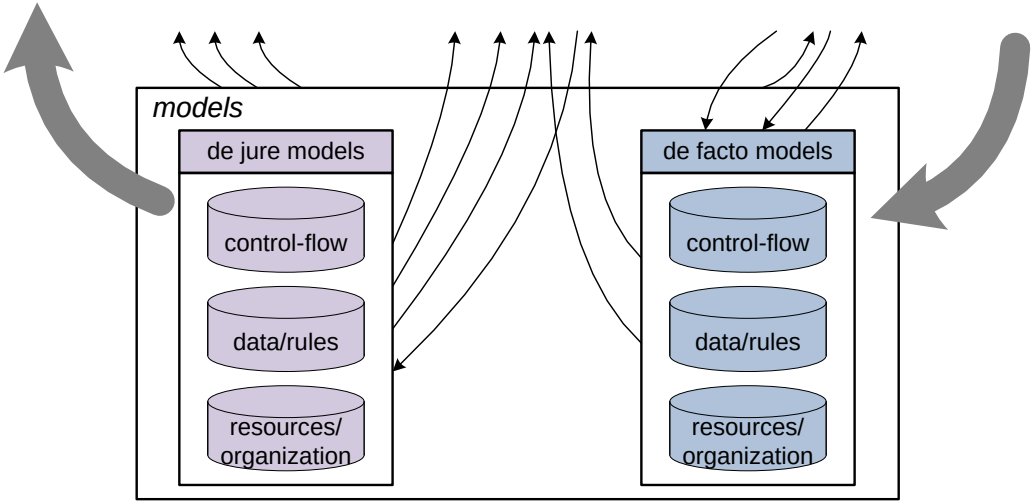
If a case is still running, i.e., the case is still "alive" (pre mortem), then it may be possible that information in the event log about this case (i.e., current data) can be exploited to ensure the correct or efficient handling of this case.

"Pre mortem" event data refer to cases that have not yet completed.



	student-related event data	sales-related event data	patient-related event data
post mortem	Understanding where and why students drop out or deviate. Should the curriculum be redesigned? What are the bottlenecks?	Understanding where and why customers are lost. Where are the bottlenecks in the sales process? How to redesign the sales process?	Are patients treated in time? Why do different doctors operate in different ways? How to save costs? How to improve service levels?
pre mortem	What advice can we give a particular student that is likely do drop out? How to signal the lecturer that the exam is likely to be a "massacre" due to inactivity of students?	When to trigger a customer with a partially completed order? When to reroute an order to an account manager?	Predicting the most likely time until surgery. Which doctor should be selected to treat the patient? Should the patient be moved to another hospital (for logistic reasons).

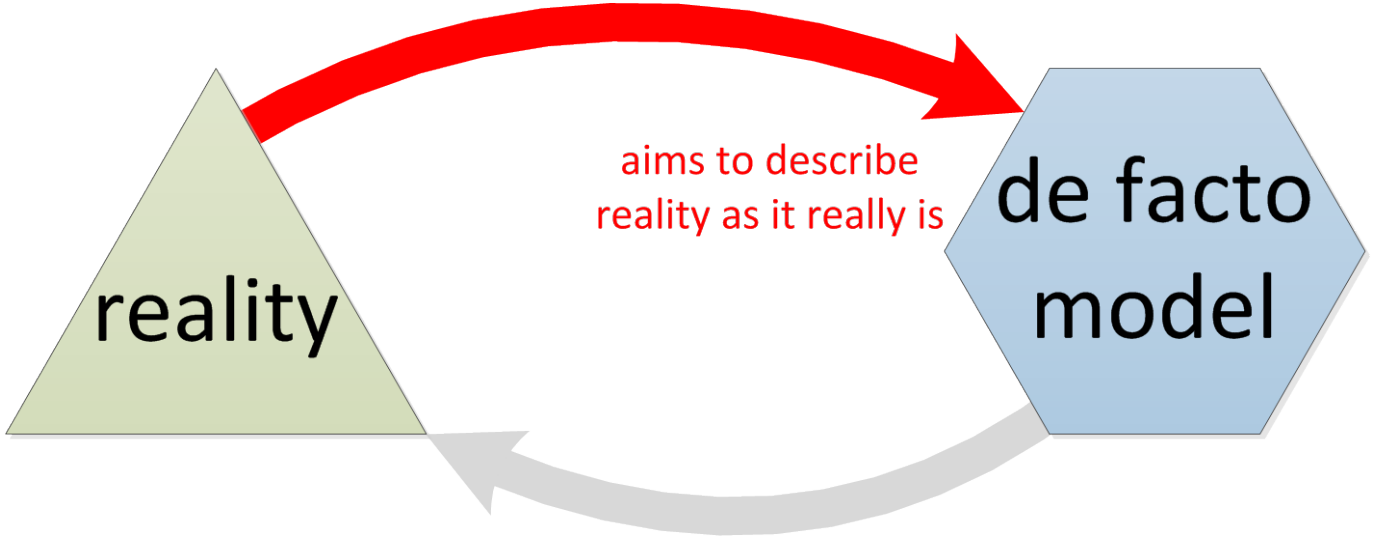
Different types of models



- **Models may cover one or more perspectives**
  - **control-flow**
  - **data/rules**
  - **resources/org.**
  - **time**
  - **costs**
  - **…**
- **“de jure” models and “de facto” models**

TU/e

A de facto model is descriptive

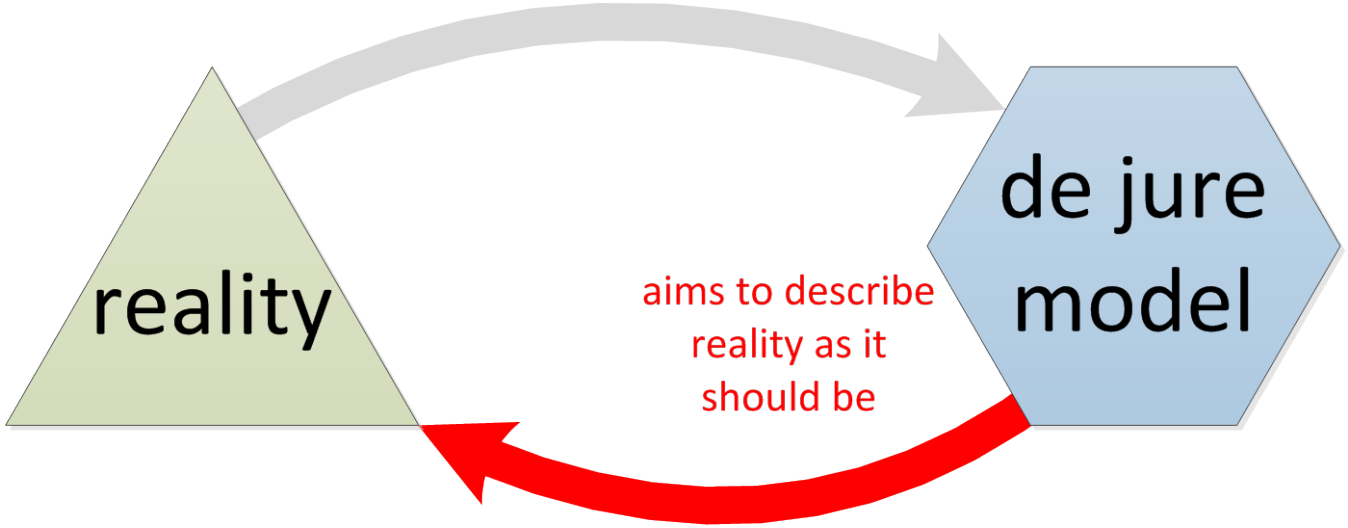


- Its purpose is to describe the "as is" process, and not to steer or control reality. De facto models aim to capture reality.
- Insights may be used for reengineering, operational support, etc.

TU/e

A de jure model is normative

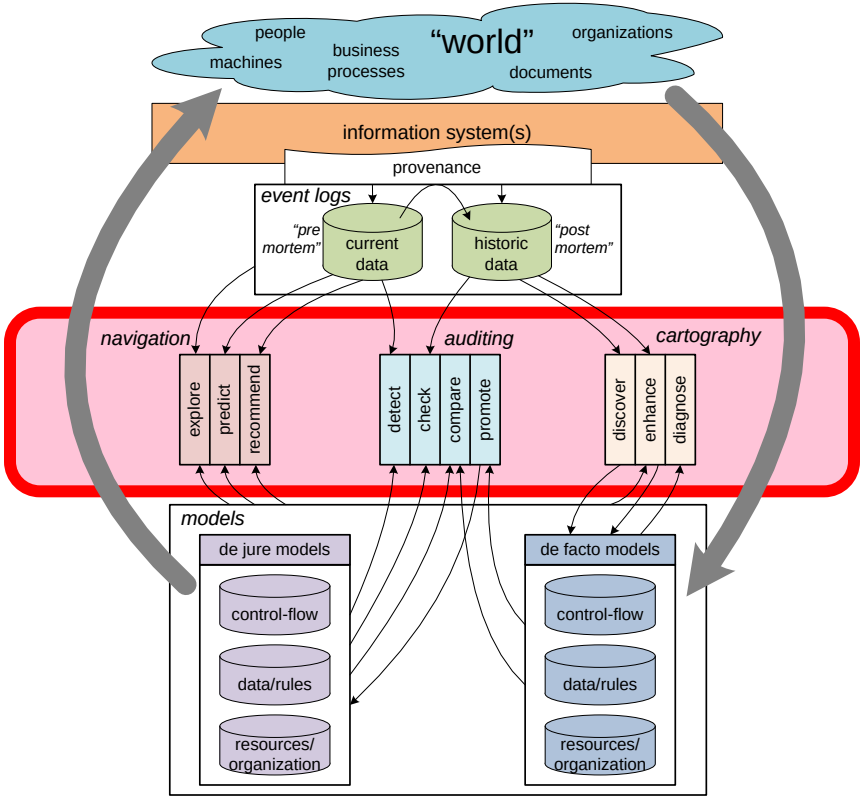
It specifies how things should be done or handled



- For example, a process model used to configure a BPM system is normative and forces people to work in a particular way.
- In other situations, normative models may be ignored by workers ("wallpaper models").

TU/e

Ten process mining activities

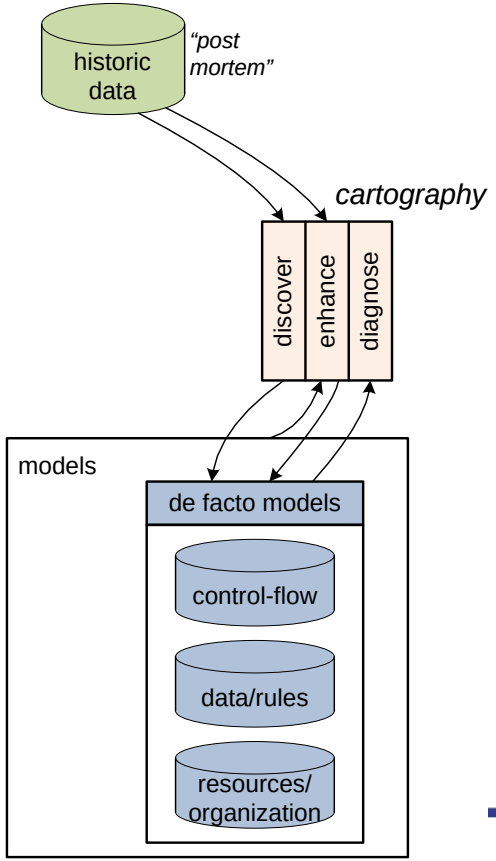


not intended to be complete

TU/e

Cartography: Process models as maps

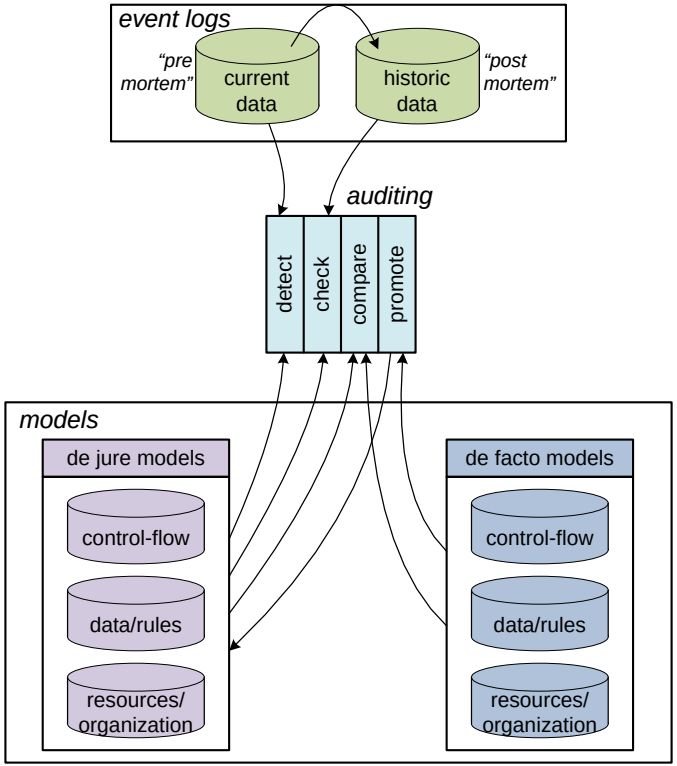
- **Discover**. This activity is concerned with the extraction of (process) models.
- **Enhance**. When existing process models (either discovered or hand-made) can be related to events logs, it is possible to enhance (**extend** and **repair**) these models.
- **Diagnose**. This activity does not directly use event logs and focuses on classical model-based analysis.



TU/e

Auditing: Confronting model and reality

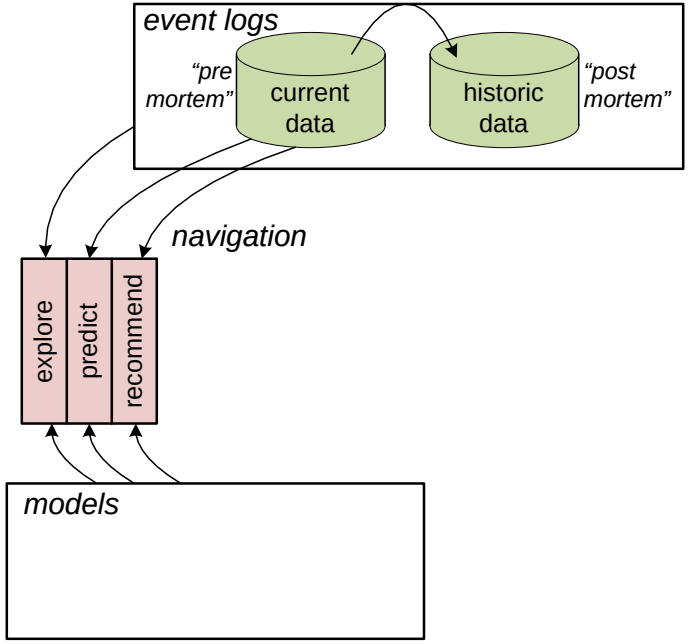
- **Detect**. Compares de jure models with current "pre mortem" data. The moment a predefined rule is violated, an alert is generated (**online**).
- **Check**. The goal of this activity is to pinpoint deviations and quantify the level of compliance (**offline**).
- **Compare**. De facto models can be compared with de jure models to see in what way reality deviates from what was planned or expected.
- **Promote**. Promote parts of the de facto model to a new de jure model.



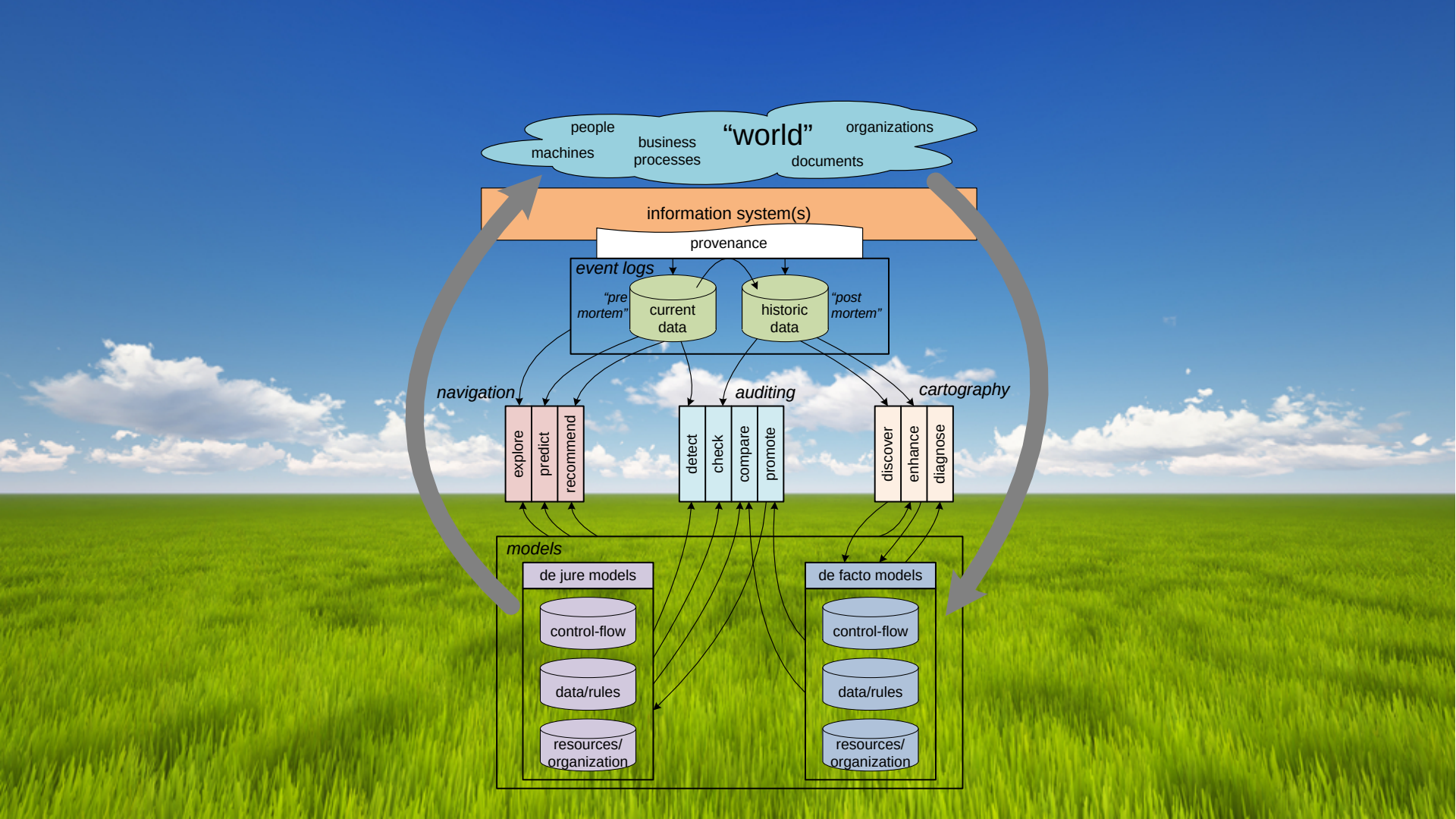
TU/e

Navigation: Supporting and guiding process execution

- **Explore**. The combination of event data and models can be used to explore business processes at run-time.
- **Predict**. By combining information about running cases with models, it is possible to make predictions about the future, e.g., the remaining flow time and the probability of success.
- **Recommend**. The information used for predicting the future can also be used to recommend suitable actions (e.g. to minimize costs or time).



TU/e


people
machines
business processes
"world"
organizations
documents
information system(s)
provenance
event logs
"pre mortem"
current data
historic data
"post mortem"
navigation
auditing
cartography
explore
predict
recommend
detect
check
compare
promote
discover
enhance
diagnose
models
de jure models
control-flow
data/rules
resources/ organization
de facto models
control-flow
data/rules
resources/ organization

Part I: Preliminaries

- Chapter 1: Introduction
- Chapter 2: Process Modeling and Analysis
- Chapter 3: Data Mining

Part II: From Event Logs to Process Models

- Chapter 4: Getting the Data
- Chapter 5: Process Discovery: An Introduction
- Chapter 6: Advanced Process Discovery Techniques

Part III: Beyond Process Discovery

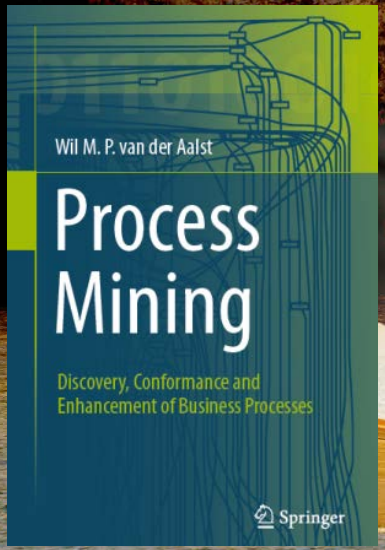
- Chapter 7: Conformance Checking
- Chapter 8: Mining Additional Perspectives
- Chapter 9: Operational Support

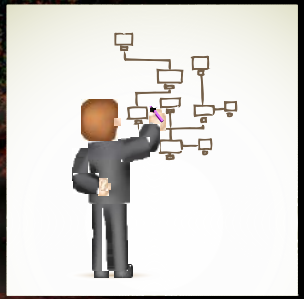
Part IV: Putting Process Mining [illegible]

- Chapter 10: Tool Support
- Chapter [illegible]: Analyzing "Lasagna Processes"
- Chapter 12: Analyzing "Spaghetti Processes"

Part V: Reflection

- Chapter 13: Cartography and Navigation
- Chapter 14: Epilogue





TU/e