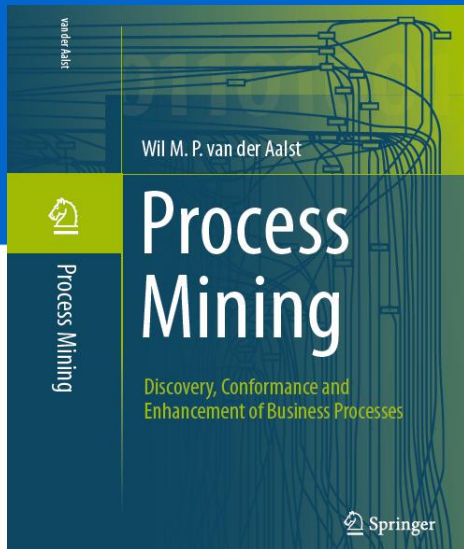


Operational Support

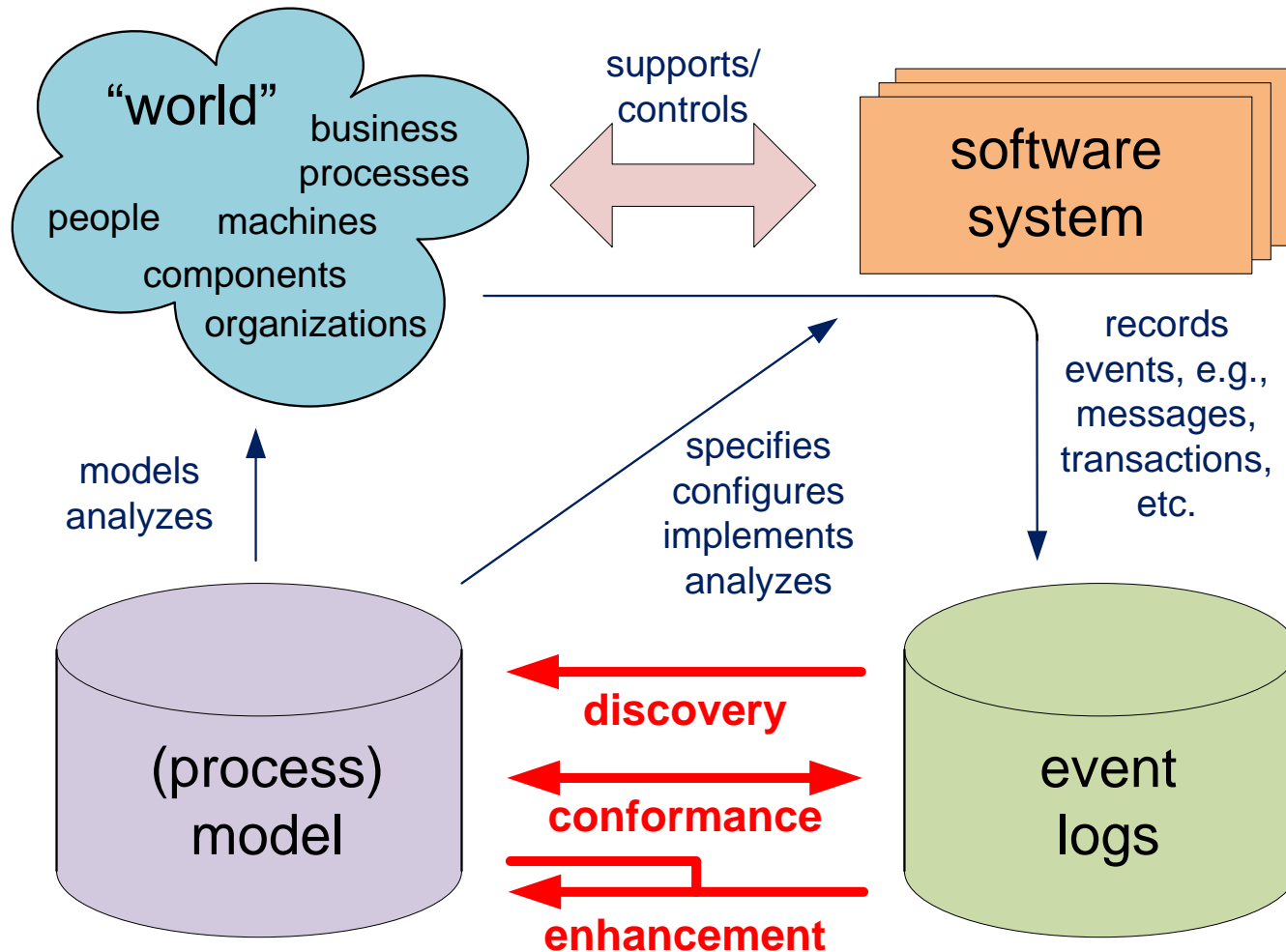
prof.dr.ir. Wil van der Aalst
www.processmining.org



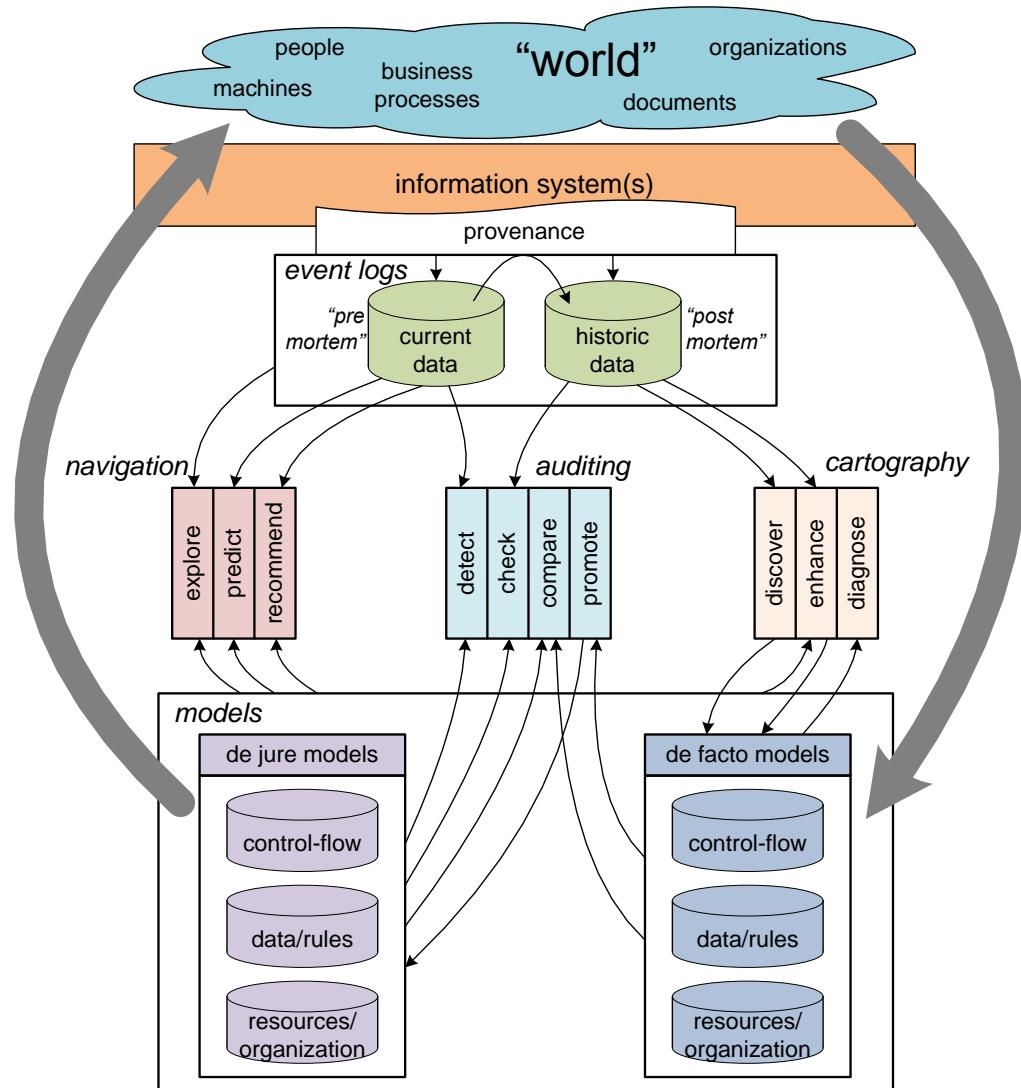
TU / **e** Technische Universiteit
Eindhoven
University of Technology

Where innovation starts

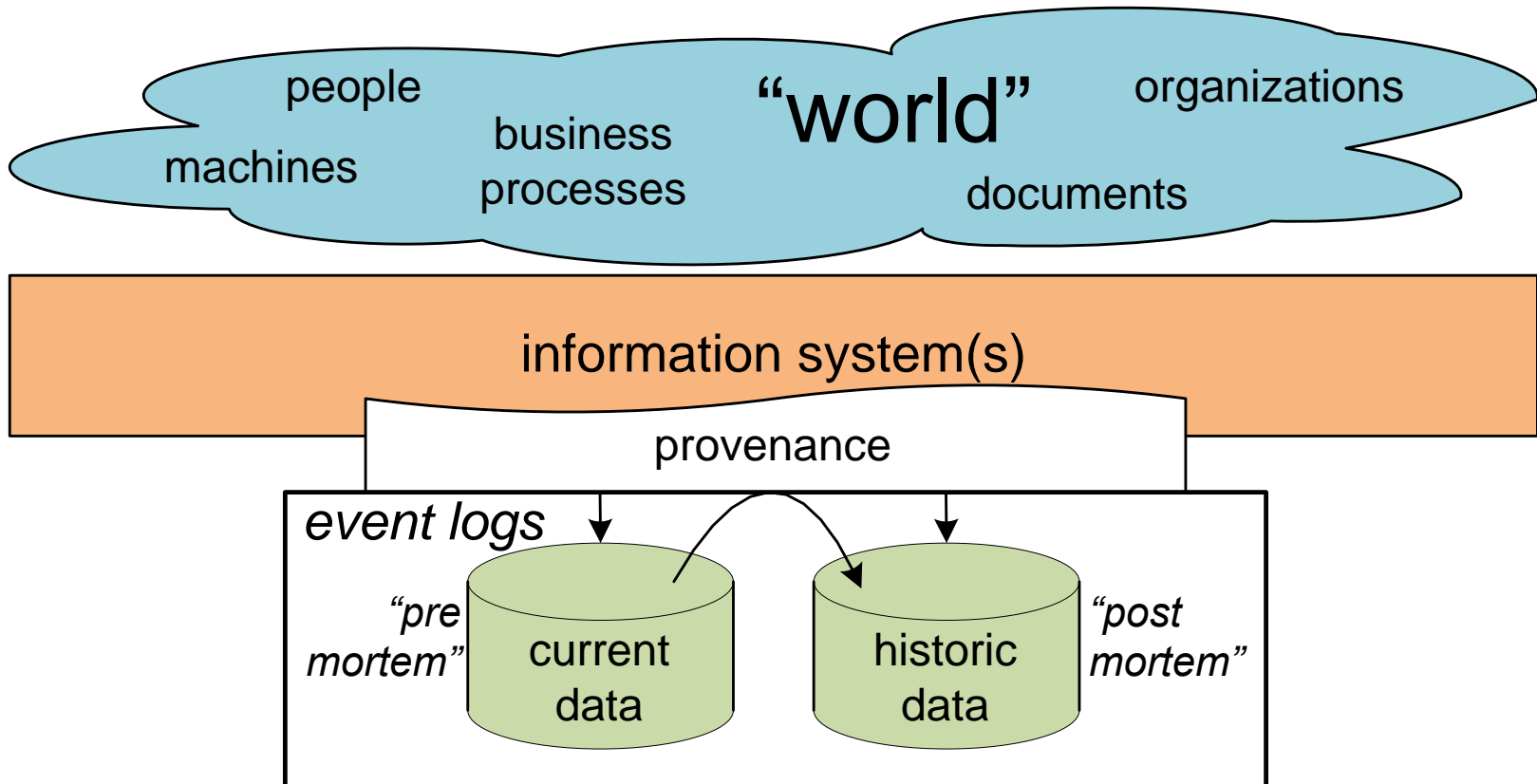
Process mining spectrum



Refined process mining framework



Business process provenance



Two types of event data: post and pre mortem

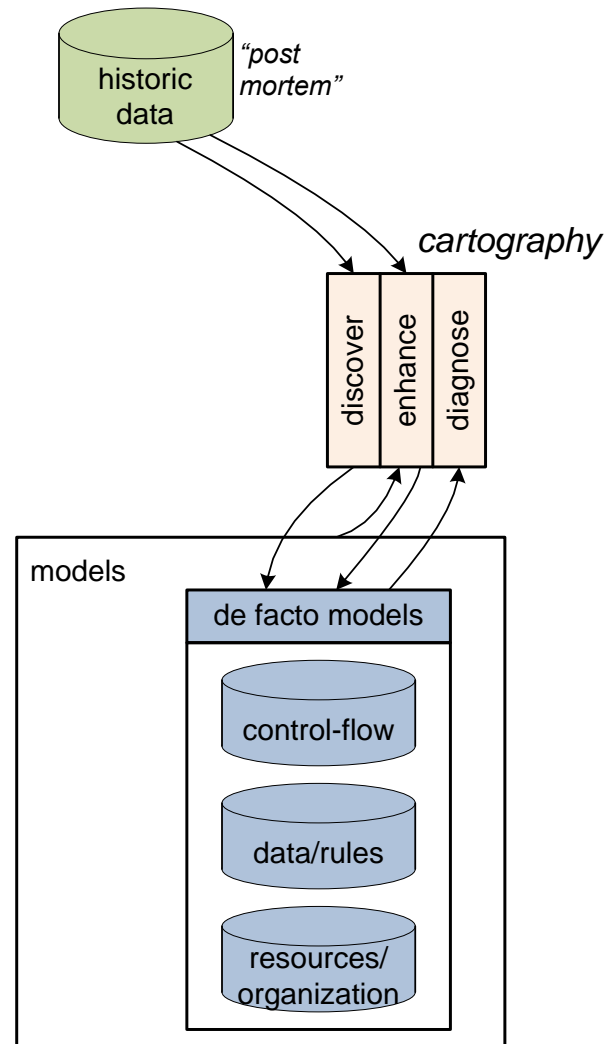
- **“Post mortem” event data** refer to information about cases that have completed, i.e., these data can be used for process improvement and auditing, but **not for influencing** the cases they refer to.
- **“Pre mortem” event data** refer to cases that have not yet completed. If a case is still running, 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.

Two types of models: “de jure models” and “de facto models”

- A **de jure model** is **normative**, i.e., 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.
- A **de facto model** is **descriptive** and its goal is not to steer or control reality. Instead, de facto models aim to capture reality.

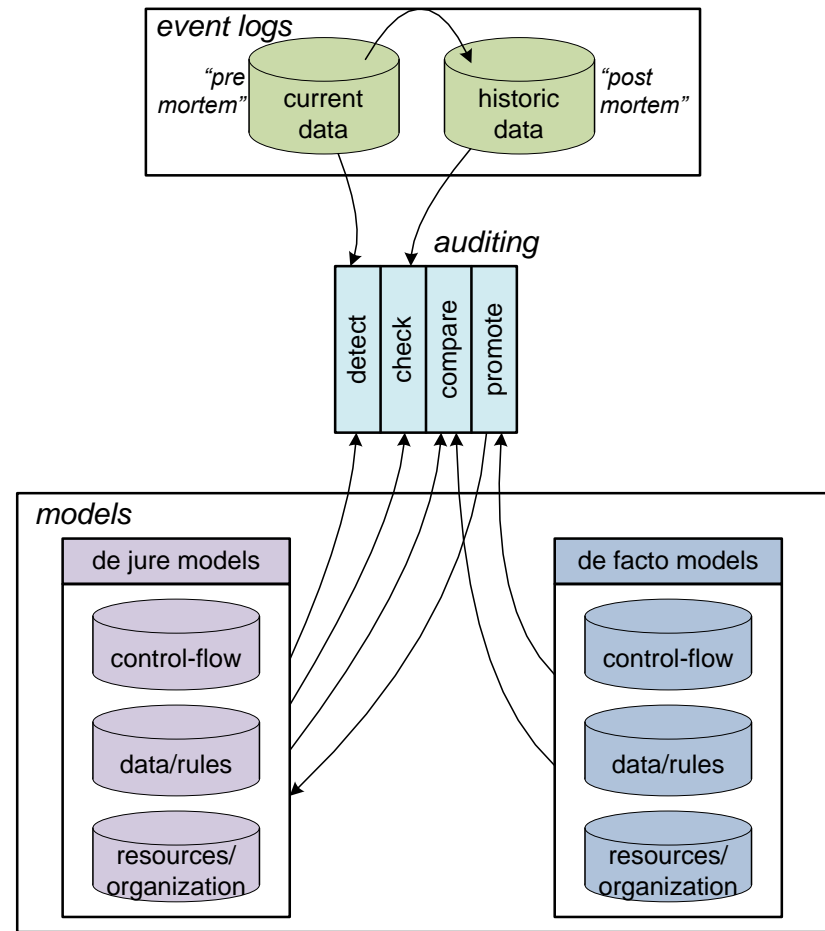
Cartography

- **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 these models.
- **Diagnose.** This activity does not directly use event logs and focuses on classical model-based analysis.



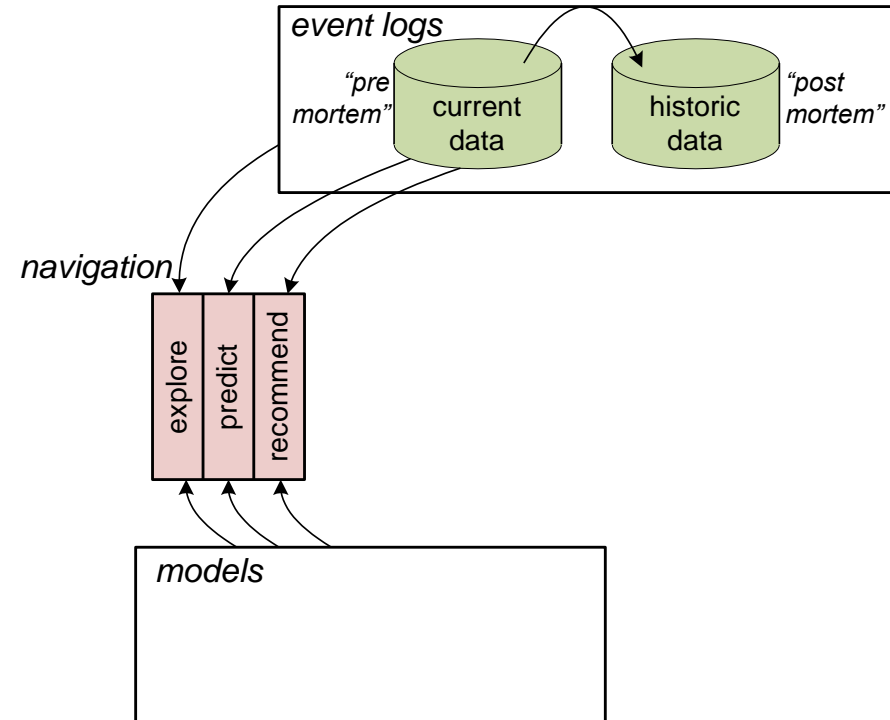
Auditing

- **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.

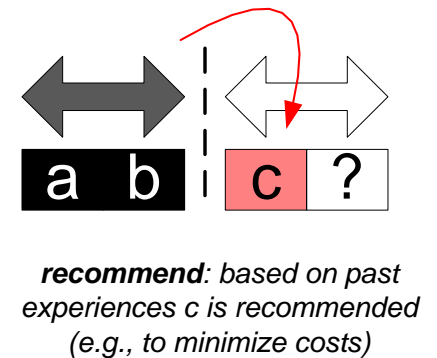
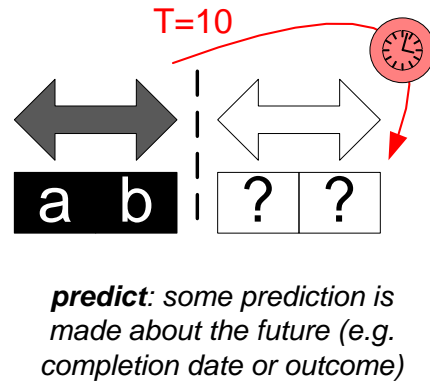
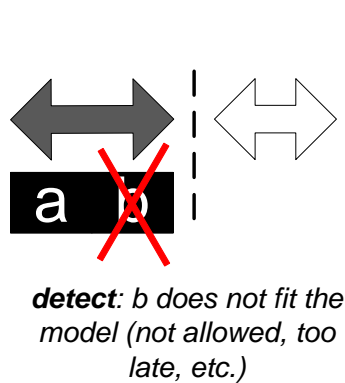
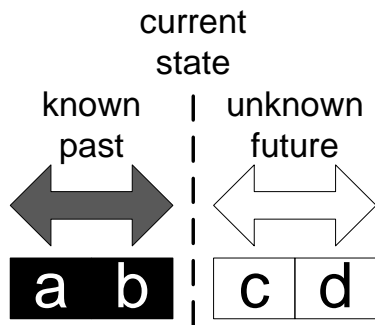


Navigation

- **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).



Operational support: online process mining using “pre mortem” event data



Let us focus one time

case id trace

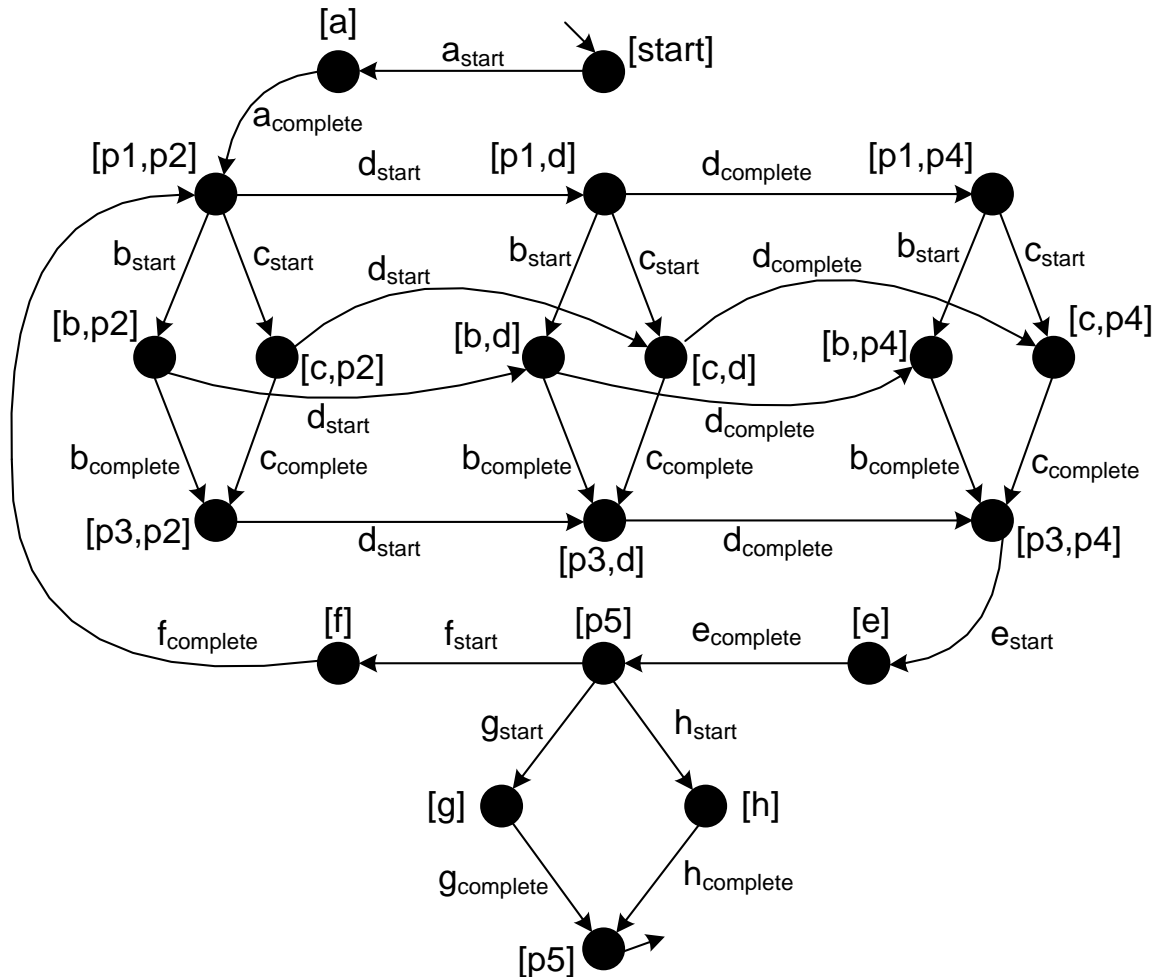
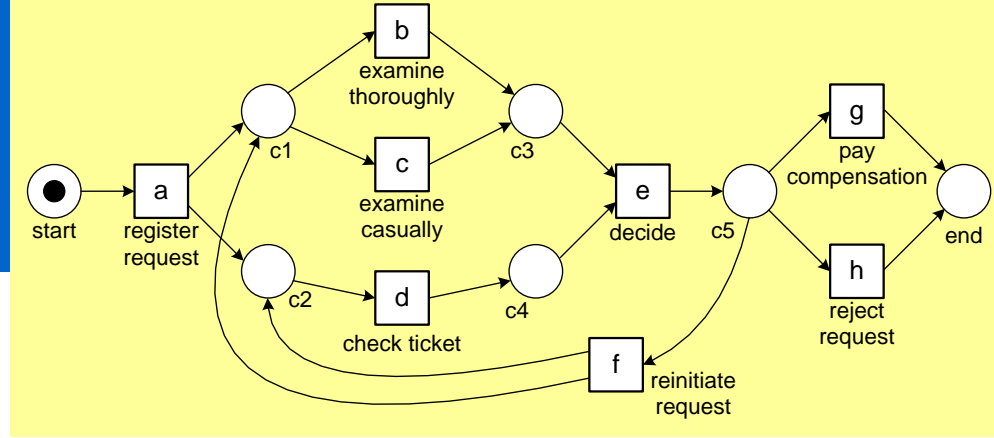
1 $\langle a_{start}^{12}, a_{complete}^{19}, b_{start}^{25}, d_{start}^{26}, b_{complete}^{32}, d_{complete}^{33}, e_{start}^{35}, e_{complete}^{40}, h_{start}^{50}, h_{complete}^{54} \rangle$

2 $\langle a_{start}^{17}, a_{complete}^{23}, d_{start}^{28}, c_{start}^{30}, d_{complete}^{32}, c_{complete}^{38}, e_{start}^{50}, e_{complete}^{59}, g_{start}^{70}, g_{complete}^{73} \rangle$

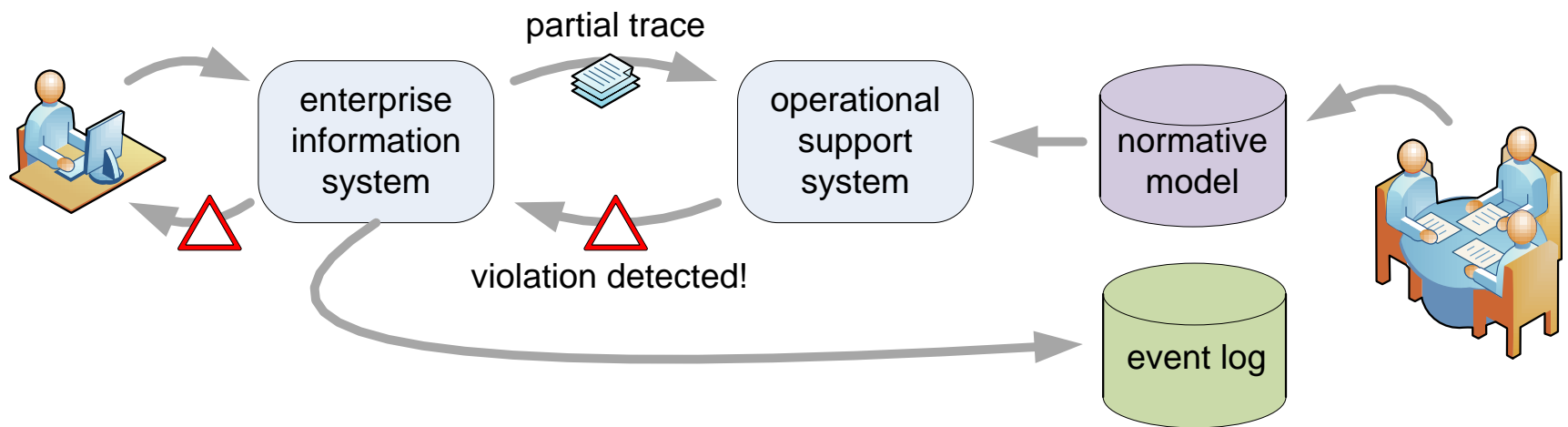
3 $\langle a_{start}^{25}, a_{complete}^{30}, c_{start}^{32}, c_{complete}^{35}, d_{start}^{35}, d_{complete}^{40}, e_{start}^{45}, e_{complete}^{50}, f_{start}^{50}, f_{complete}^{55},$
 $b_{start}^{60}, d_{start}^{62}, b_{complete}^{65}, d_{complete}^{67}, e_{start}^{80}, e_{complete}^{87}, g_{start}^{90}, g_{complete}^{98} \rangle$

... ...

Transition system (with start/complete)



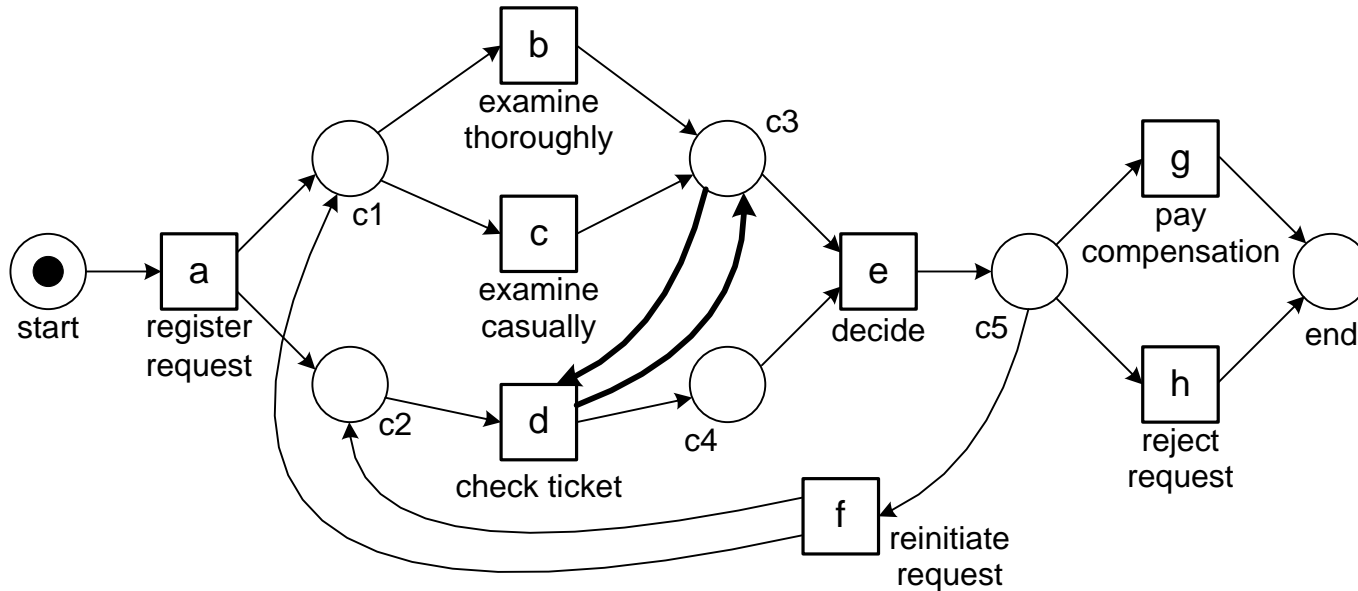
Operational support: Detect



Example

$\langle a_{start}^{12}, a_{complete}^{19}, b_{start}^{25}, d_{start}^{26} \rangle$

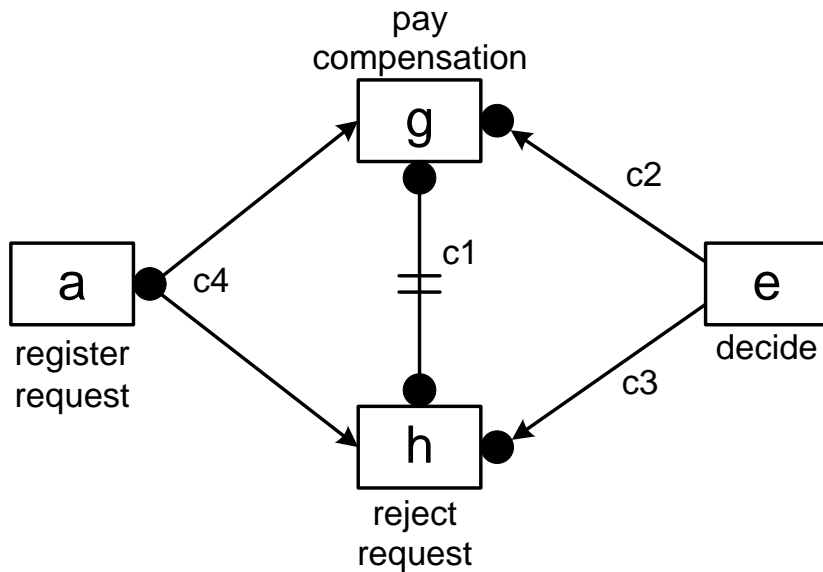
alert!!!!



case id trace

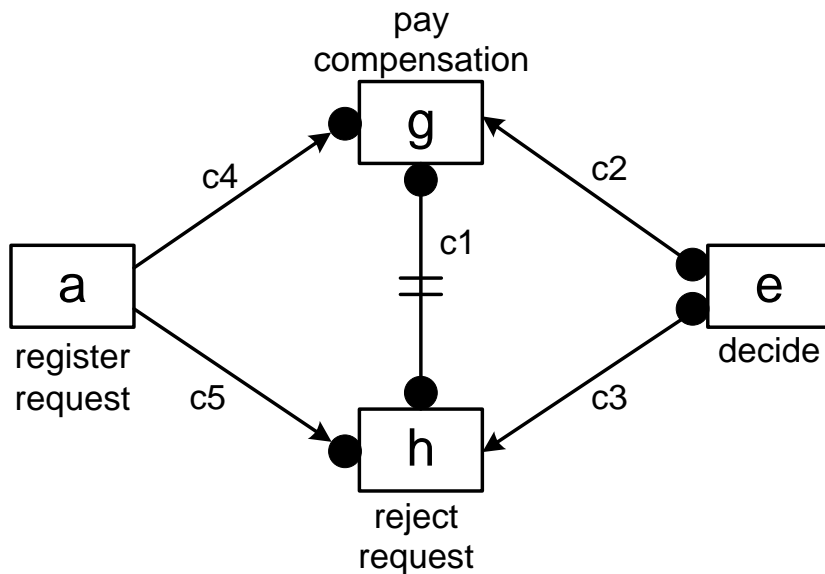
1	$\langle a_{start}^{12}, a_{complete}^{19}, b_{start}^{25}, d_{start}^{26}, b_{complete}^{32}, d_{complete}^{33}, e_{start}^{35}, e_{complete}^{40}, h_{start}^{50}, h_{complete}^{54} \rangle$
2	$\langle a_{start}^{17}, a_{complete}^{23}, d_{start}^{28}, c_{start}^{30}, d_{complete}^{32}, c_{complete}^{38}, e_{start}^{50}, e_{complete}^{59}, g_{start}^{70}, g_{complete}^{73} \rangle$
3	$\langle a_{start}^{25}, a_{complete}^{30}, c_{start}^{32}, c_{complete}^{35}, d_{start}^{35}, d_{complete}^{40}, e_{start}^{45}, e_{complete}^{50}, f_{start}^{50}, f_{complete}^{55}, b_{start}^{60}, d_{start}^{62}, b_{complete}^{65}, d_{complete}^{67}, e_{start}^{80}, e_{complete}^{87}, g_{start}^{90}, g_{complete}^{98} \rangle$
...	...

Declare specifications for detecting violations



- **Satisfied:** the LTL formula evaluates to true for the current partial trace.
- **Temporarily violated:** the LTL formula evaluates to false, however, there is a longer trace that evaluates to true.
- **Permanently violated:** the LTL formula evaluates to false for current trace and all its extensions

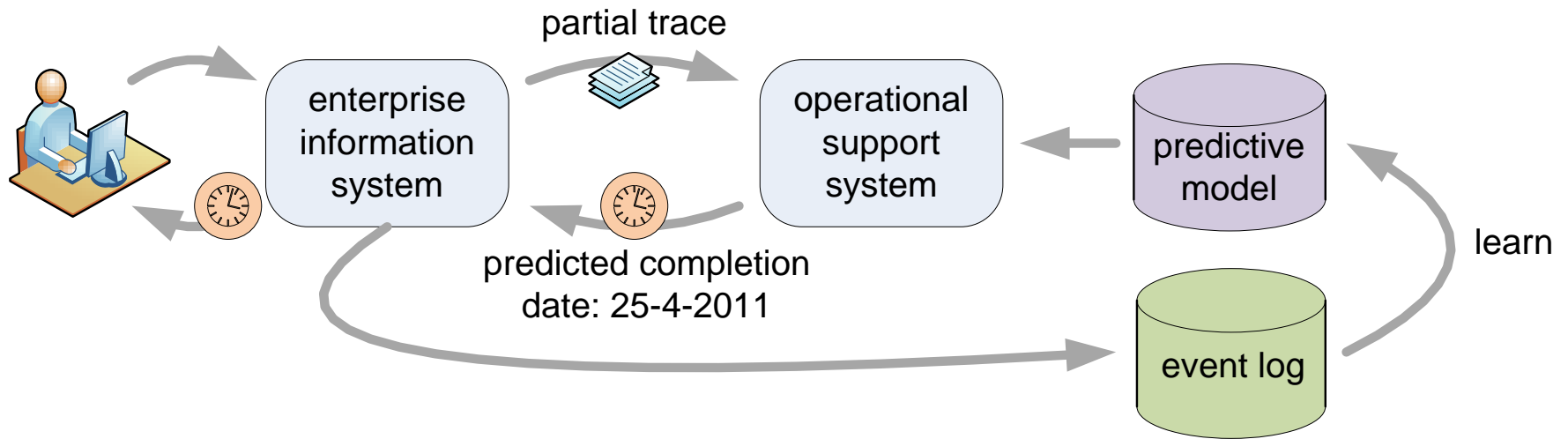
Conflicting constraints



- A Declare specification is **satisfied** for a case if all of its constraints are satisfied.
- A Declare specification is **temporarily violated** by a case if for the current partial trace at least one of the constraints is violated, however, there is a possible future in which all constraints are satisfied.
- A Declare specification is **permanently violated** by a case if no such future exists.

Note that c1, c2, and c3 imply that e cannot be executed without permanently violating the specification.

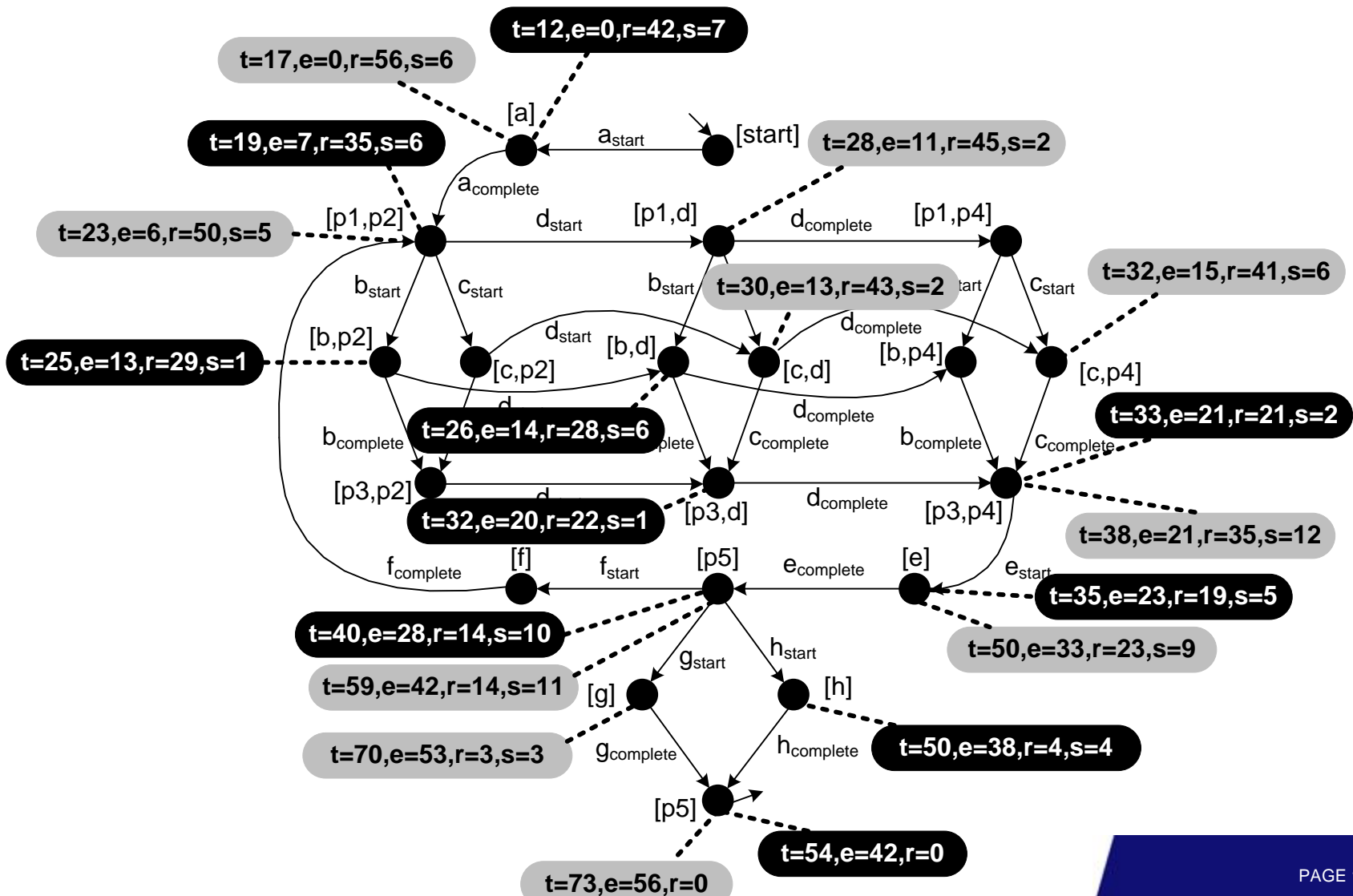
Operational support: Predict



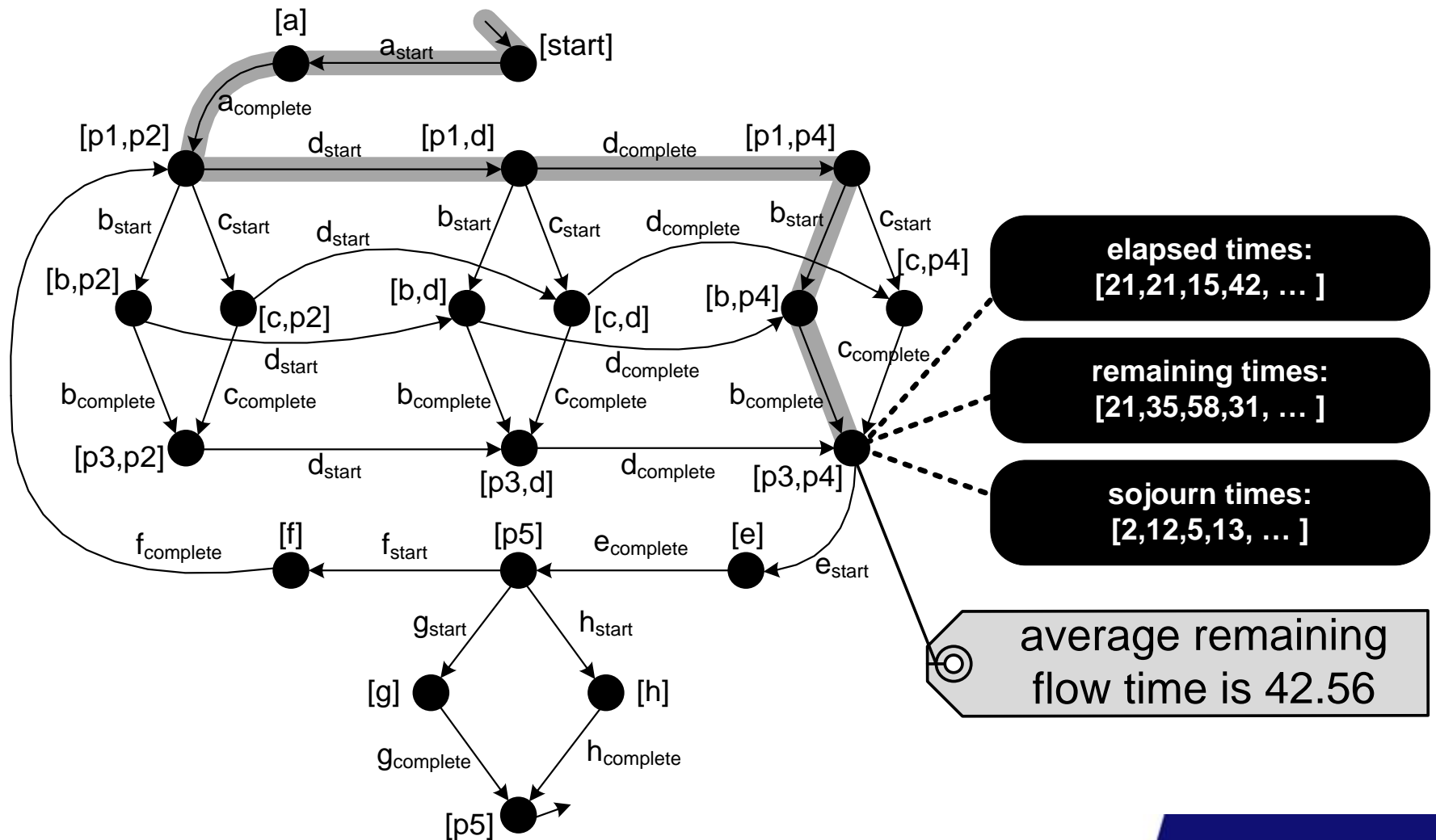
Examples of predictions

- the predicted remaining flow time is 14 days;
- the predicted probability of meeting the legal deadline is 0.72;
- the predicted total cost of this case is 4500 euro;
- the predicted probability that activity a will occur is 0.34;
- the predicted probability that person r will work on this case is 0.57;
- the predicted probability that a case will be rejected is 0.67; and
- the predicted total service time is 98 minutes.

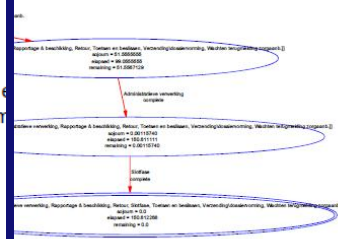
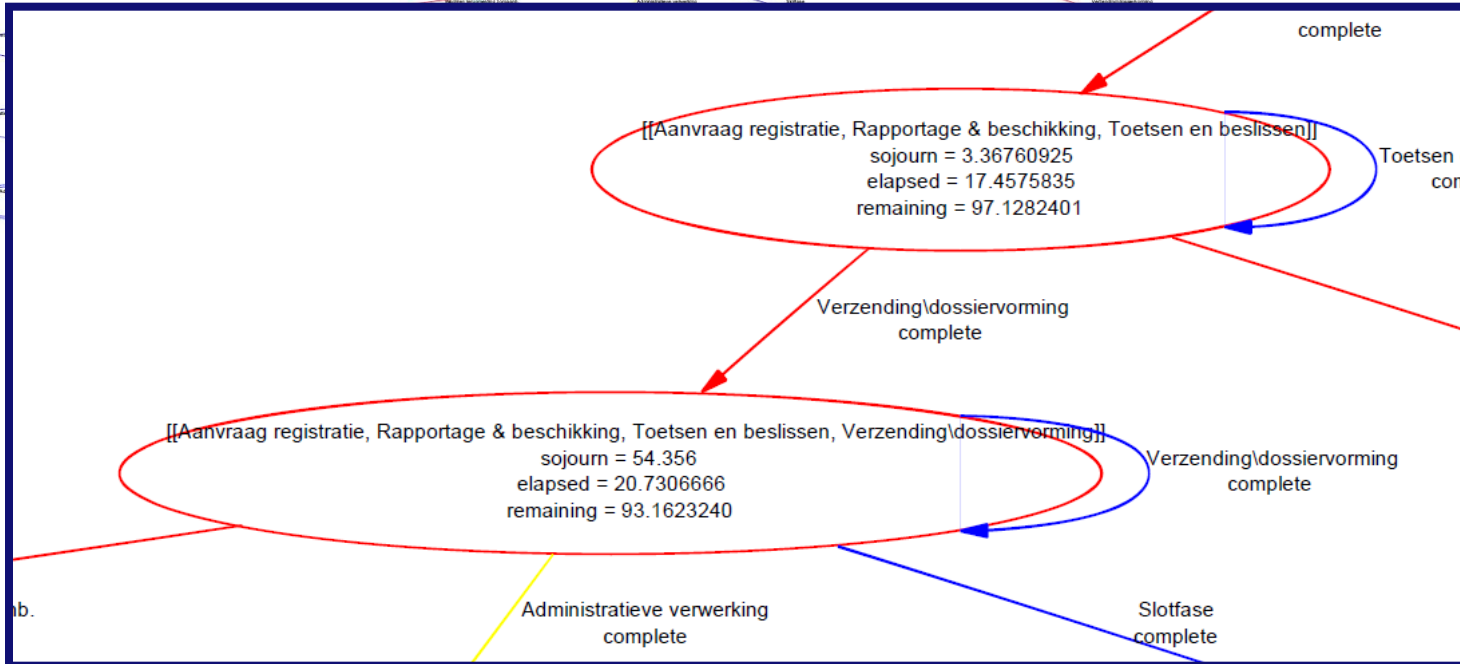
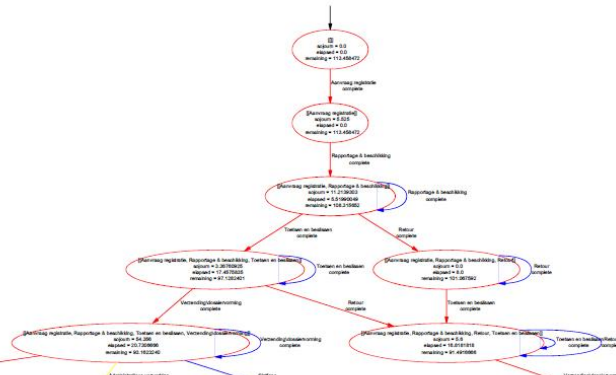
Annotated transition system



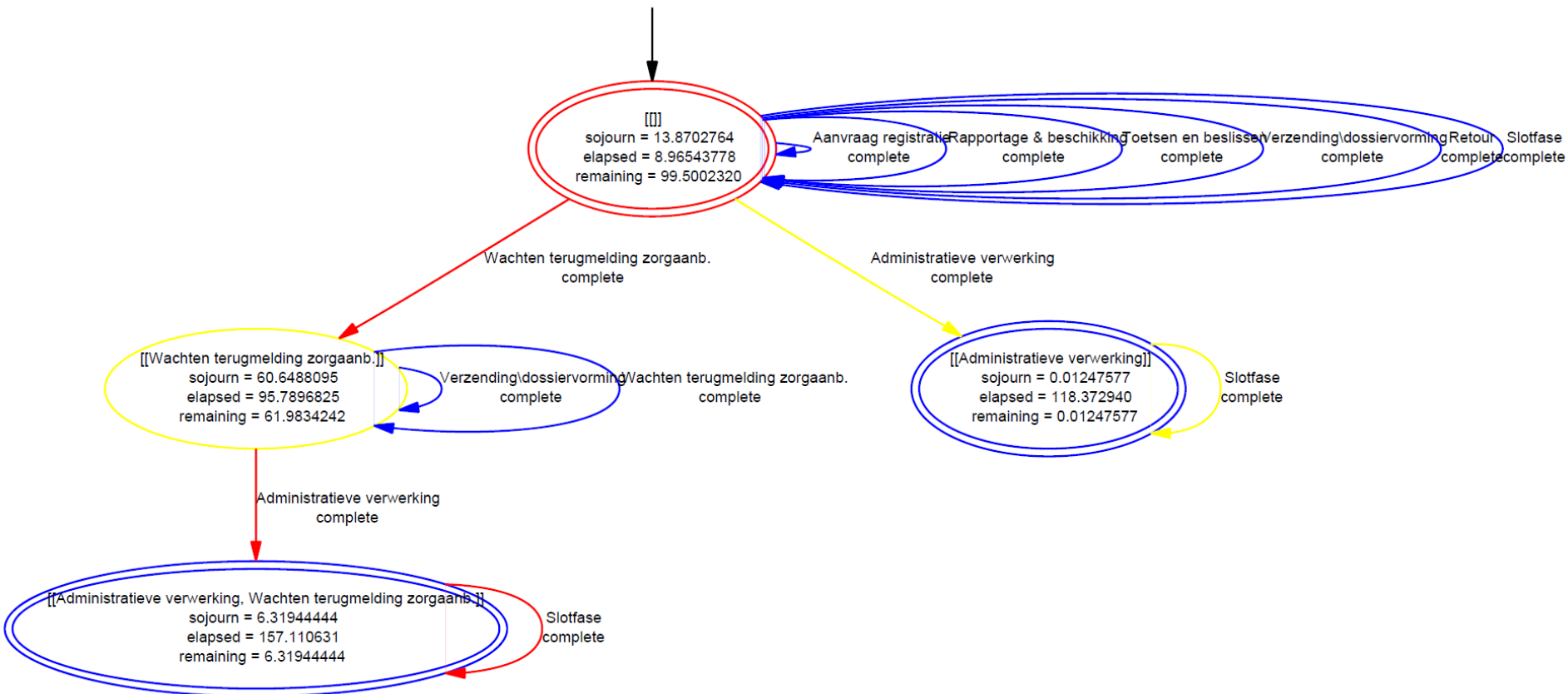
Collect results per state



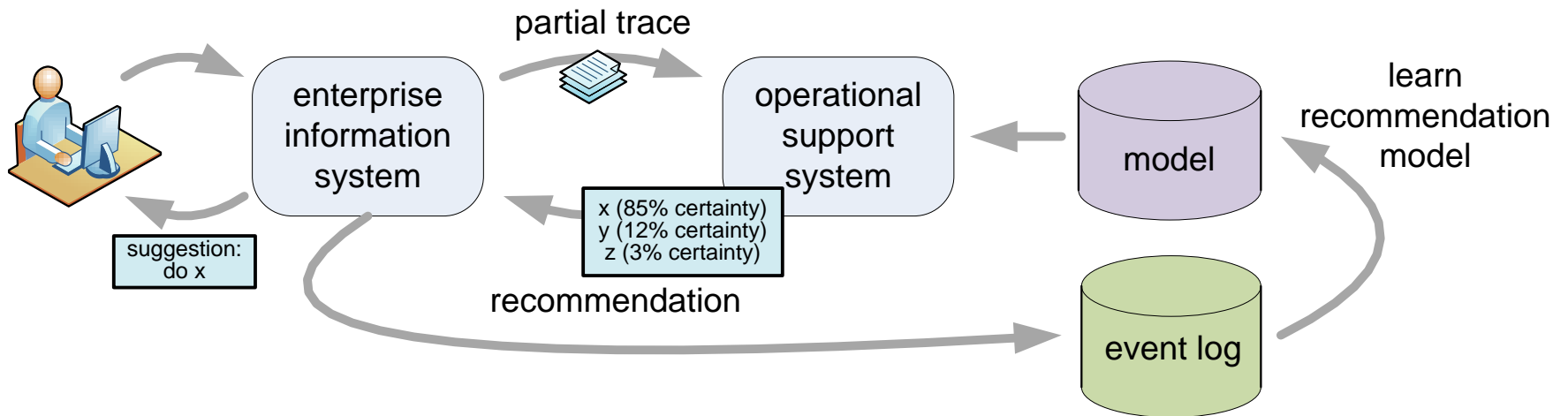
Example: Predicting the Remaining Processing Time in a Municipality



Same event log but a coarser abstraction



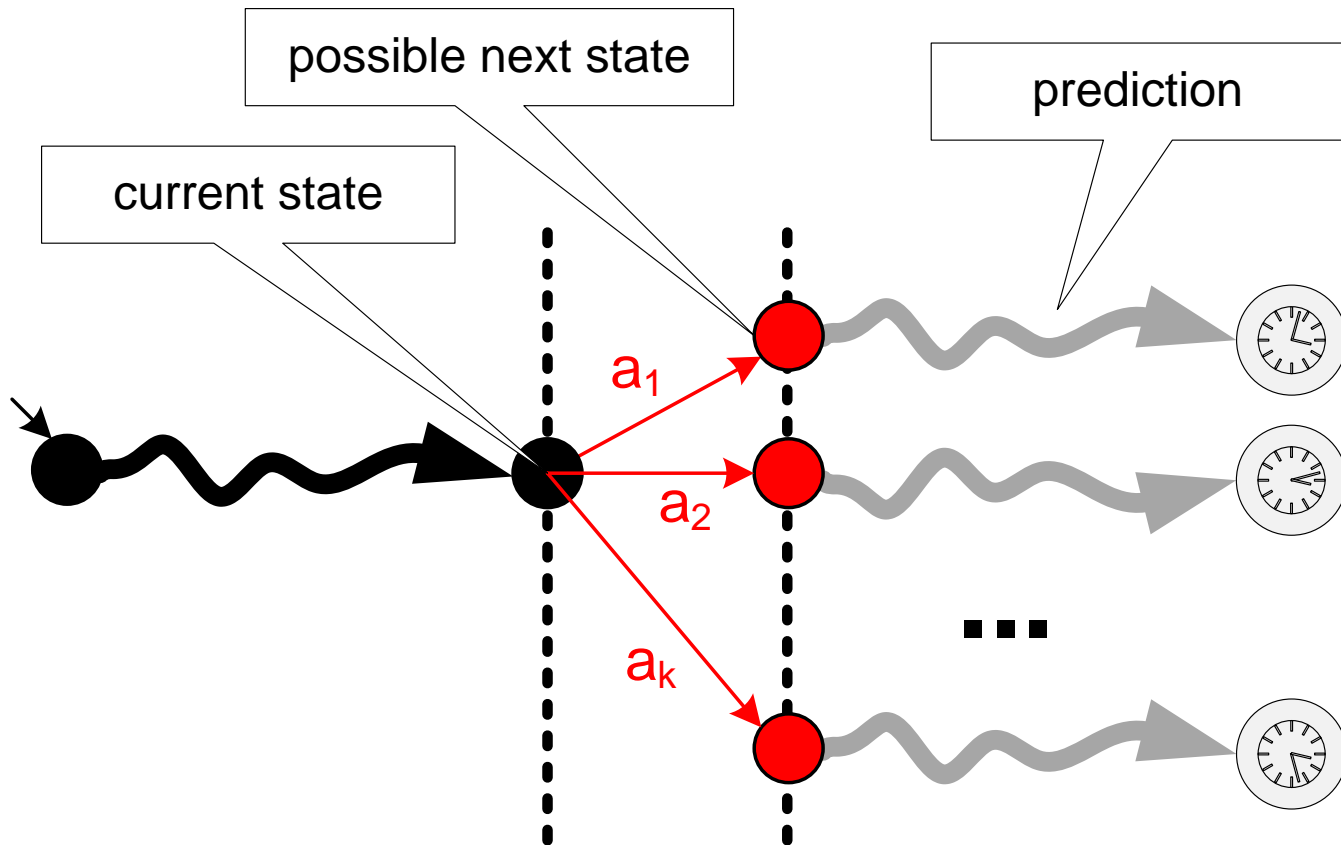
Operational support: Recommend



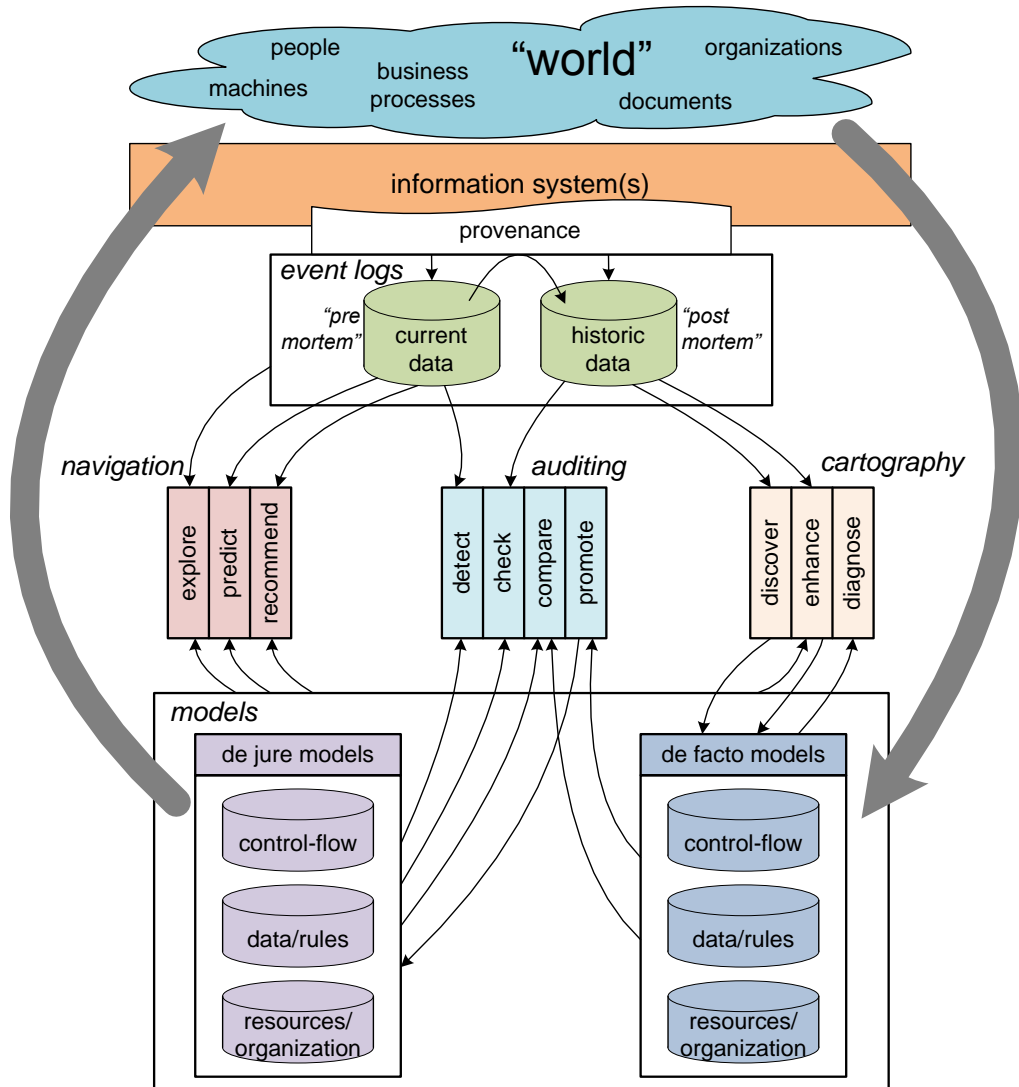
Recommend

- **Possible recommendations:**
 - next activity;
 - suitable resource; or
 - routing decision.
- **A recommendation is always given with respect to a specific goal.**
- **Examples of goals are:**
 - minimize the remaining flow time;
 - minimize the total costs;
 - maximize the fraction of cases handled within 4 weeks;
 - maximize the fraction of cases that is accepted; and
 - minimize resource usage.

Relation between prediction and recommendation

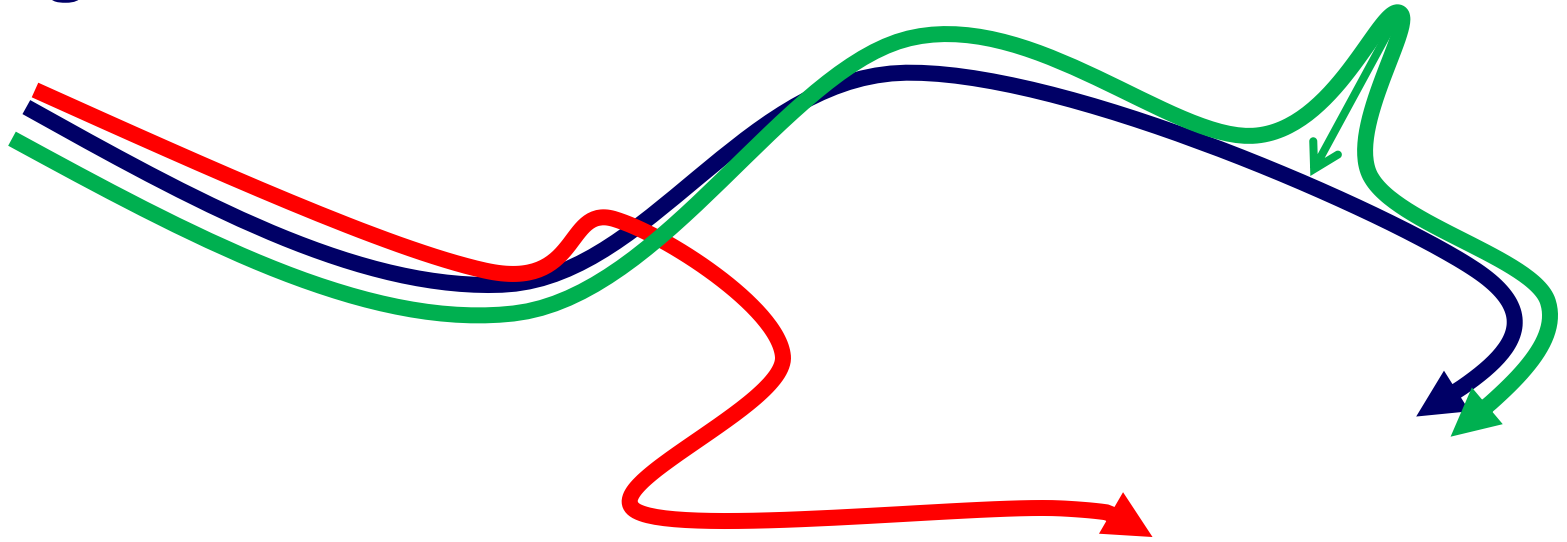


Process mining spectrum



Three Key Observations

#1 Alignments are essential!



- conformance checking to diagnose deviations
- squeezing reality into the model to do model-based analysis

move on
model

<i>a</i>	<i>c</i>	\gg	<i>d</i>	\gg	<i>f</i>	\gg
<i>a</i>	<i>c</i>	<i>b</i>	<i>d</i>	τ	\gg	<i>h</i>
<i>t1</i>	<i>t4</i>	<i>t3</i>	<i>t5</i>	<i>t7</i>		<i>t10</i>

move on
log

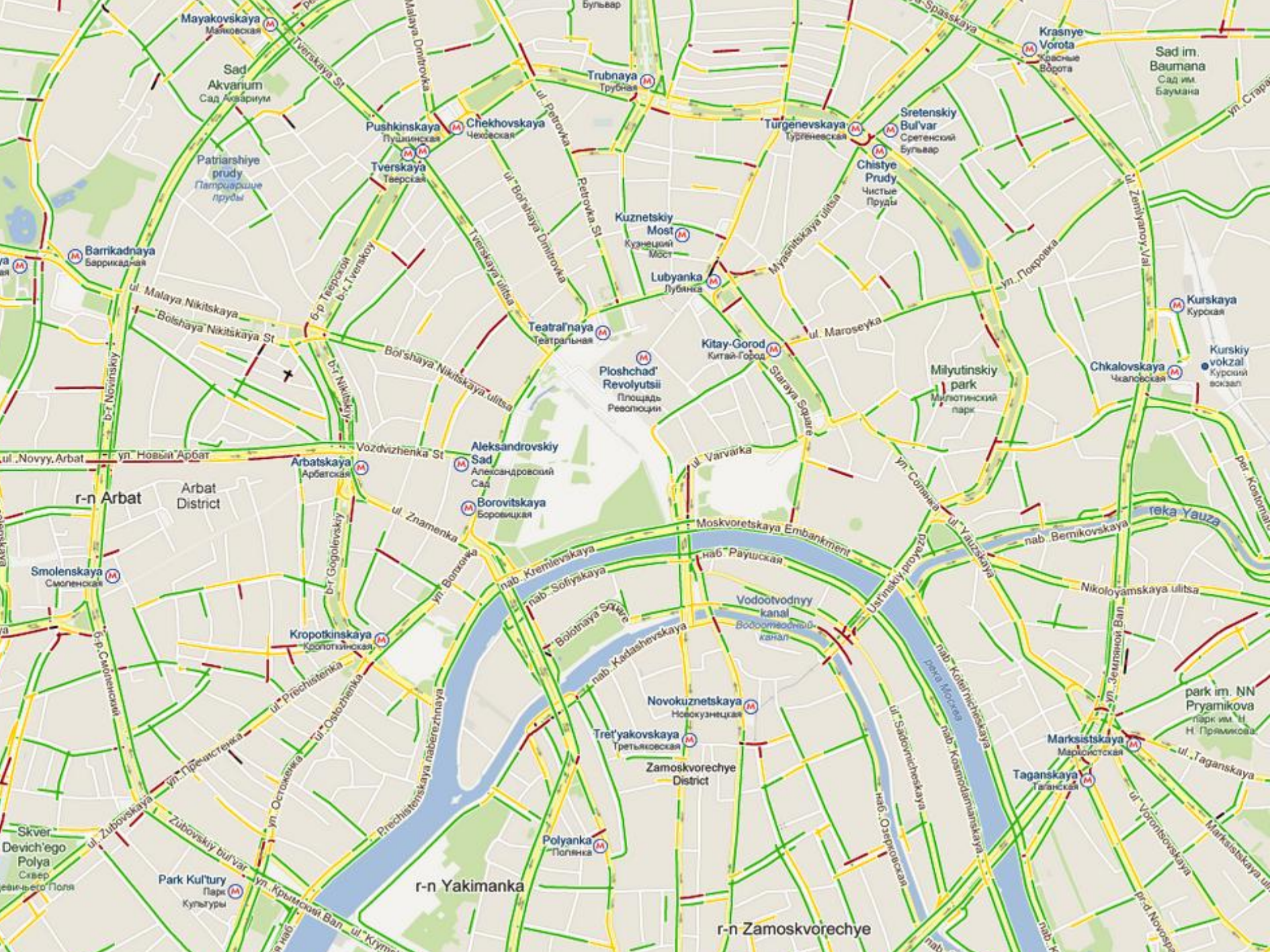
move on model
(harmless)

move on
model

#2 Models are like the glasses required to see and understand event data!



```
<Data>  
<Attribute name="Complicatie">-Asystolie</Attribute>  
<Attribute name="DatumComplicatie">2004-07-19T03:09.000+02:00</Attribute>  
<Attribute name="OpnameNummer">42914</Attribute>  
<Attribute name="TijdComplicatie">13:25</Attribute>  
<Attribute name="Initialen">mk</Attribute>  
</Data>  
<Event type="schedule"></Event type>  
<Timestamp>2004-07-19T03:09.000+02:00</Timestamp>  
<Originator>mke</Originator>  
</AuditTrailEntry>  
<Data>  
<Attribute name="Complicatie">Acute Tubulus</Attribute>  
<Attribute name="DatumComplicatie">2004-07-19T14:25:42.000+02:00</Attribute>  
<Attribute name="OpnameNummer">42914</Attribute>  
<Attribute name="TijdComplicatie">13:25</Attribute>  
<Attribute name="Initialen">abi</Attribute>  
</Data>  
<Event type="schedule"></Event type>  
<Timestamp>2004-07-19T14:25:42.000+02:00</Timestamp>  
<Originator>abi</Originator>  
</AuditTrailEntry>  
<Data>  
<Attribute name="Complicatie">Ischemische</Attribute>  
<Attribute name="DatumComplicatie">2004-07-19T14:26:00.000+02:00</Attribute>  
<Attribute name="OpnameNummer">42914</Attribute>  
<Attribute name="TijdComplicatie">13:26</Attribute>  
<Attribute name="Initialen">abi</Attribute>  
</Data>  
<Event type="schedule"></Event type>  
<Timestamp>2004-07-19T14:26:00.000+02:00</Timestamp>  
<Originator>abi</Originator>  
</AuditTrailEntry>
```

Mayakovskaya
Маяковская

Sad Akvarium
Сад Аквариум

Patriarshiye prudy
Патриаршие пруды

Barikadnaya
Барикадная

Bolshaya Nikitskaya St
Большая Никитская Ст

r-n Arbat
Арбатский район

Arbat District

Smolenskaya
Смоленская

Skver Devich'ego Polya
Сквер Девицкого Поля

Park Kultury
Парк Культуры

r-n Yakimanka
Якиманка район

r-n Zamoskvorechye
Замоскворечье район

Pushkinskaya
Пушкинская

Chekhovskaya
Чеховская

Tverskaya
Тверская

Trubnaya
Трубная

Turgenevskaya
Тургеневская

Sretenskiy Bulvar
Сретенский Бульвар

Chistye Prudy
Чистые Пруды

Kuznetskiy Most
Кузнецкий Мост

Lubyanka
Лубянка

Teatral'naya
Театральная

Ploshchad' Revolyutsii
Площадь Революции

Kitay-Gorod
Китай-Город

ul. Maroseyka
Ул. Маросейка

Milyutinskiy park
Милютинский парк

Kurskaya
Курская

Kurskiy vokzal
Курский вокзал

Chkalovskaya
Чкаловская

Arbatskaya
Арбатская

Aleksandrovskiy Sad
Александровский Сад

Borovitskaya
Боровицкая

ul. Varvarka
Ул. Варварка

Moskvoretskaya Embankment
Москворецкая набережная

nab. Raushskaya
Наб. Раушская

nab. Kremlevskaya
Наб. Кремлевская

nab. Sofiyuskaya
Наб. Софийская

nab. Katalashevskaya
Наб. Котловская

Novokuznetskaya
Новокузнецкая

Tret'yakovskaya
Третьяковская

Zamoskvorechye District

Polyanka
Полянка

Vodootvodnyy kanal
Водоотводный канал

ul. Solov'yeva
Ул. Соловьева

ul. Yauzskaya
Ул. Яузская

nab. Bemkovskaya
Наб. Бемковская

Nikoloyamskaya ulitsa
Николаямская улица

park im. NN Pryamukova
Парк им. Н.Н. Прямукиной

Marksistskaya
Марьинская

Taganskaya
Таганская

ul. Taganskaya
Ул. Таганская

ul. Voronovskaya
Ул. Воронцовская

ul. Voronovskaya
Ул. Воронцовская

ul. Voronovskaya
Ул. Воронцовская

ul. Voronovskaya
Ул. Воронцовская

Wil M. P. van der Aalst

Process Mining

Discovery, Conformance and Enhancement of Business Processes

More and more information about business processes is recorded by information systems in the form of so-called "event logs". Despite the omnipresence of such data, most organizations diagnose problems based on fiction rather than facts. Process mining is an emerging discipline based on process model-driven approaches and data mining. It not only allows organizations to fully benefit from the information stored in their systems, but it can also be used to check the conformance of processes, detect bottlenecks, and predict execution problems.

Wil van der Aalst delivers the first book on process mining. It aims to be self-contained while covering the entire process mining spectrum from process discovery to operational support. In Part I, the author provides the basics of business process modeling and data mining necessary to understand the remainder of the book. Part II focuses on process discovery as the most important process mining task. Part III moves beyond discovering the control flow of processes and highlights conformance checking, and organizational and time perspectives. Part IV guides the reader in successfully applying process mining in practice, including an introduction to the widely used open-source tool ProM. Finally, Part V takes a step back, reflecting on the material presented and the key open challenges.

Overall, this book provides a comprehensive overview of the state of the art in process mining. It is intended for business process analysts, business consultants, process managers, graduate students, and BPM researchers.

Features and Benefits:

- First book on process mining, bridging the gap between business process modeling and business intelligence.
- Written by one of the most influential and most-cited computer scientists and the best-known BPM researcher.
- Self-contained and comprehensive overview for a broad audience in academia and industry.
- The reader can put process mining into practice immediately due to the applicability of the techniques and the availability of the open-source process mining software ProM.

Computer Science

ISBN 978-3-642-19344-6



9 783642 193446

► springer.com

van der Aalst



Process Mining

Wil M. P. van der Aalst

Process Mining

Discovery, Conformance and
Enhancement of Business Processes

www.processmining.org

www.win.tue.nl/ieeetfpm/

 Springer