Process Mining: Beyond Business Intelligence

Gartner Business Process Management Summit, February 2009, London

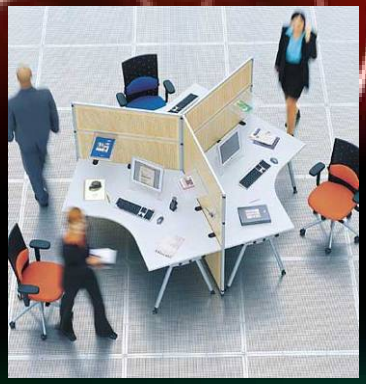
prof.dr.ir. Wil van der Aalst

www.processmining.org

ProM

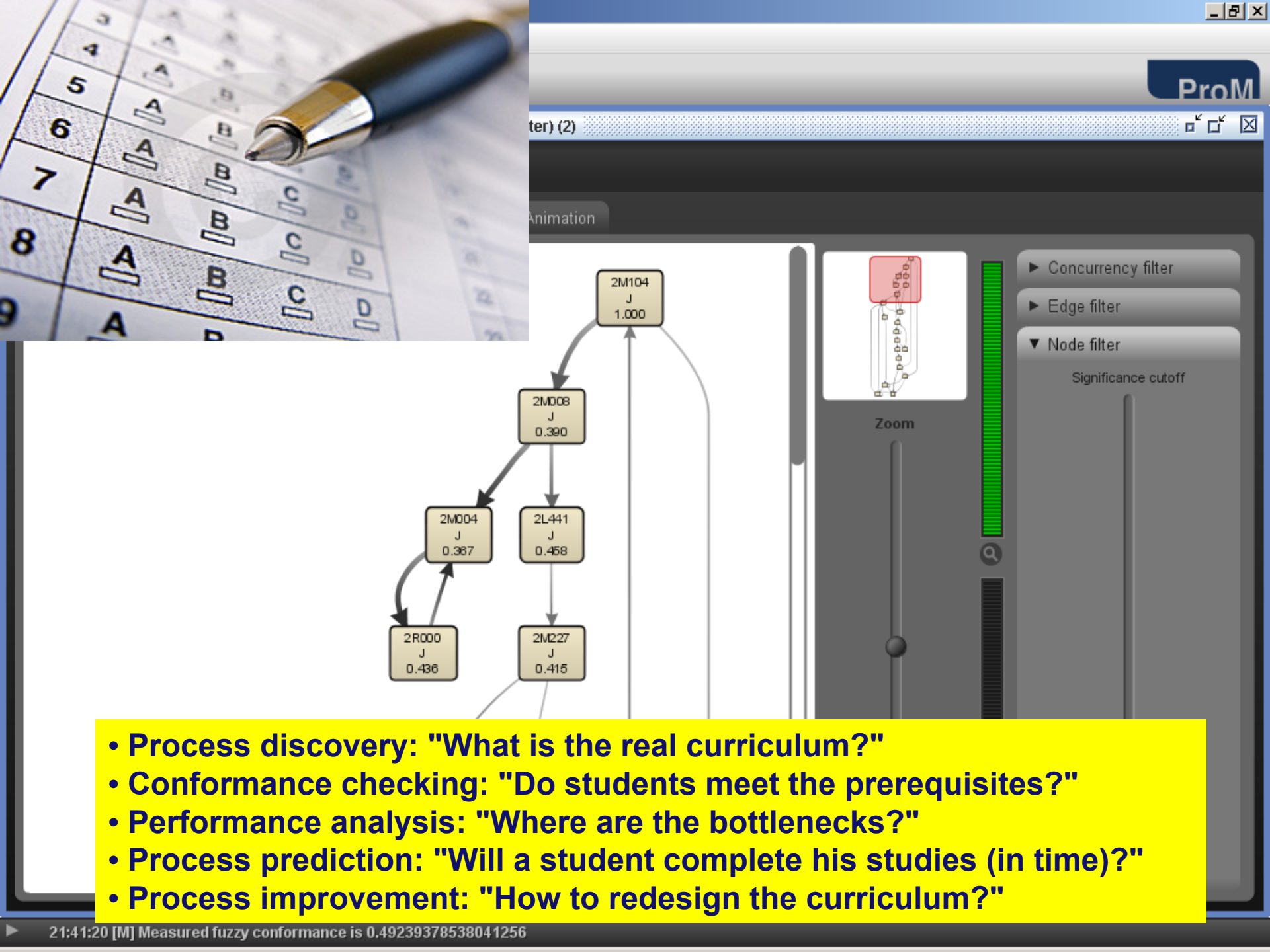
TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts



Process Mining

- Process discovery: "What is really happening?"
- Conformance checking: "Do we do what was agreed upon?"
- Performance analysis: "Where are the bottlenecks?"
- Process prediction: "Will this case be late?"
- Process improvement: "How to redesign this process?"
- Etc.



- **Process discovery: "What is the real curriculum?"**
- **Conformance checking: "Do students meet the prerequisites?"**
- **Performance analysis: "Where are the bottlenecks?"**
- **Process prediction: "Will a student complete his studies (in time)?"**
- **Process improvement: "How to redesign the curriculum?"**

Outline

- Trends in BPM
- Process Mining: The Basics
  - Input data
  - Discovery
  - Conformance
  - Software support
- Process Mining: Applications
- Process Mining: TomTom for Business Processes
- Conclusion

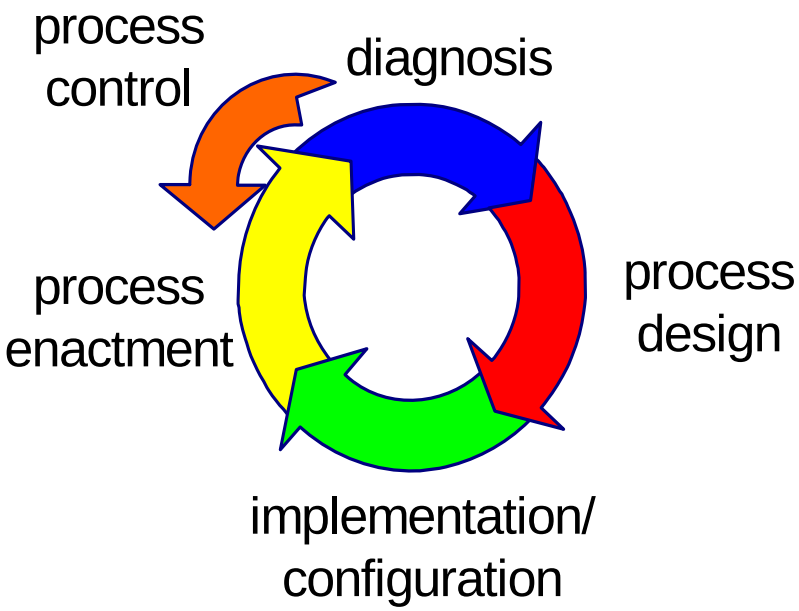
Trends in BPM

Technische Universiteit
Eindhoven
University of Technology

Where innovation starts

History

- **The first workflow management systems (called "office automation systems") were implemented in seventies, cf. Petri-net-based systems such as Officetalk (Xerox Parc, Skip Ellis) and SCOOP (Wharton, Michael Zisman).**
- **Mid nineties: "explosion" of workflow products.**
- **Shift from workflow automation to business process management.**




MS Workflow Foundation
Global 360 BPM Suite
YAWL
FileNet
InConcert
Fujitsu Interstage
Axxerion
BWise
Software AG/webMethods
XPDL
IBM WebSphere
casewise
COSA
BPEL
UML
Savvion BusinessManager
ADs
BPM|one
TIBCO iProcess Suite
jBPM
EPCs
FlowConnect
SAP Workflow
BPMN
Pegasystems SmartBPM Suite
Ensemble
Bizagi
TeamWARE
Oracle BPEL
Promatis
BiZZdesigner
Ultimus BPM Suite

Workflow Patterns Initiative

- Initiative started in late 90-ties.
- Collections:
  - 43 control-flow patterns (process/routing)
  - 40 data patterns
  - 43 resource patterns (work distr. and organization)
  - exception, flexibility, service interaction, ... patterns
- Frequently used as a tool in selection processes.
- Influenced standards (BPMN, BPEL, etc.) and systems.
- See www.workflowpatterns.com (+/- 500 unique visitors per day)

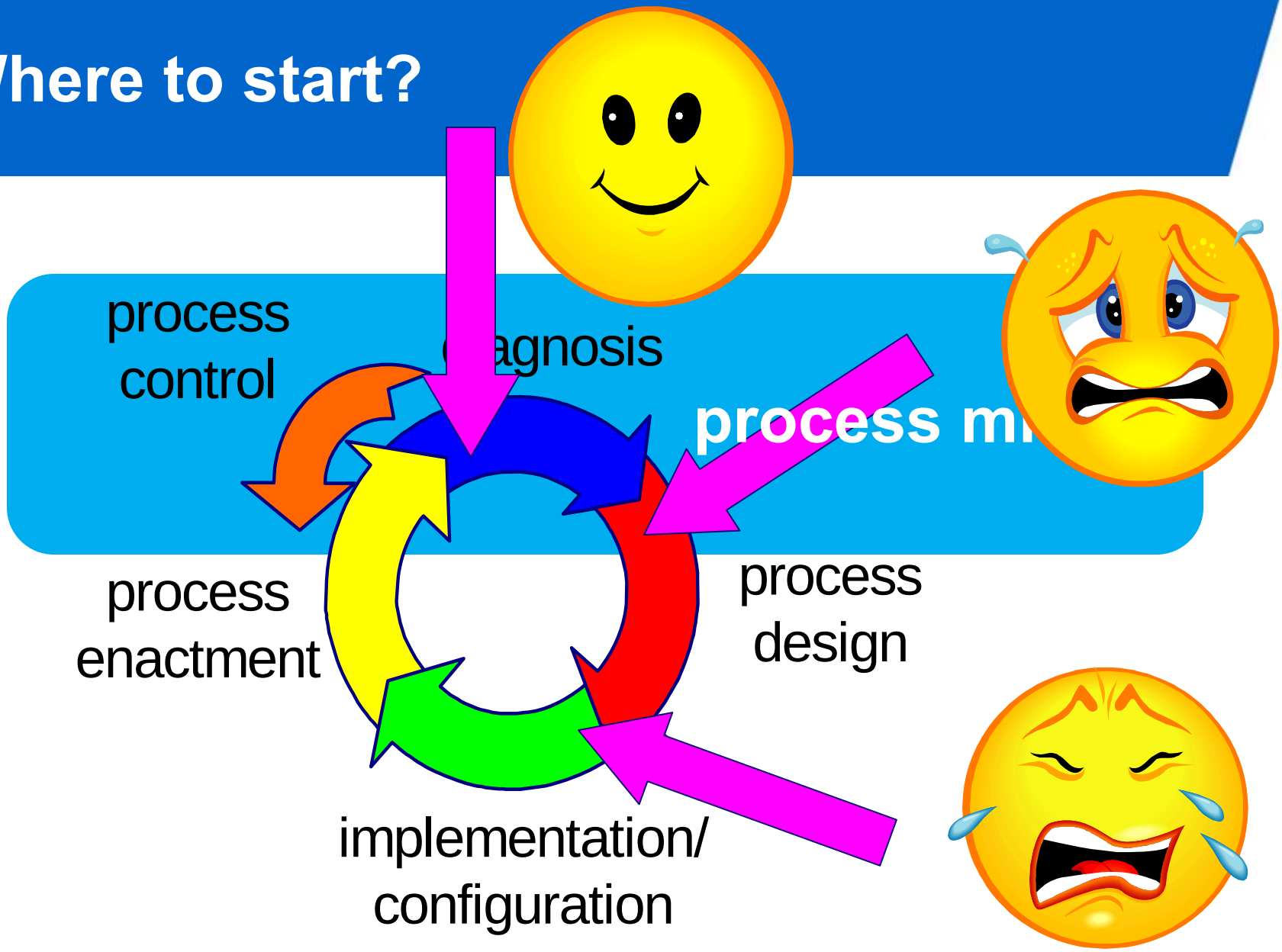
Problem is NOT the automation of structured processes!

Alignment
(Avoiding PowerPoint reality)

Ensuring compliance

Supporting flexibility

Where to start?

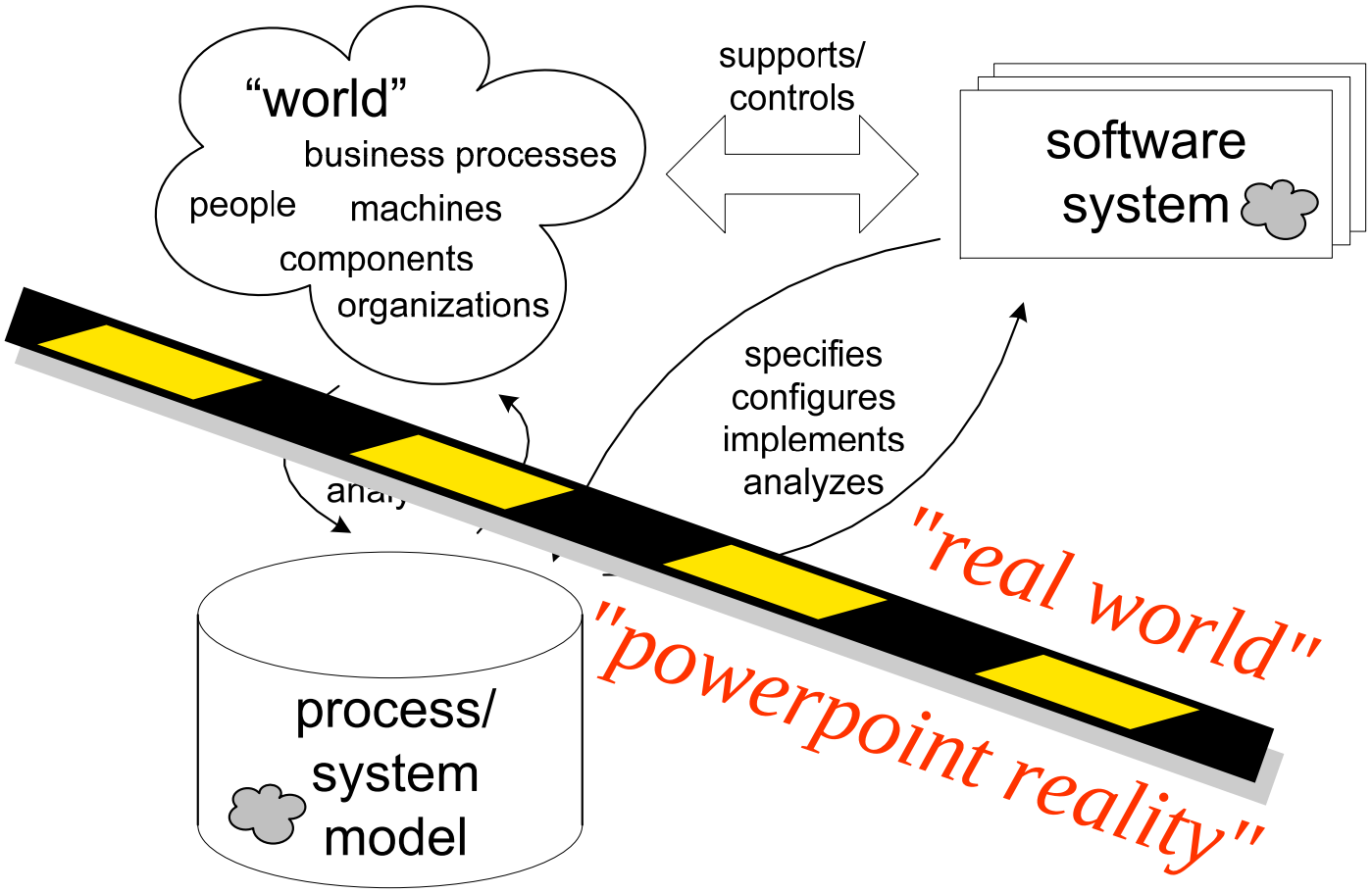


Process Mining: The Basics

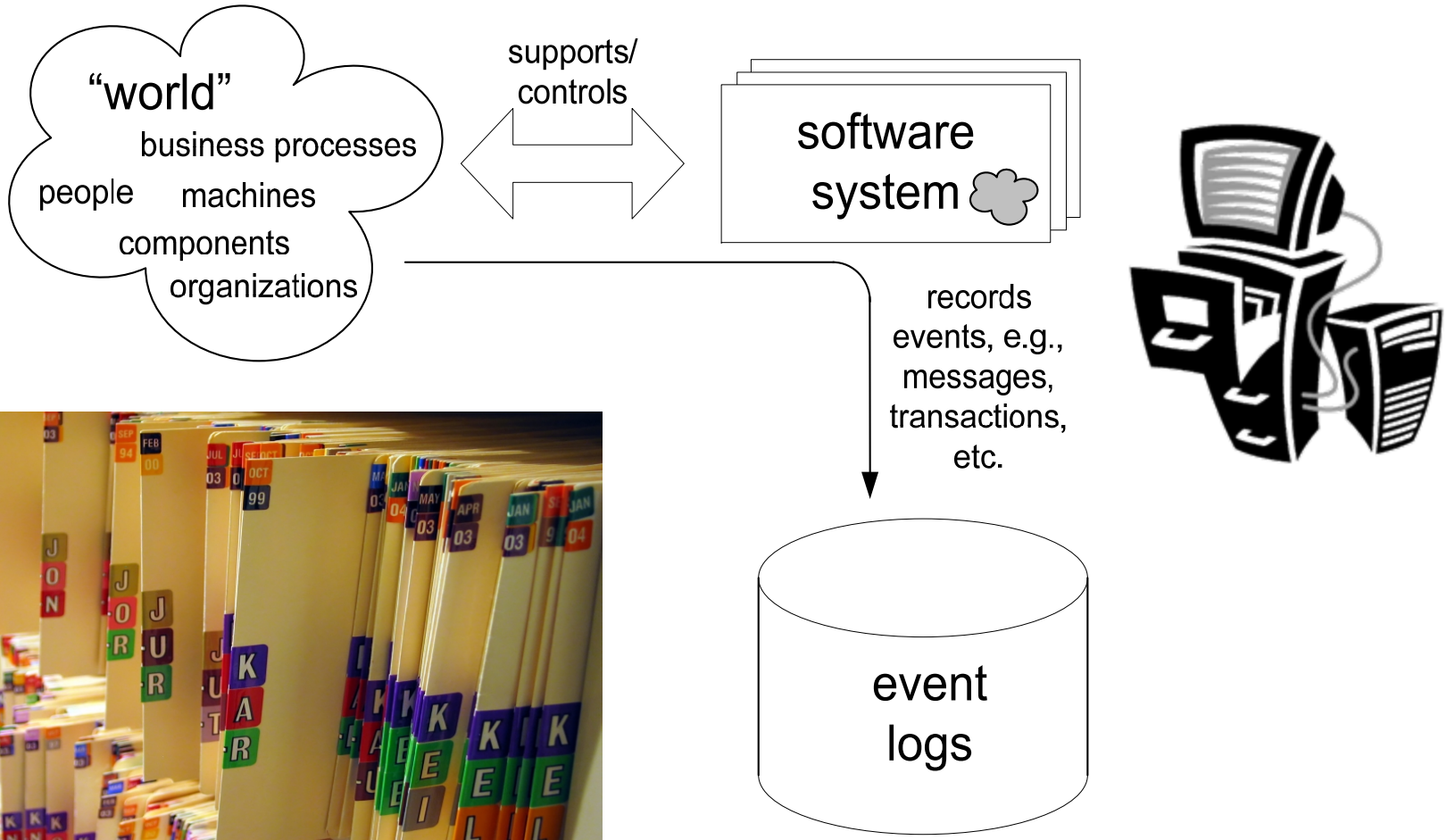
TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts

Role of models



Event logs are a reflection of reality



Examples:

staffware. THE POWER OF PROCESS

IBM

FileNet An IBM Company

océ

Pallas Athena

PHILIPS sense and simplicity

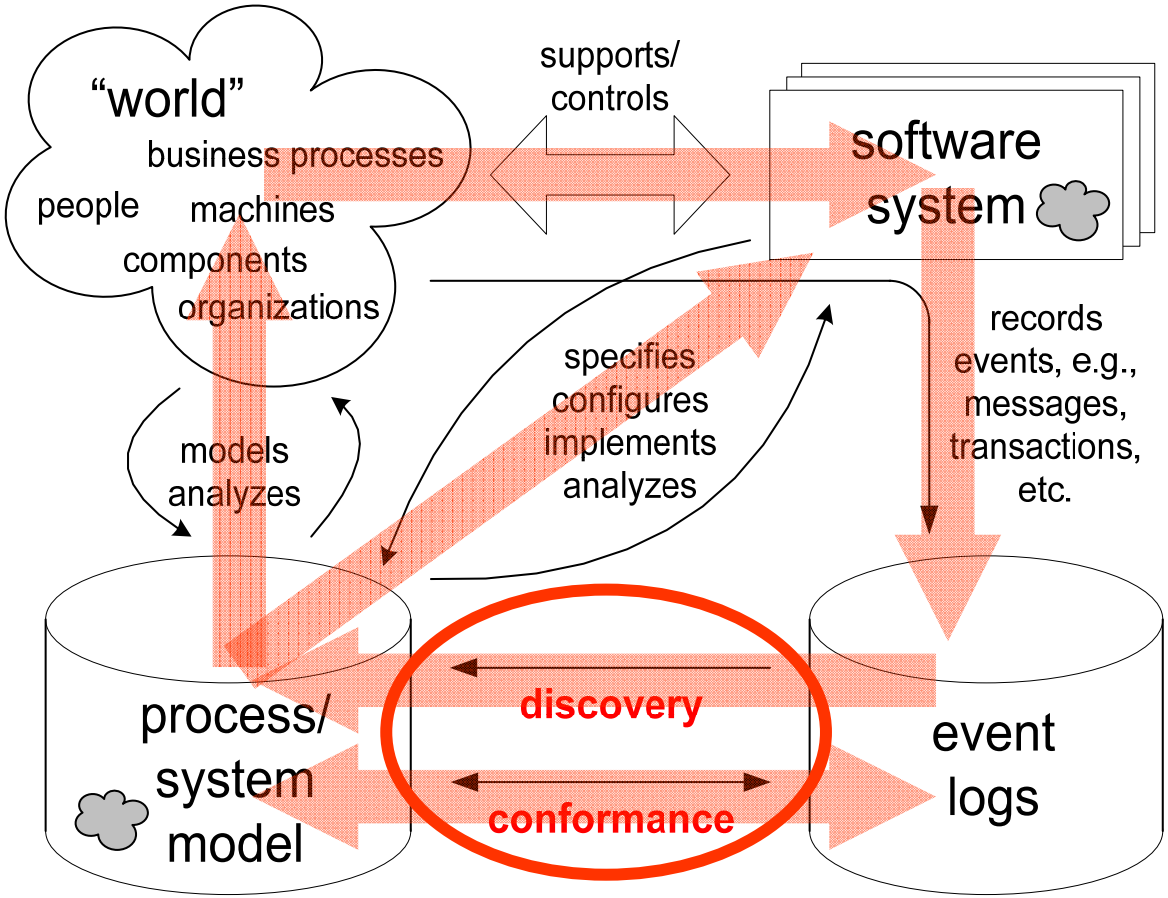
SAP

ORACLE

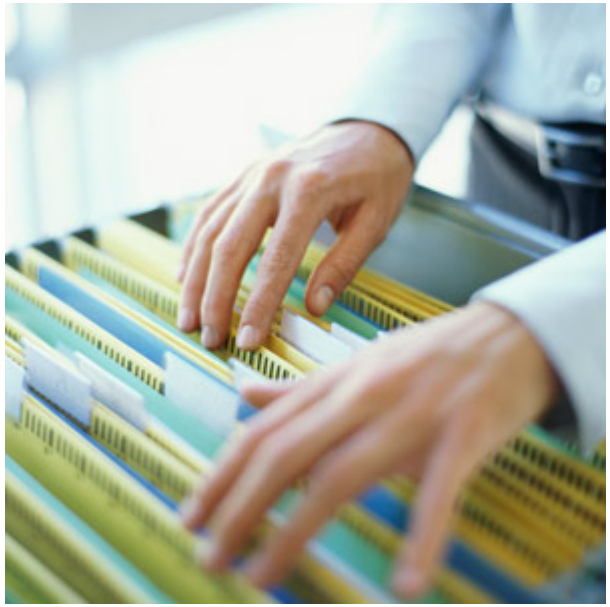
ASML

WebSphere. software

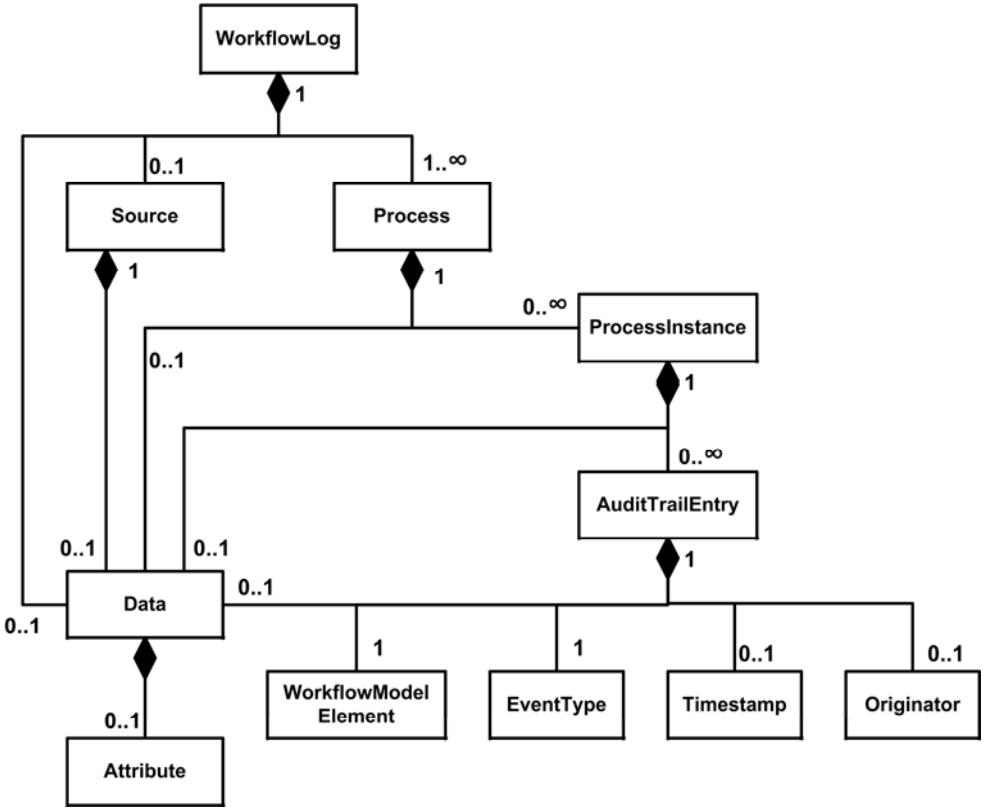
Process mining: Linking events to models



Starting point: event logs

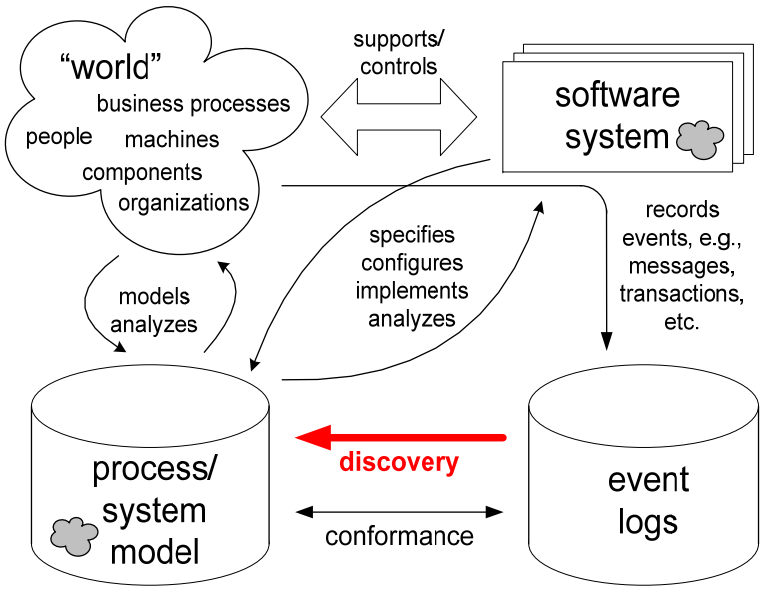


event logs, audit trails, databases, message logs, etc.



unified event log (MXML)

Discovery

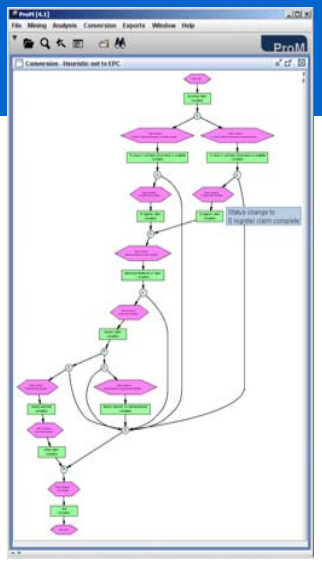

"world"
business processes
people
machines
components
organizations
supports/
controls
software
system
records
events, e.g.,
messages,
transactions,
etc.
specifies
configures
implements
analyzes
models
analyzes
process/
system
model
discovery
conformance
event
logs

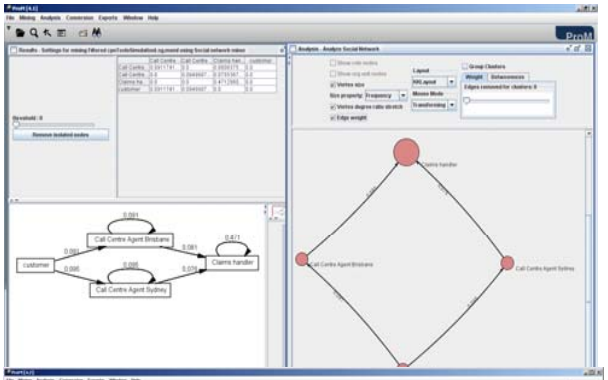
What to discover?

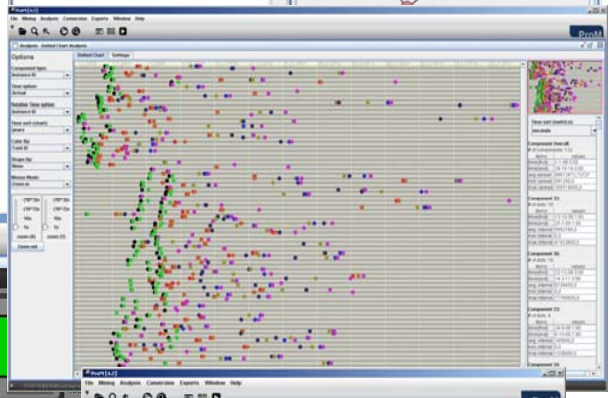
- process models (Petri nets, EPCs, BPMN, etc.),
- organizational models,
- social networks,
- sequence diagrams,
- business rules,
- bottlenecks,
- simulation models,
- etc.

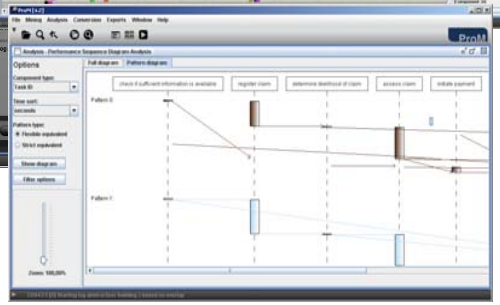
i.e., beyond "slice and dice" and showing KPIs on a dashboard ...

MXML Log
- **instances:** 3512
- **audit trail entries:** 46138









ProM supports +40 types of model discovery!

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Heuristics miner

incoming claim (complete) 3512

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Conversion - Heuristic net to Petri net

incoming claim complete

B check if sufficient information is available complete

S check if sufficient information is available complete

B register claim complete

S register claim complete

determine likelihood of claim complete

assess claim complete

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

ProM

Conversion - Heuristic net to EPC

incoming claim complete

B check if sufficient information is available complete

S check if sufficient information is available complete

B register claim complete

S register claim complete

Status change to S register claim complete

determine likelihood of claim complete

ProM [4.1]

File Mining Analysis Conversion Exports Window Help

ProM

Conversion - Labeled WF net to YAWL model

incoming claim complete Task

[default]

true [1]

B check if sufficient information is available complete Task

S check if sufficient information is available complete Task

true [0]

[default]

B register claim complete Task

S register claim complete Task

determine likelihood of claim complete Task

true [0]

assess claim complete Task

[default]

true [0]

advise claimant on reimbursement complete Task

initiate payment complete Task

close claim complete Task

true [0]

[default]

end complete Task

true [0]

3512

ProM [4.1]

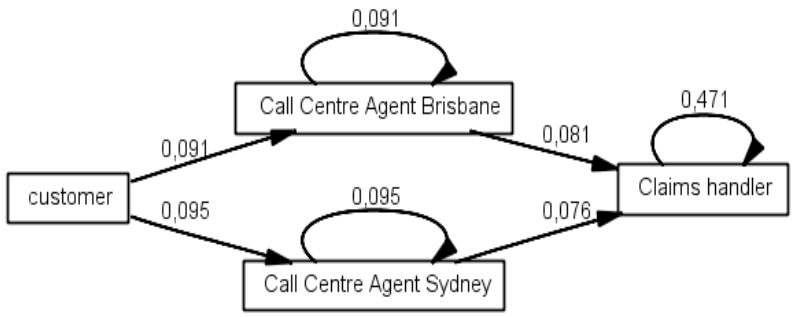
File Mining Analysis Conversion Exports Window Help

Results - Settings for mining Filtered cpnToolsSimulationLog.mxml using Social network miner

	Call Centre ...	Call Centre ...	Claims han...	customer
Call Centre...	0.0911741...	0.0	0.0808375...	0.0
Call Centre...	0.0	0.0949907...	0.0755367...	0.0
Claims ha...	0.0	0.0	0.4712960...	0.0
customer	0.0911741...	0.0949907...	0.0	0.0

threshold : 0

Remove isolated nodes



Zoom

Analysis - Analyze Social Network

Show role nodes
Show org unit nodes
Vertex size
Size property: Frequency
Vertex degree ratio stretch
Edge weight

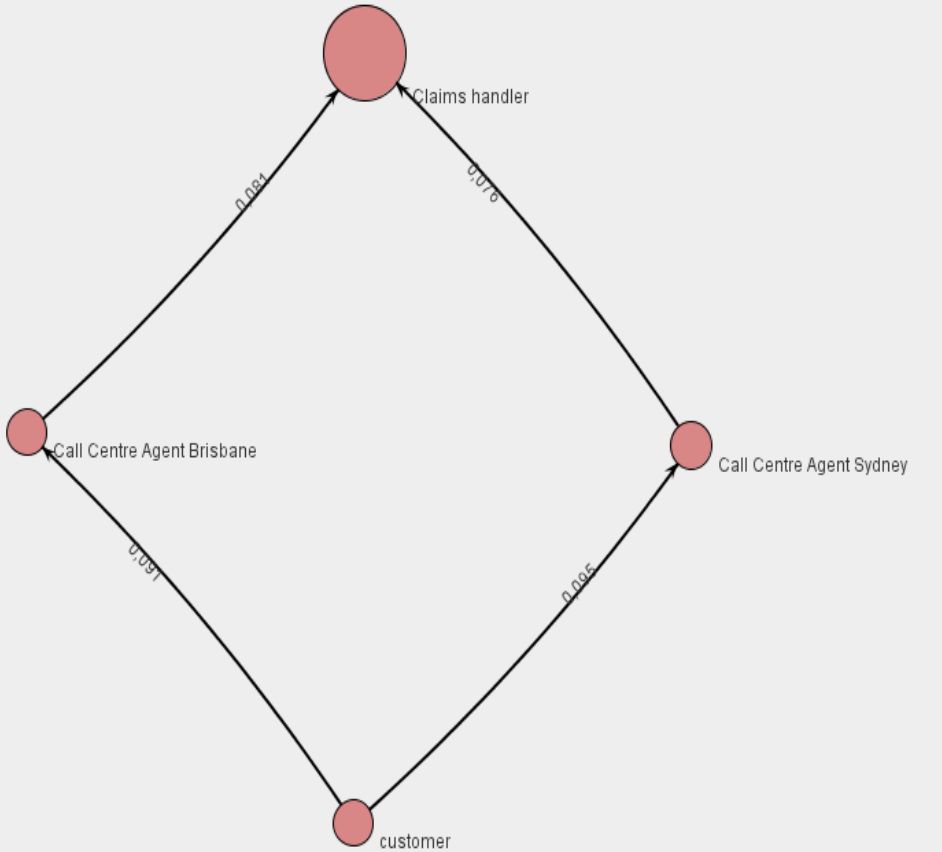
Layout: KKLayout

Mouse Mode: Transforming

Group Clusters

Weight | Betweenness

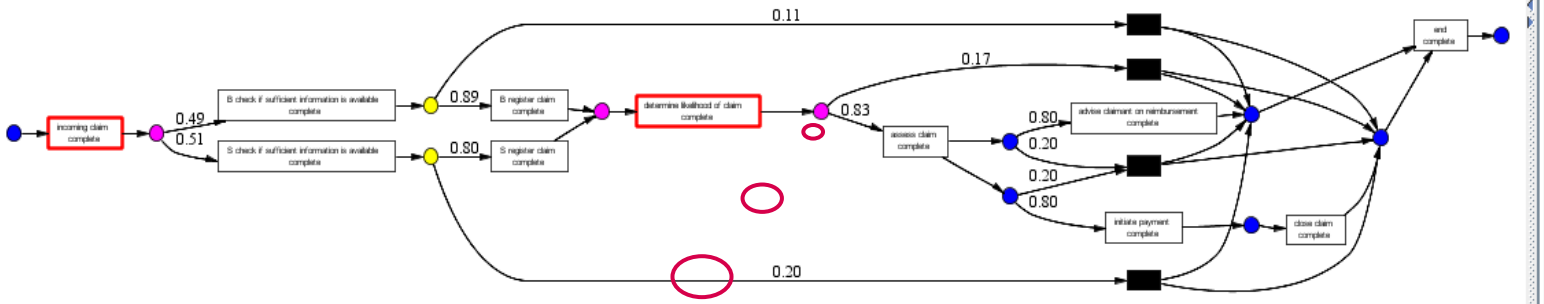
Edges removed for clusters: 0



ProM [4.1]

File Mining Analysis Conversion Exports Window Help

Analysis - Performance Analysis with Petri net



bottle-necks

throughput time

flow time from A to B

Process information:

Total number selected:

3512 cases

Number fitting:

3512 cases

Arrival rate:

0,12 cases per second

	Throughput time (seconds)
avg	11115,54
min	0,0
max	40704,0
stdev	8906,98
fast 25...	1379,19
slow 2...	23817,24
norma...	9632,87

Change Percentages | Export Time-Metrics

Performance information of the selected transitions:

Frequency: 2950 cases

	Time in between (seconds)
avg	12248,87
min	53,0
max	39706,0
stdev	8381,14

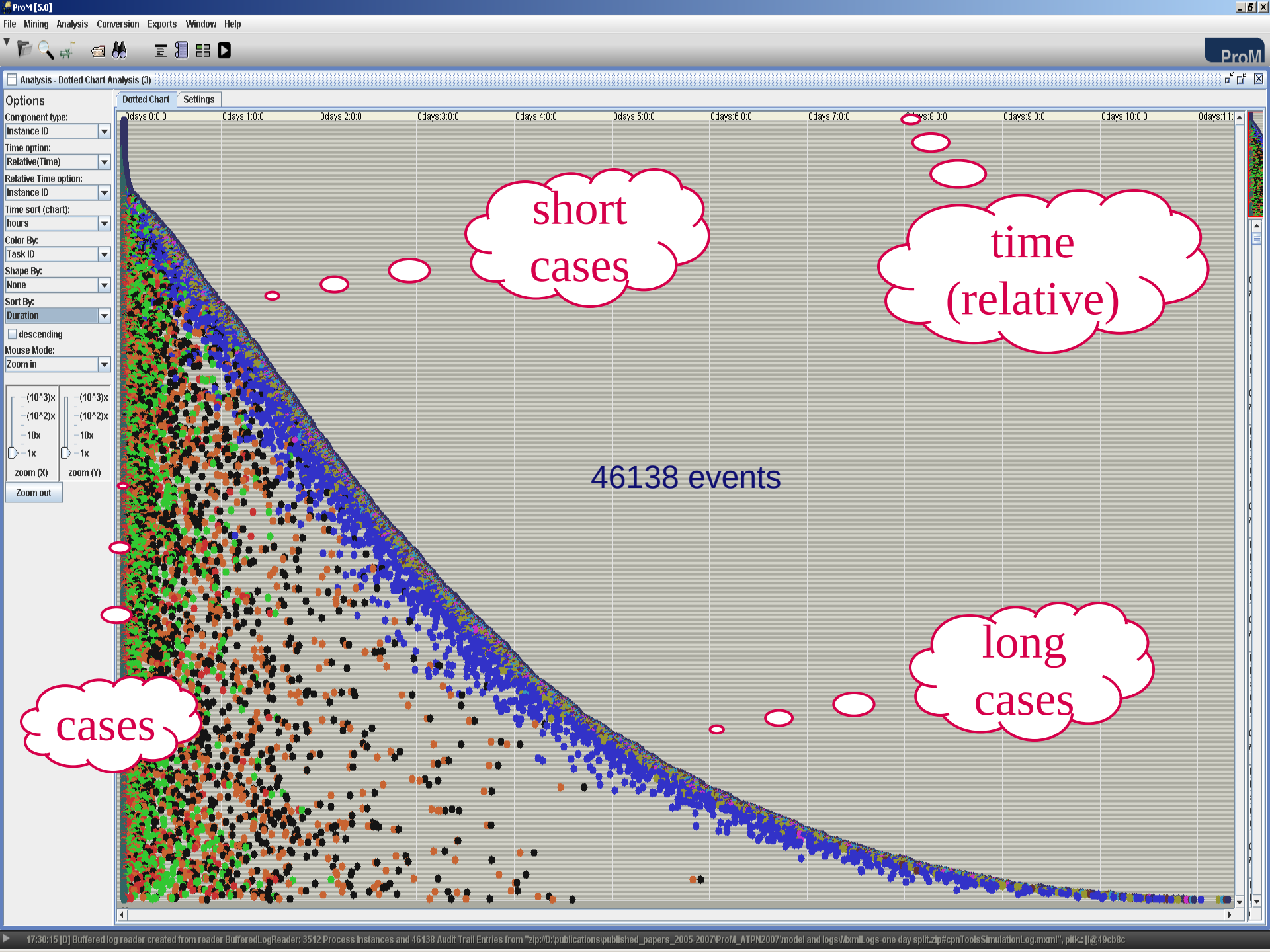
Waiting time:
High
Medium
Low
Settings

Selected:

Transition - incoming claim c...

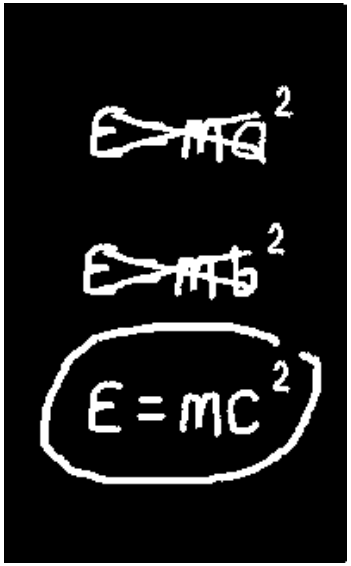
and:

Transition - determine likeliho...


ProM [5.0]
File Mining Analysis Conversion Exports Window Help
Analysis - Dotted Chart Analysis (3)
Options
Component type:
Instance ID
Time option:
Relative(Time)
Relative Time option:
Instance ID
Time sort (chart):
hours
Color By:
Task ID
Shape By:
None
Sort By:
Duration
descending
Mouse Mode:
Zoom in
(10^3)x
(10^2)x
10x
1x
zoom (X)
zoom (Y)
Zoom out
Dotted Chart
Settings
0days:0:0:0
0days:1:0:0
0days:2:0:0
0days:3:0:0
0days:4:0:0
0days:5:0:0
0days:6:0:0
0days:7:0:0
0days:8:0:0
0days:9:0:0
0days:10:0:0
0days:11:
46138 events
short cases
time (relative)
long cases
cases

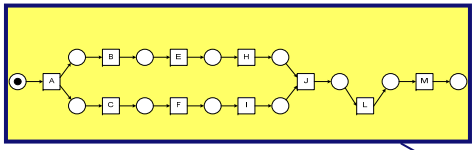
A bit of theory: Process discovery techniques

- Algorithmic techniques
  - Alpha miner
  - Alpha+, Alpha++, Alpha#
  - Heuristic miner
  - Multi phase miner
  - ...
- Genetic process mining
- Region-based process mining
  - State-based regions
  - Language based regions

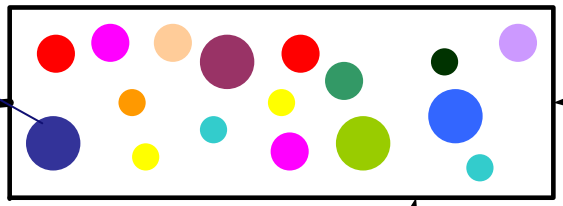


cf. www.processmining.org for an overview

Example: Genetic Mining

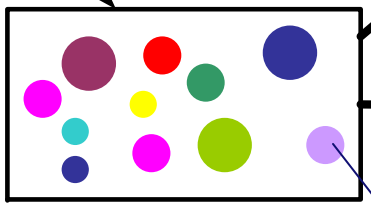


1. initial population



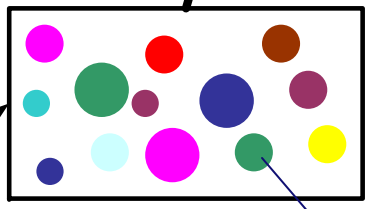
2. fitness test

3. select best parents



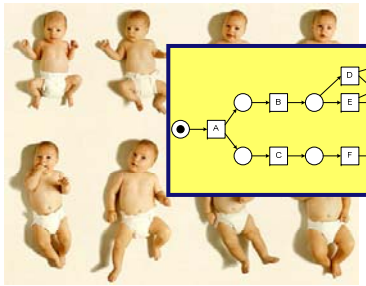
4. crossover

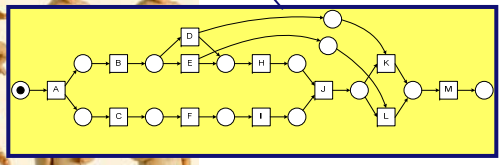
5. children

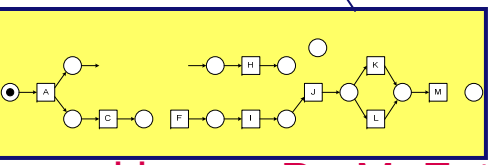


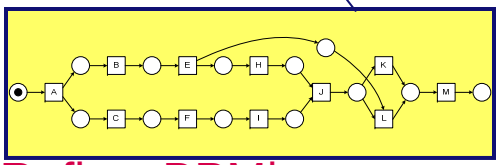
6. mutation

7. new population



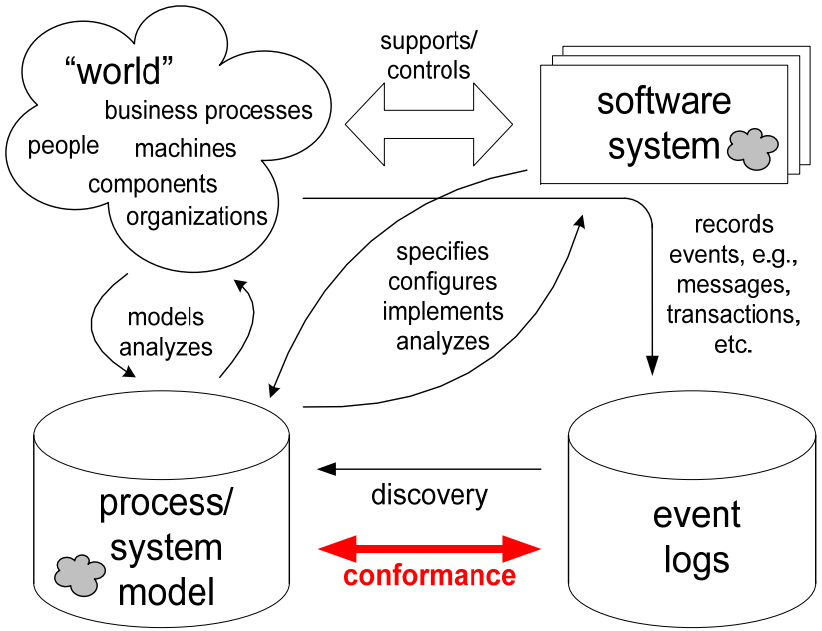






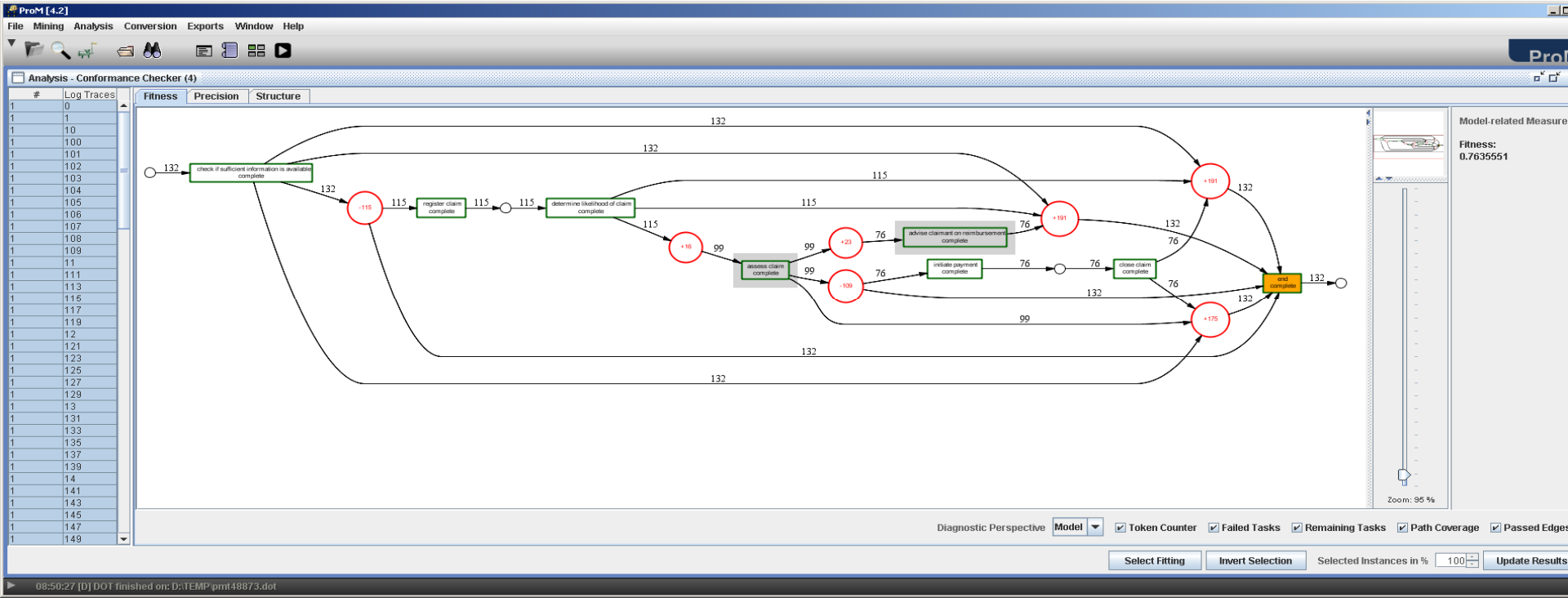
used in e.g. ProM, Futura Reflex, BPM|one

Conformance Checking



Conformance Checking

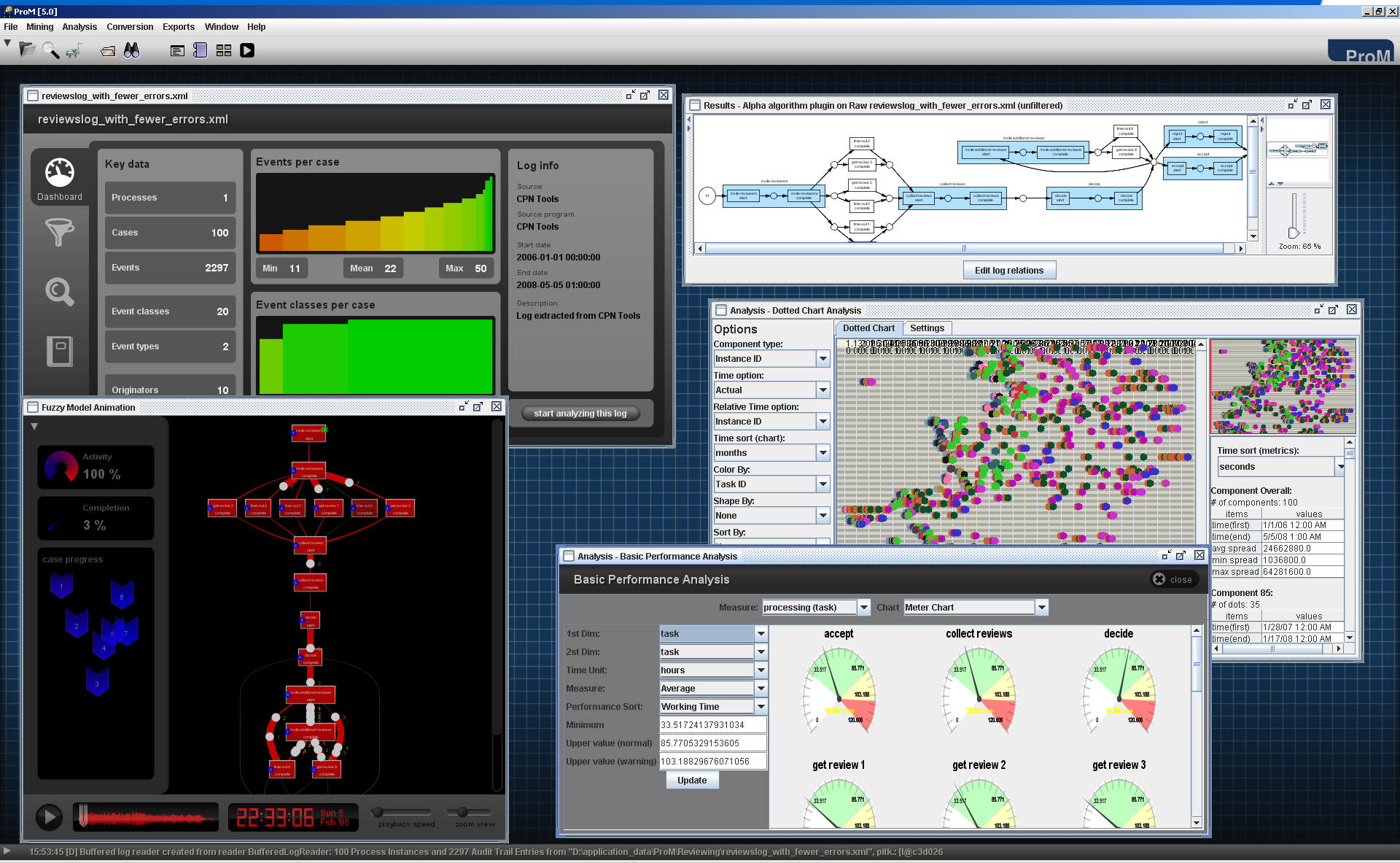
- **Compare process model and event log: highlight deviations and measure conformance.**
- **Compare constraints/business rules and event logs: check e.g. the 4-eyes principle.**



Tool support

- Open source initiative started in 2003 after several early prototypes.
- Common Public License (CPL).
- Current version: 5.0.
- ProMimport: to extract MXML from all kinds of applications
- Plug-in architecture.
- About 250 plug-ins available:
  - mining plug-ins: 38 (all mining algorithms presented and many more)
  - analysis plug-ins: 71 (e.g., verification, SNA, LTL, conformance checking, etc.)
  - import: 21 (for loading EPCs, Petri nets, YAWL, BPMN, etc.)
  - export: 44 (for storing EPCs, Petri nets, YAWL, BPMN, BPEL, etc.)
  - conversion: 45 (e.g., translating EPCs or BPMN into Petri nets)
  - filter: 24 (e.g., removing infrequent activities)

Screenshot of ProM 5.0



Business Intelligence Tools?

- Business Objects (SAP)
- Cognos Business Intelligence (IBM)
- Oracle Business Intelligence
- Hyperion (Oracle)
- SAS Business Intelligence
- Microsoft Business Intelligence
- SAP Business Intelligence (SAP BI)
- Jaspersoft (Open Source Business Intelligence)
- Pentaho BI Suite (Open Source)
- ....

- Dashboards, reports, scorecards, ...
- Slicing and dicing, data mining, ...

Process Mining Software

Futura Reflect

BPM|one

Comprehend

ARIS Process Performance Manager

Interstage Automated Business Process Discovery & Visualization

Process Discovery Focus

Enterprise Visualization Suite

Process Mining: Applications

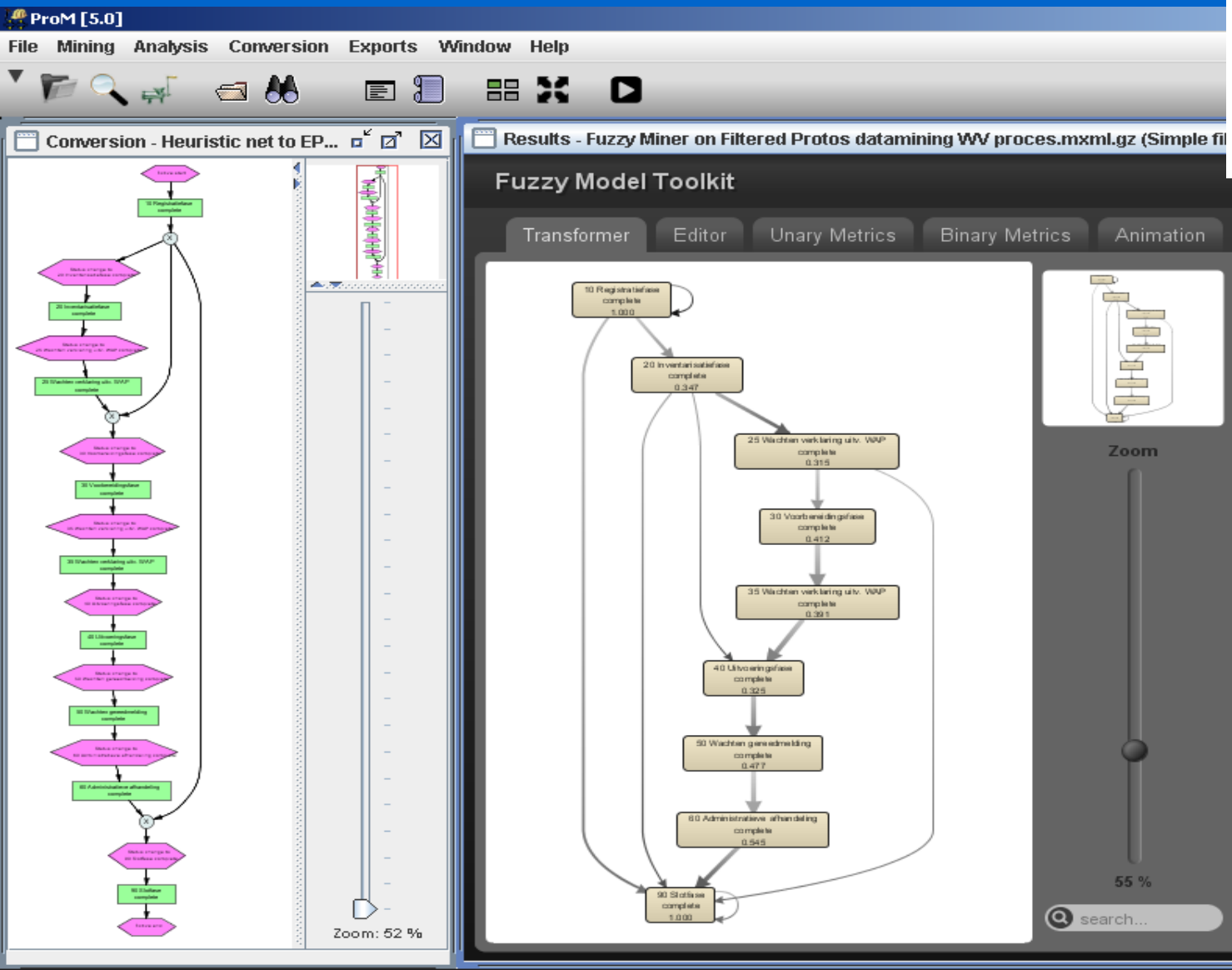
TU/e Technische Universiteit Eindhoven University of Technology

Where innovation starts

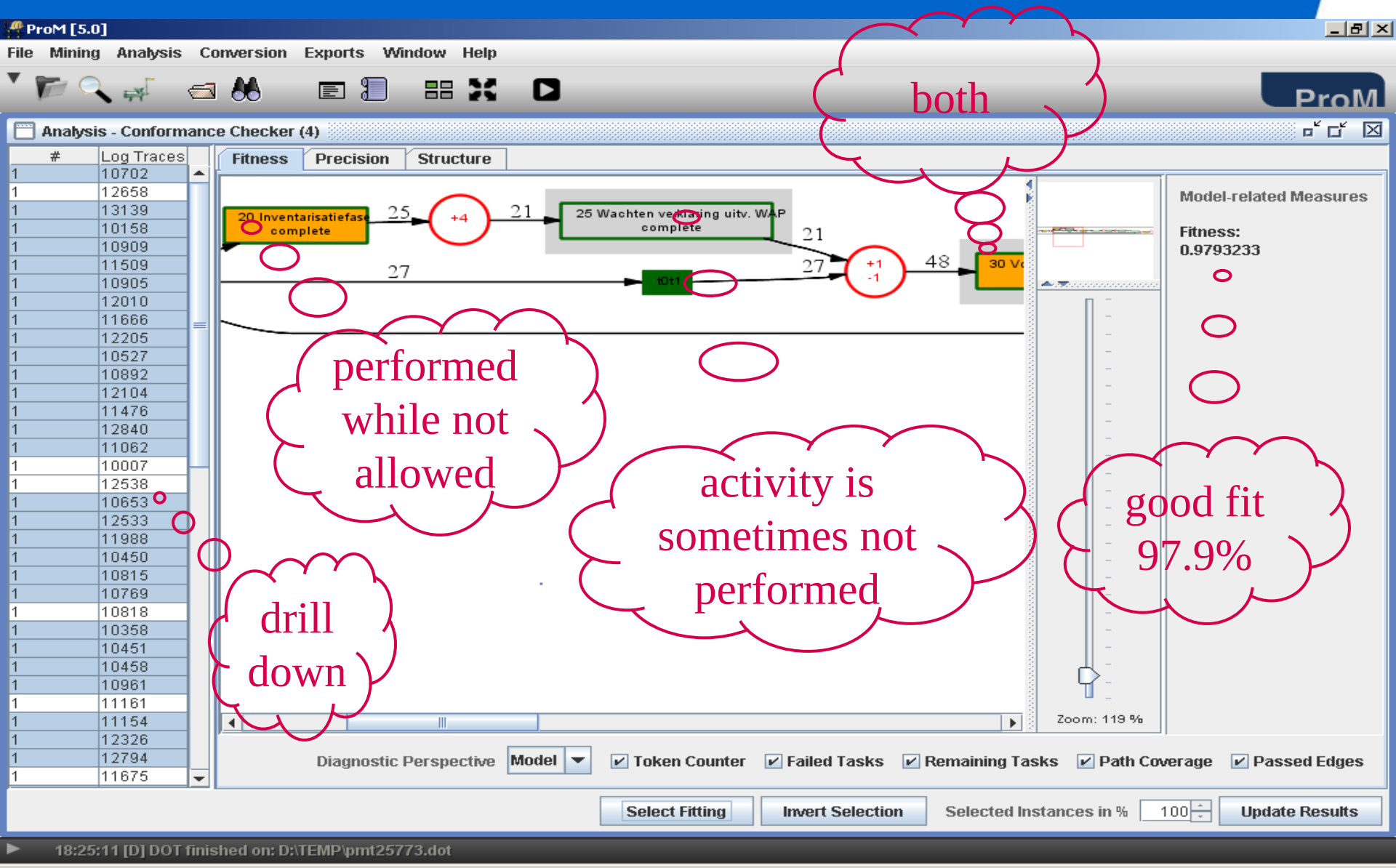
Where did we apply process mining?

- Municipalities (e.g., Alkmaar, Heusden, Harderwijk, etc.)
- Government agencies (e.g., Rijkswaterstaat, Centraal Justitieel Incasso Bureau, Justice department)
- Insurance related agencies (e.g., UWV)
- Banks (e.g., ING Bank)
- Hospitals (e.g., AMC hospital, Catharina hospital)
- Multinationals (e.g., DSM, Deloitte)
- High-tech system manufacturers and their customers (e.g., Philips Healthcare, ASML, Thales)
- Media companies (e.g. Winkwaves)
- ...

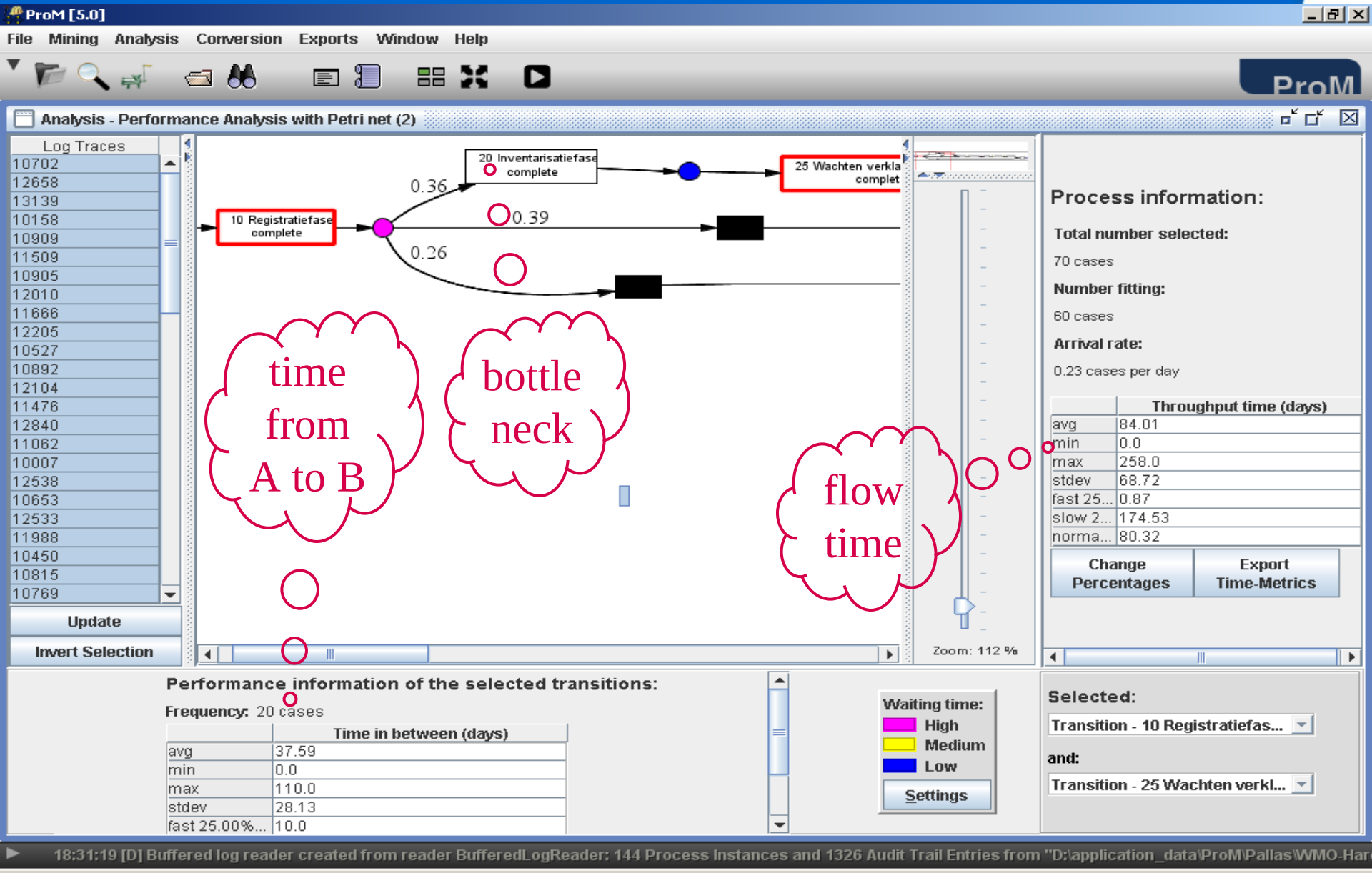
Example: A Dutch Municipality



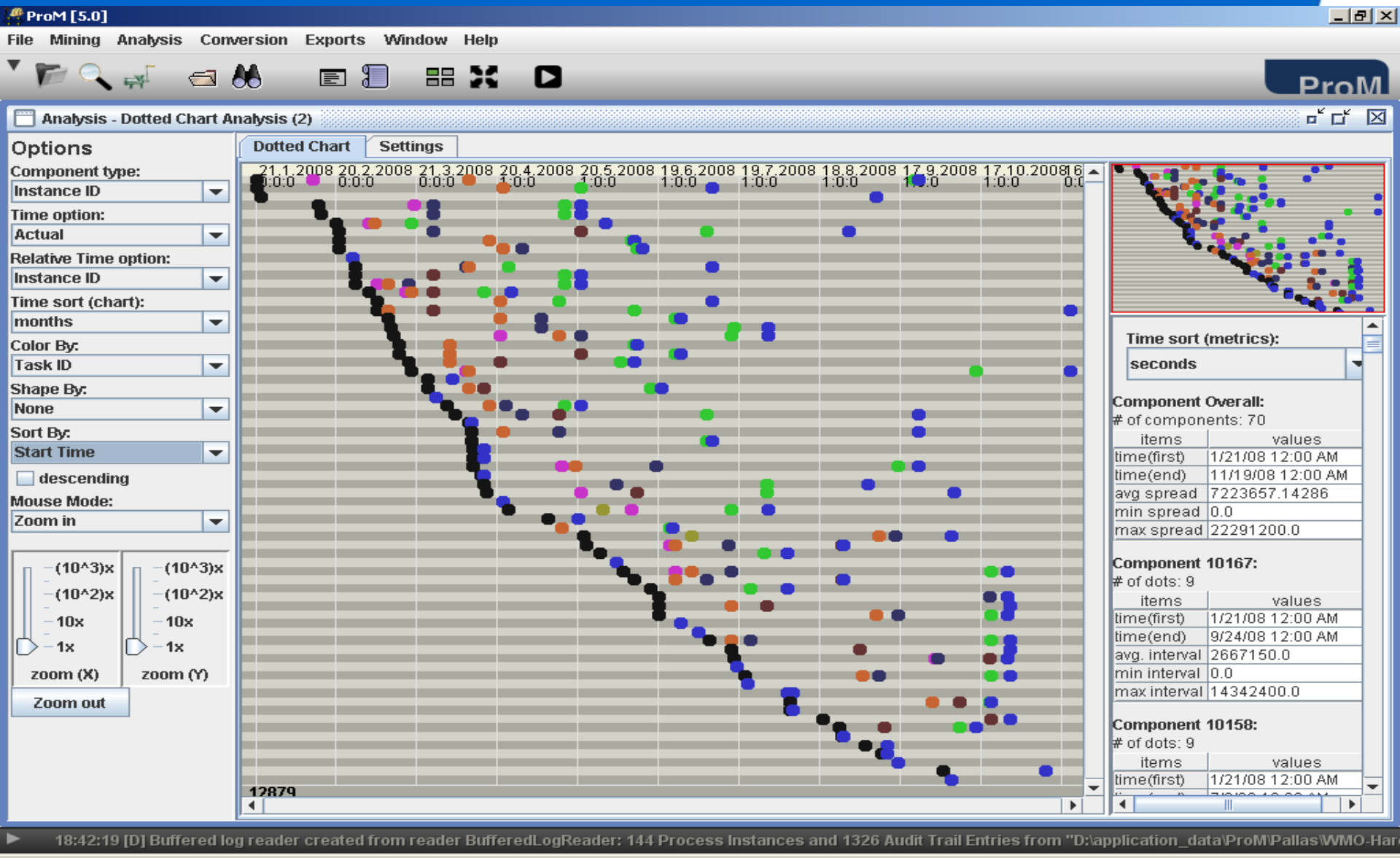
Conformance check of discovered model



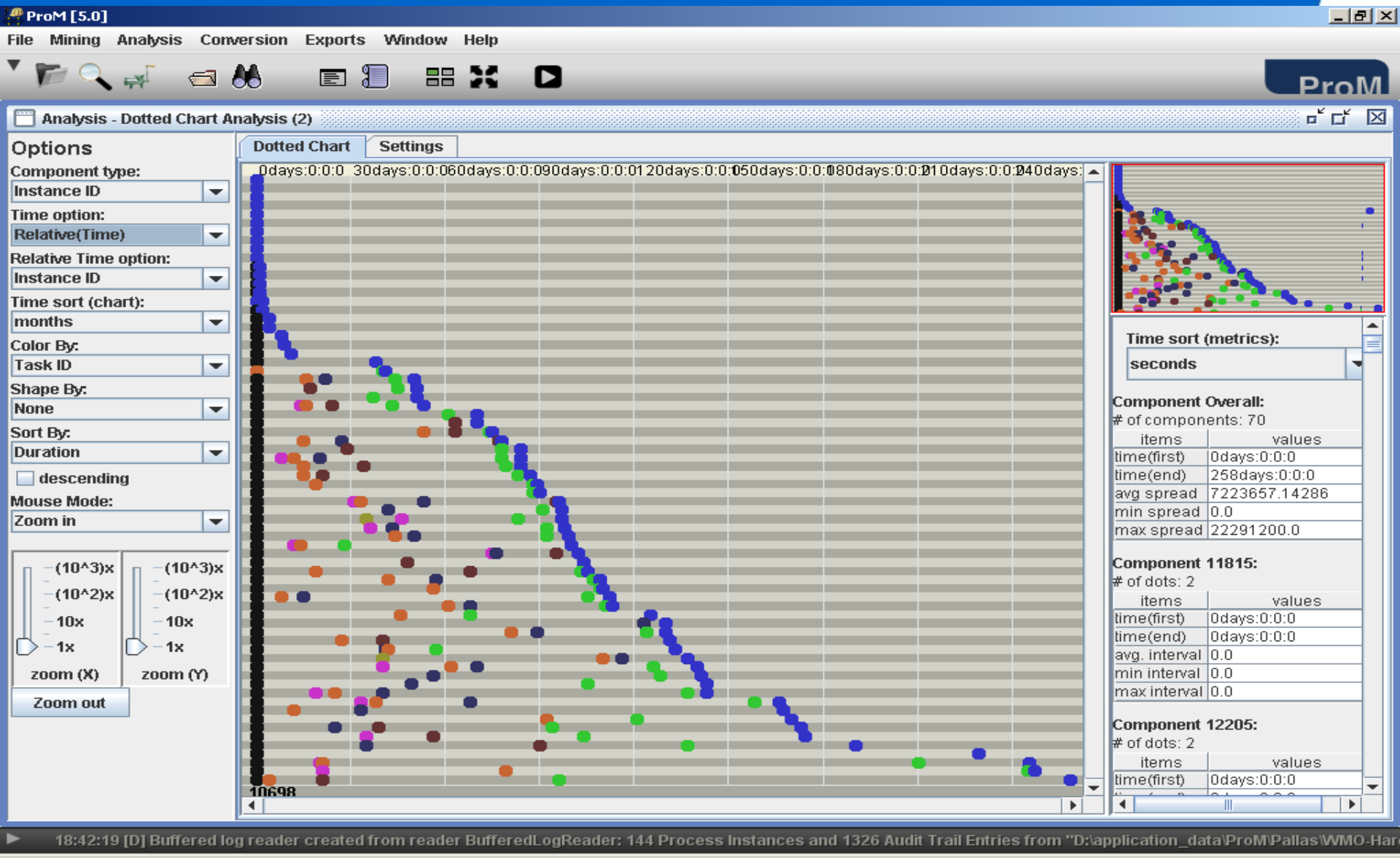
Performance analysis



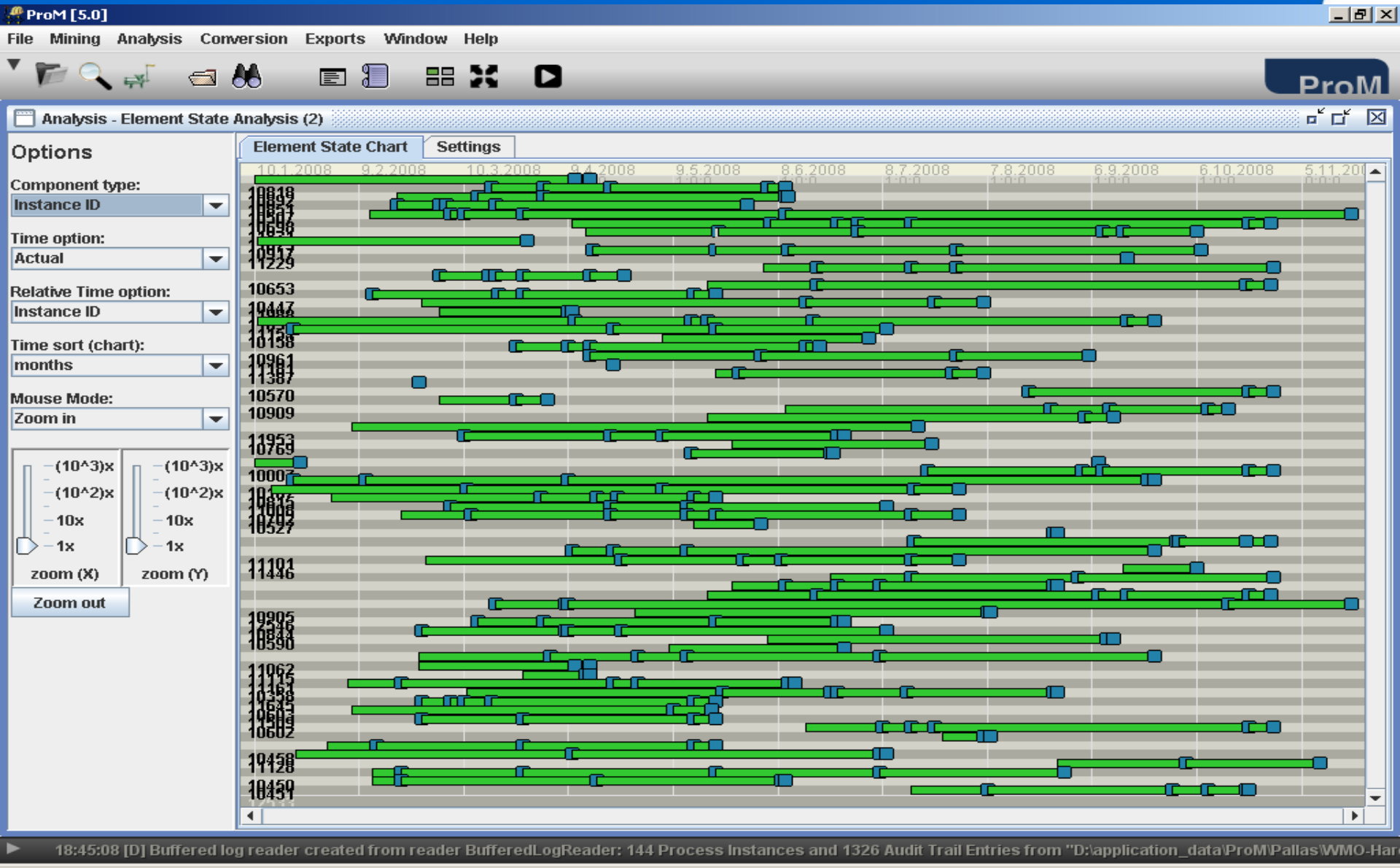
Events sorted by start time of case



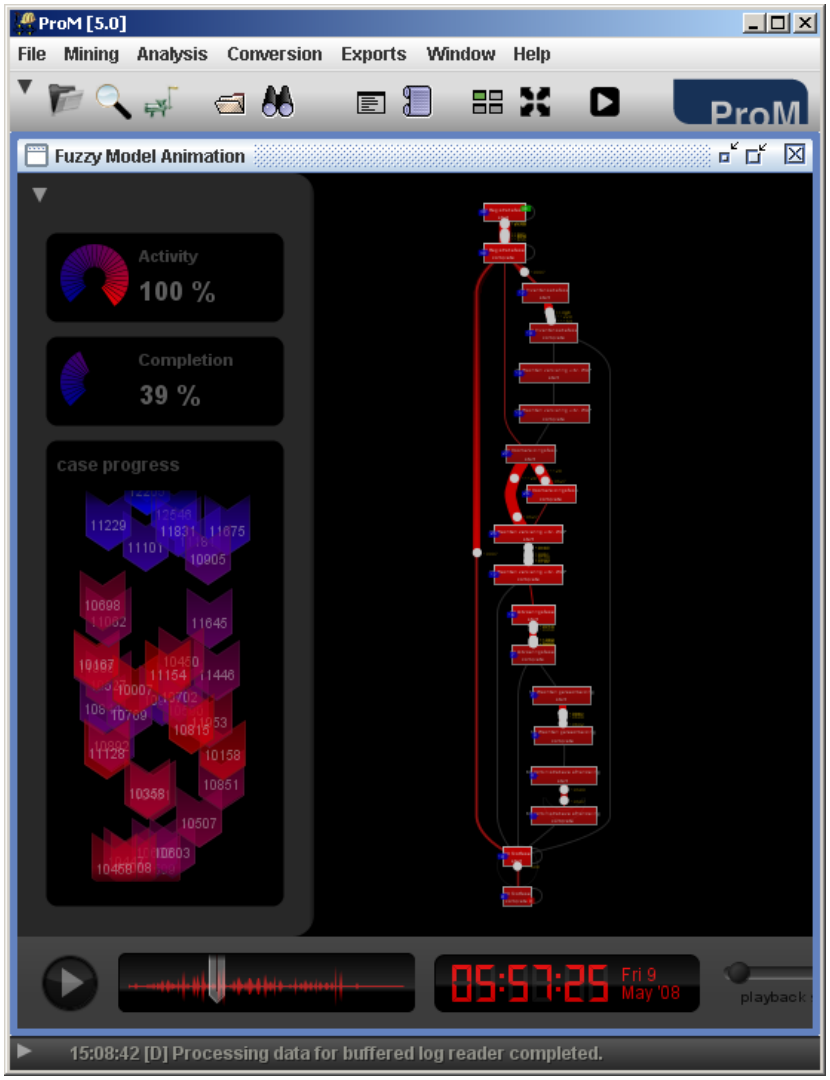
Events sorted by duration

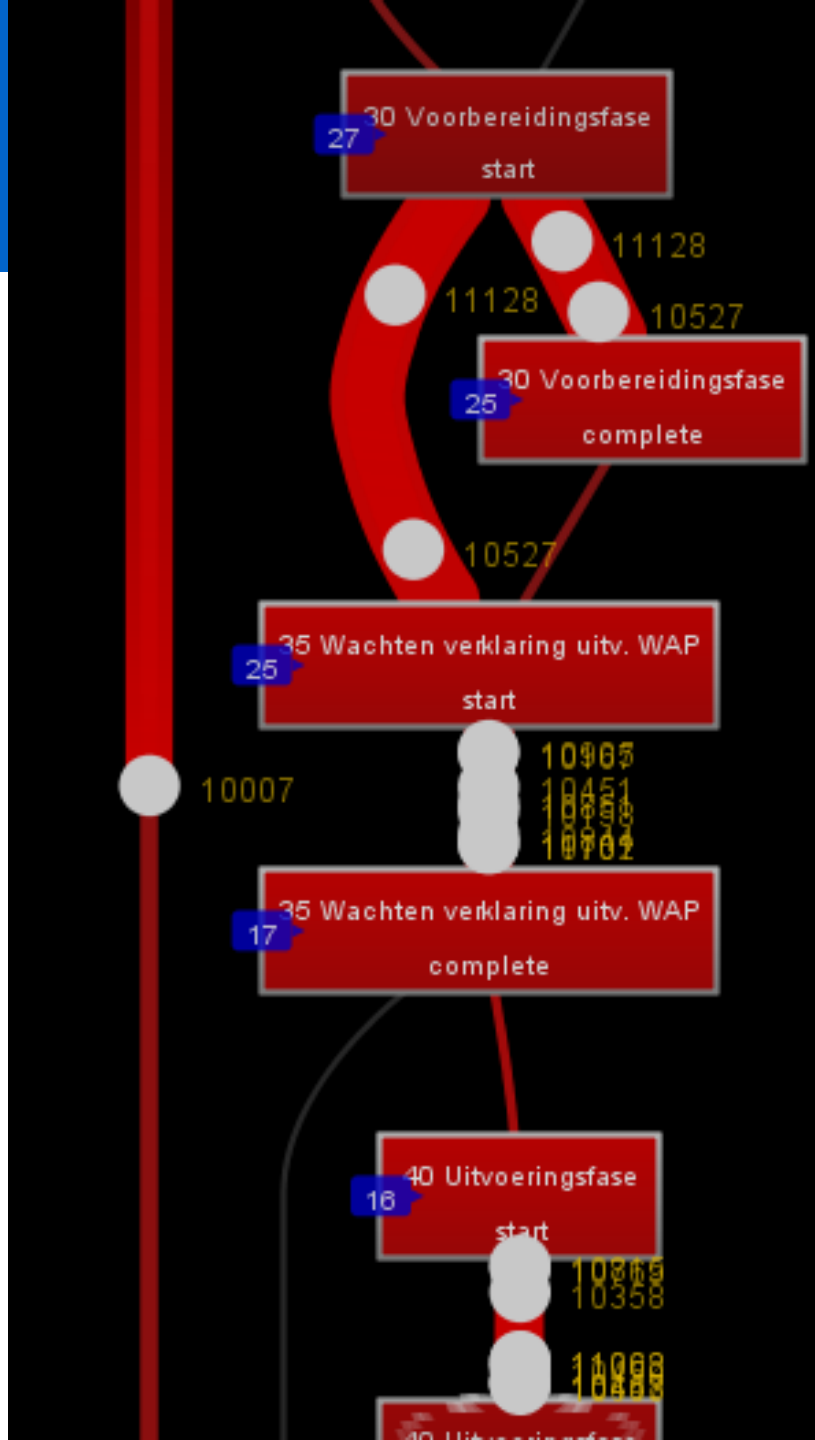


Idle time versus working time

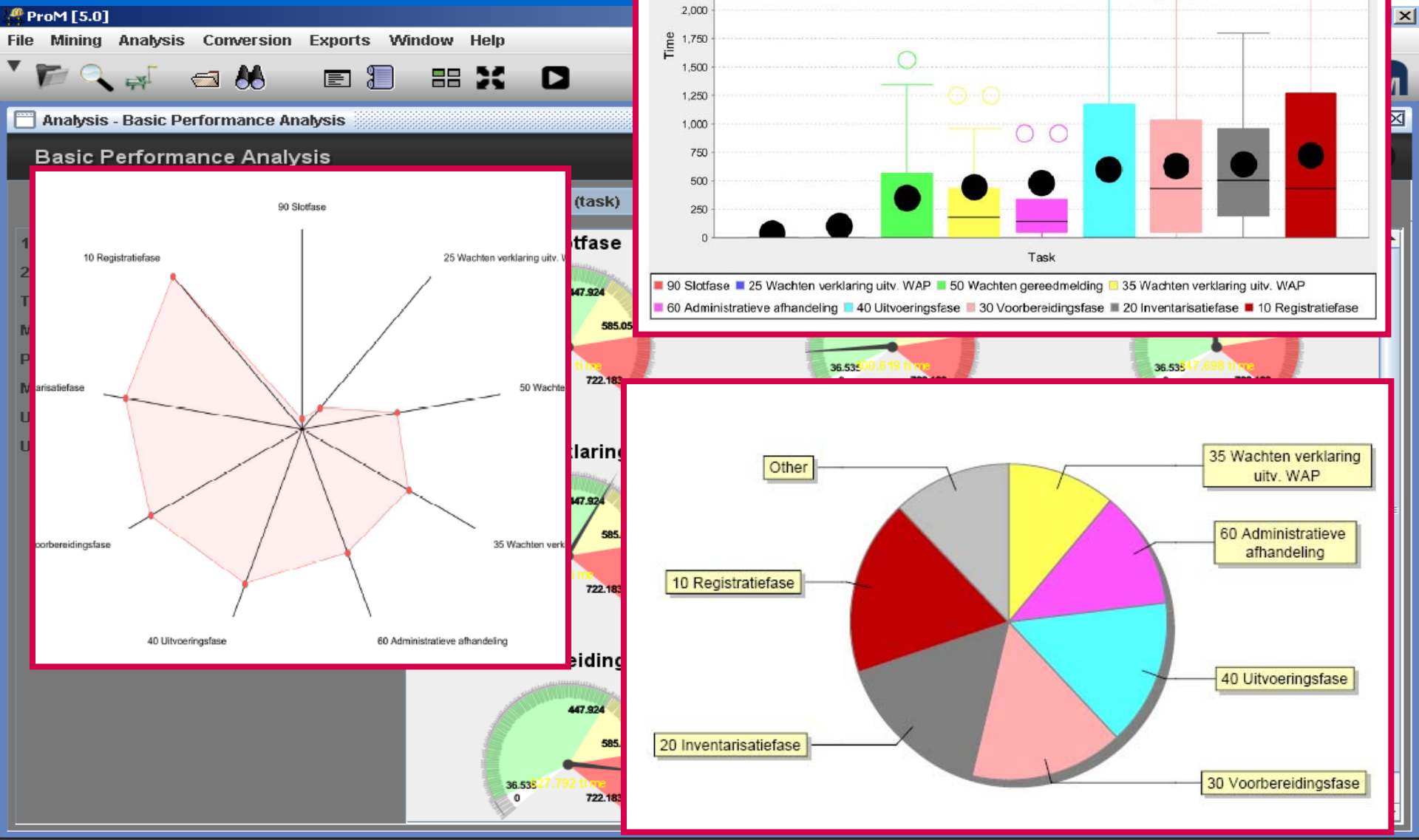


"Real" animation



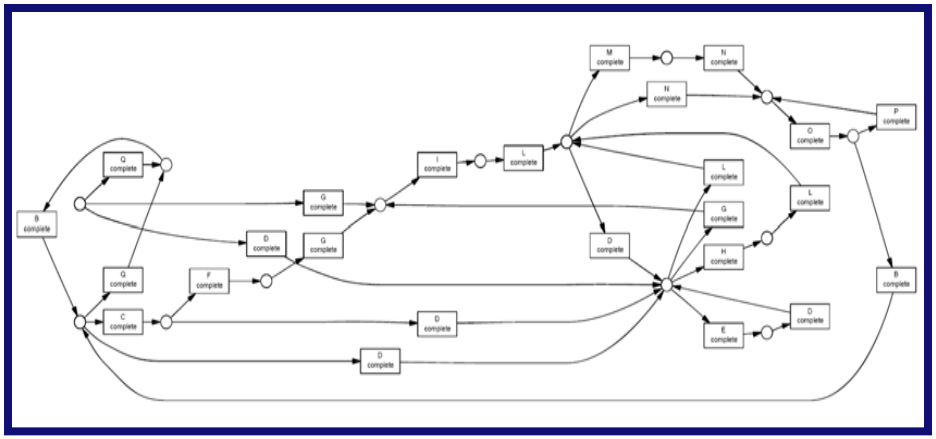


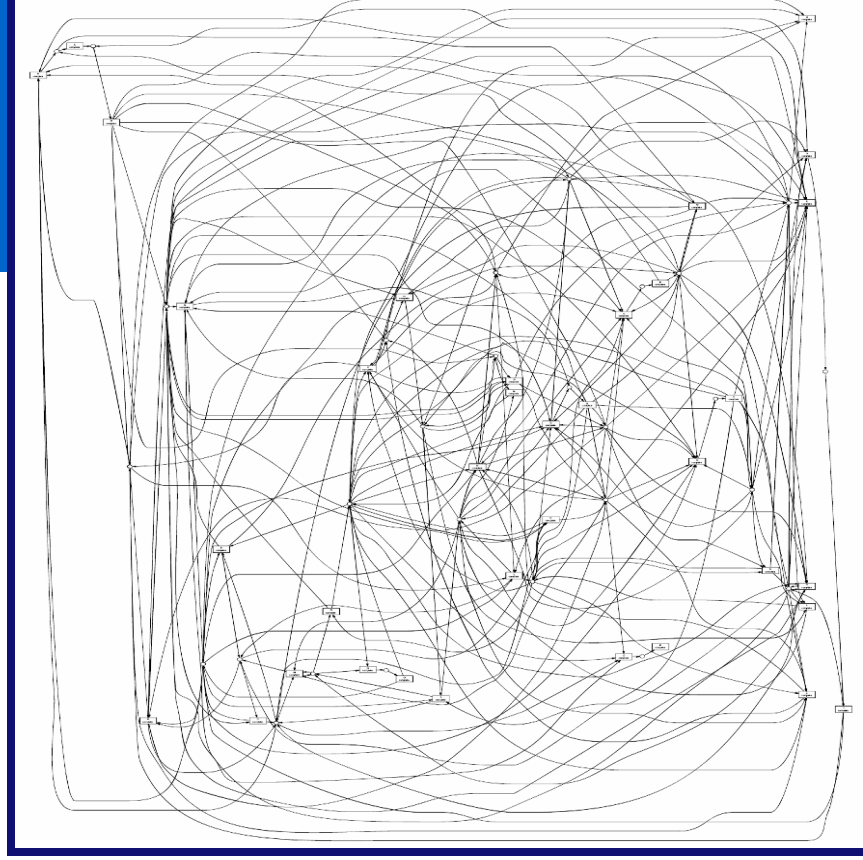
And of course ...



Reality ≠ PowerPoint (or Visio)

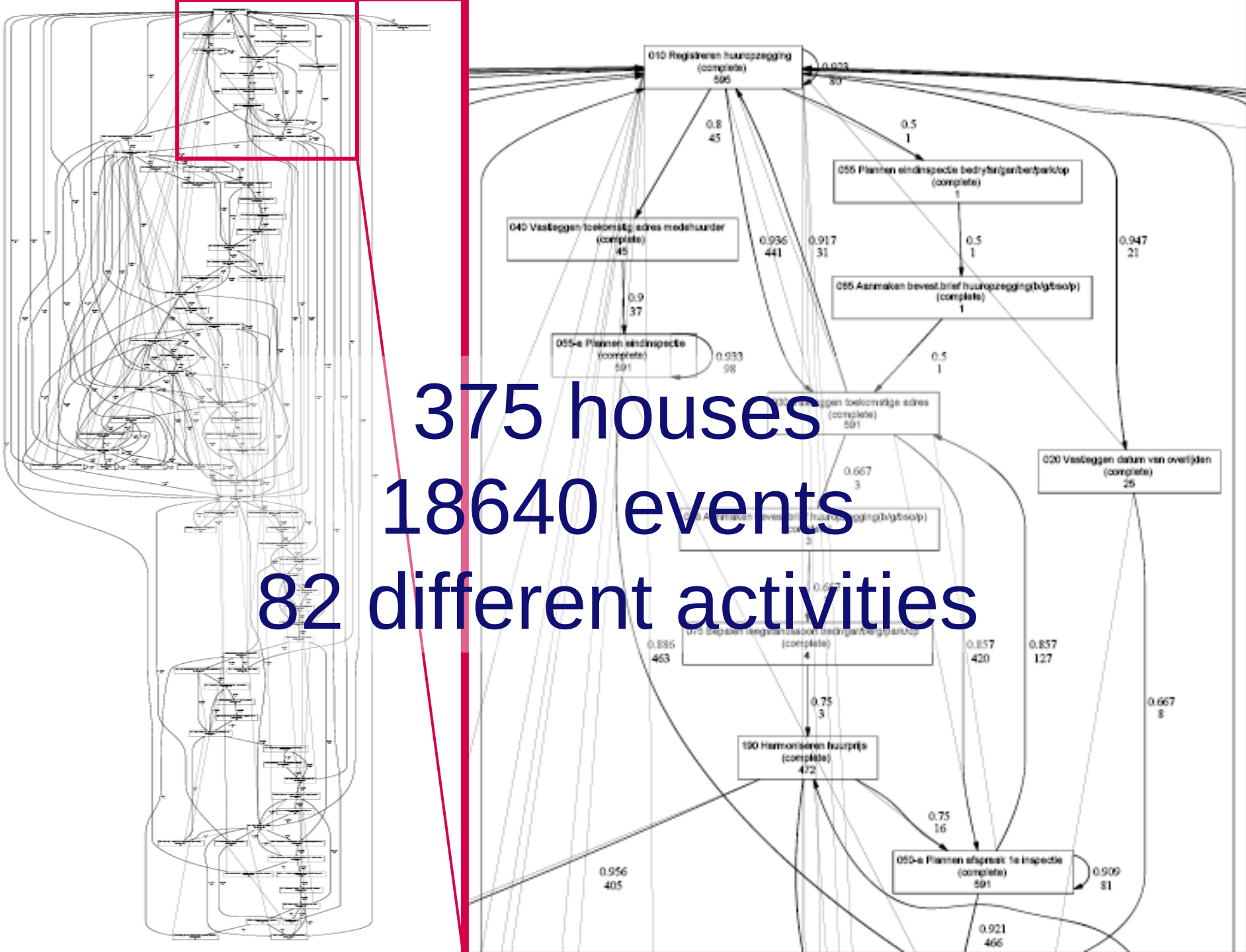
Process spectrum



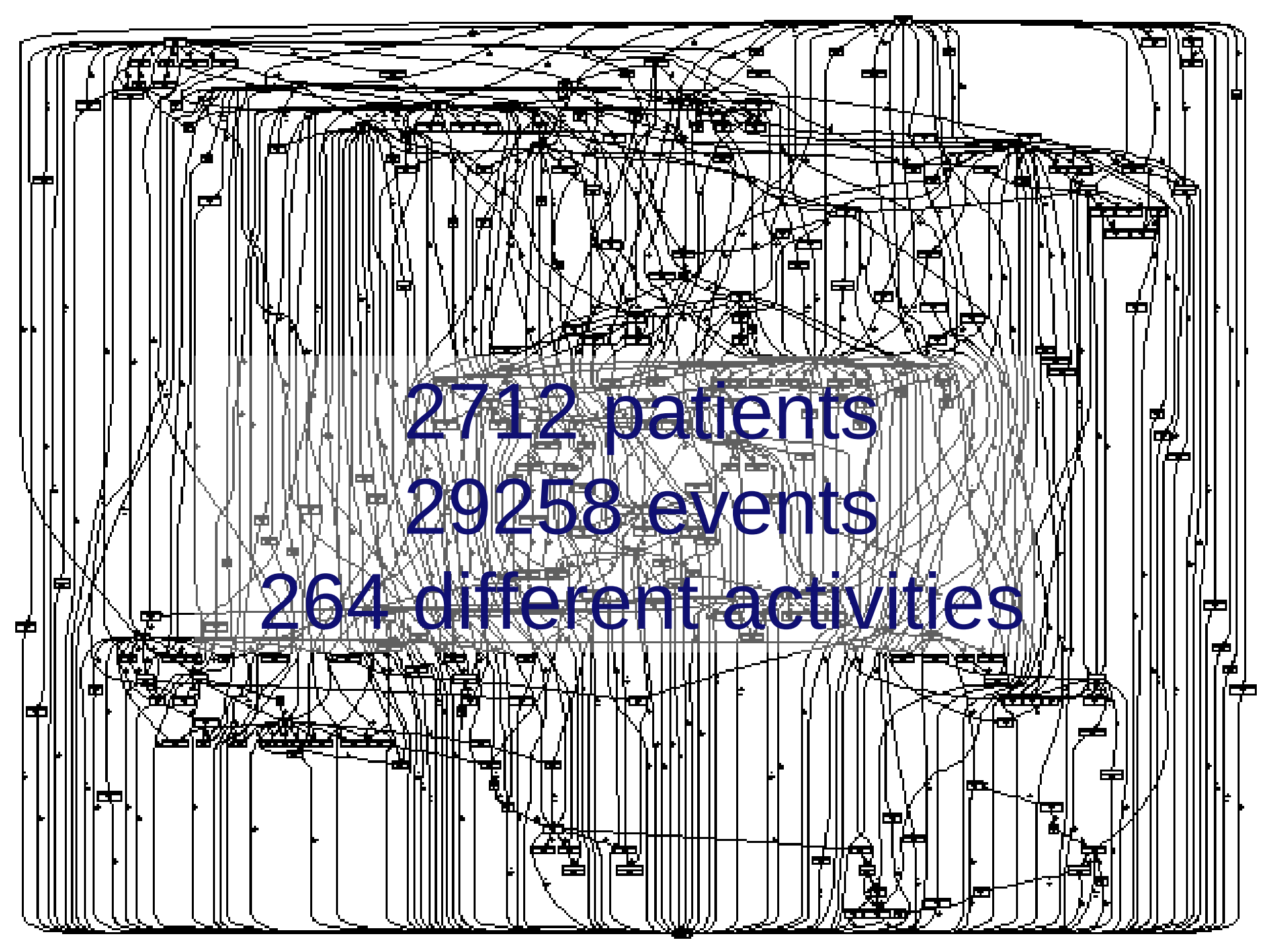


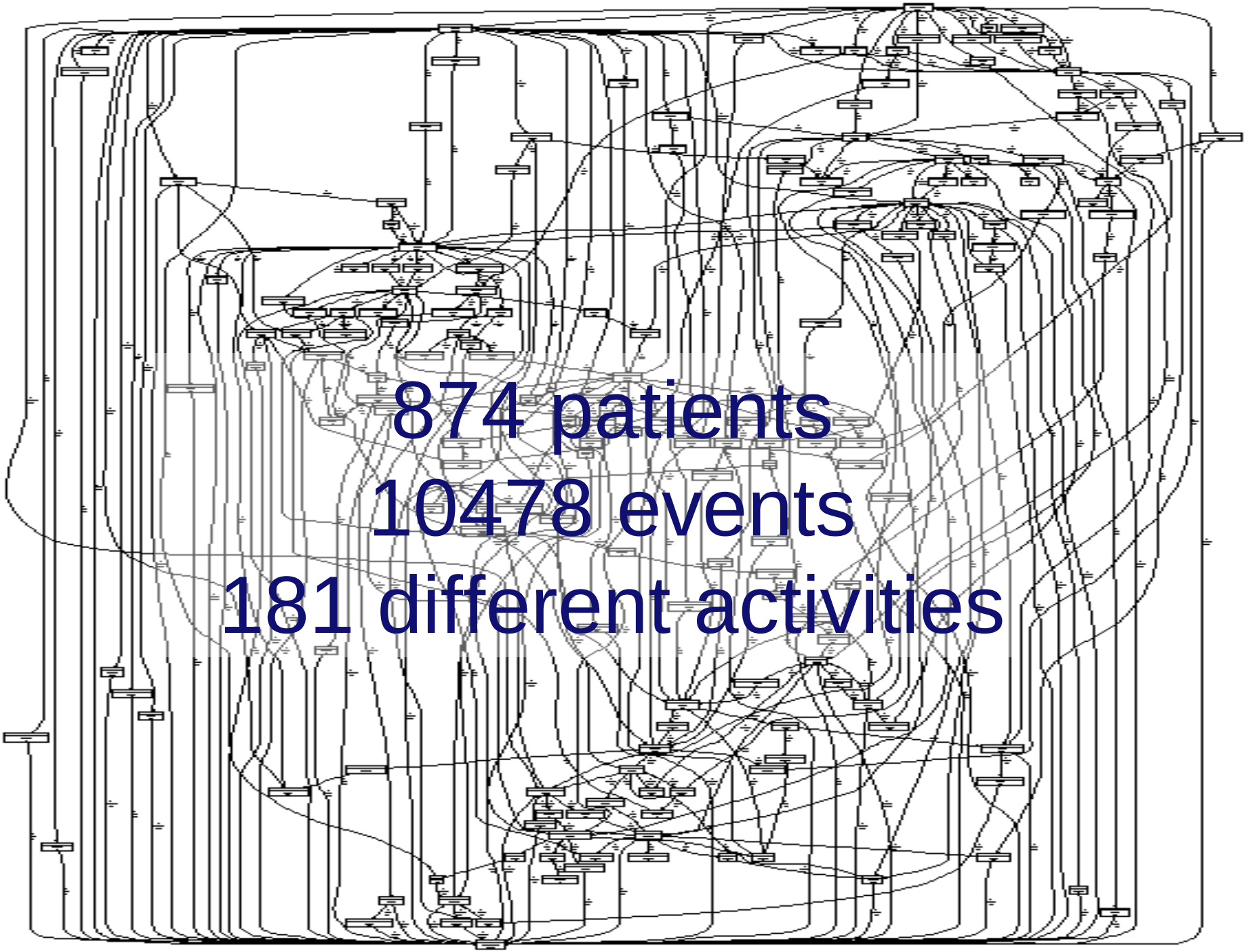
structured (Lasagna)

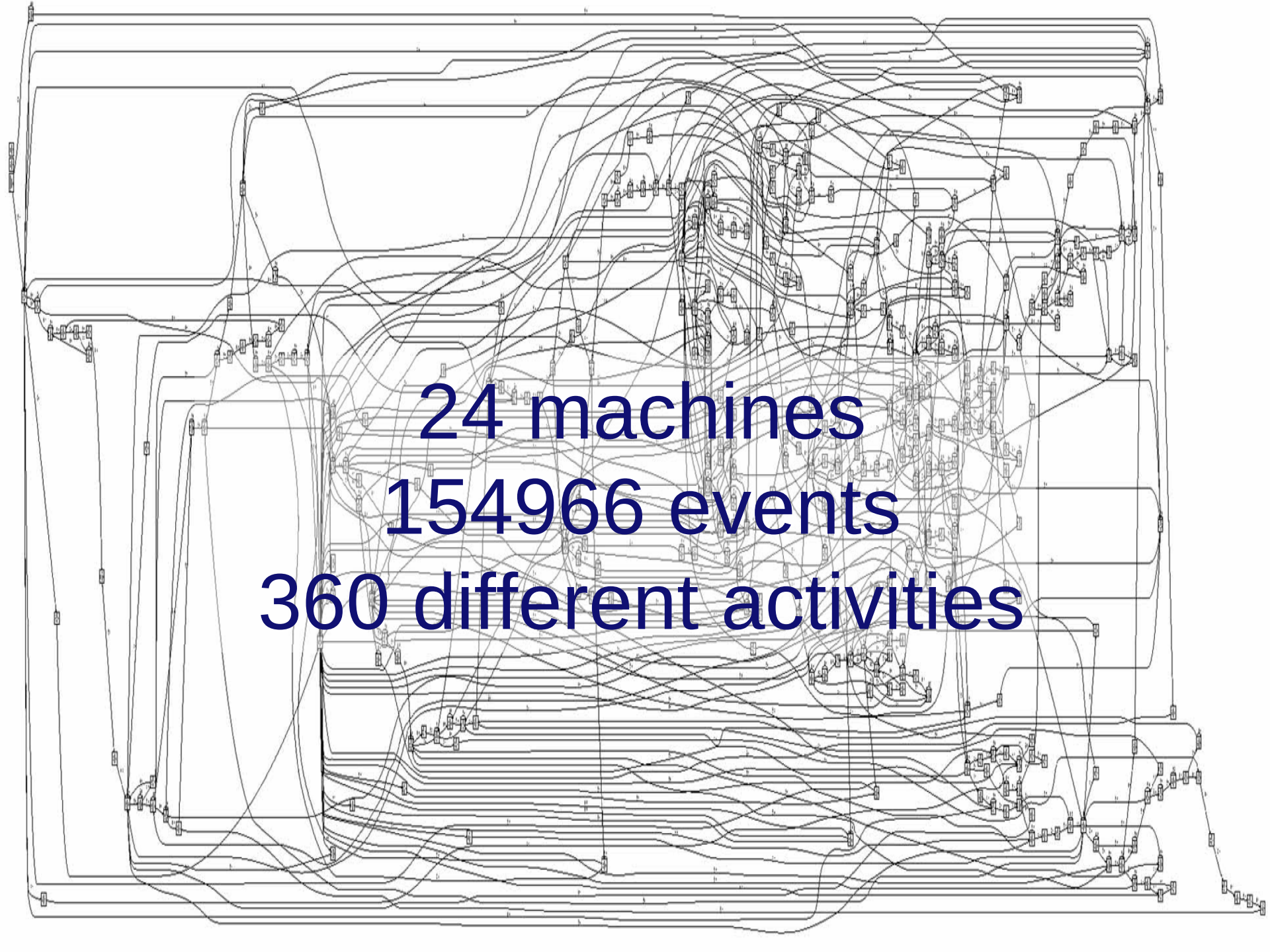
unstructured (Spaghetti)

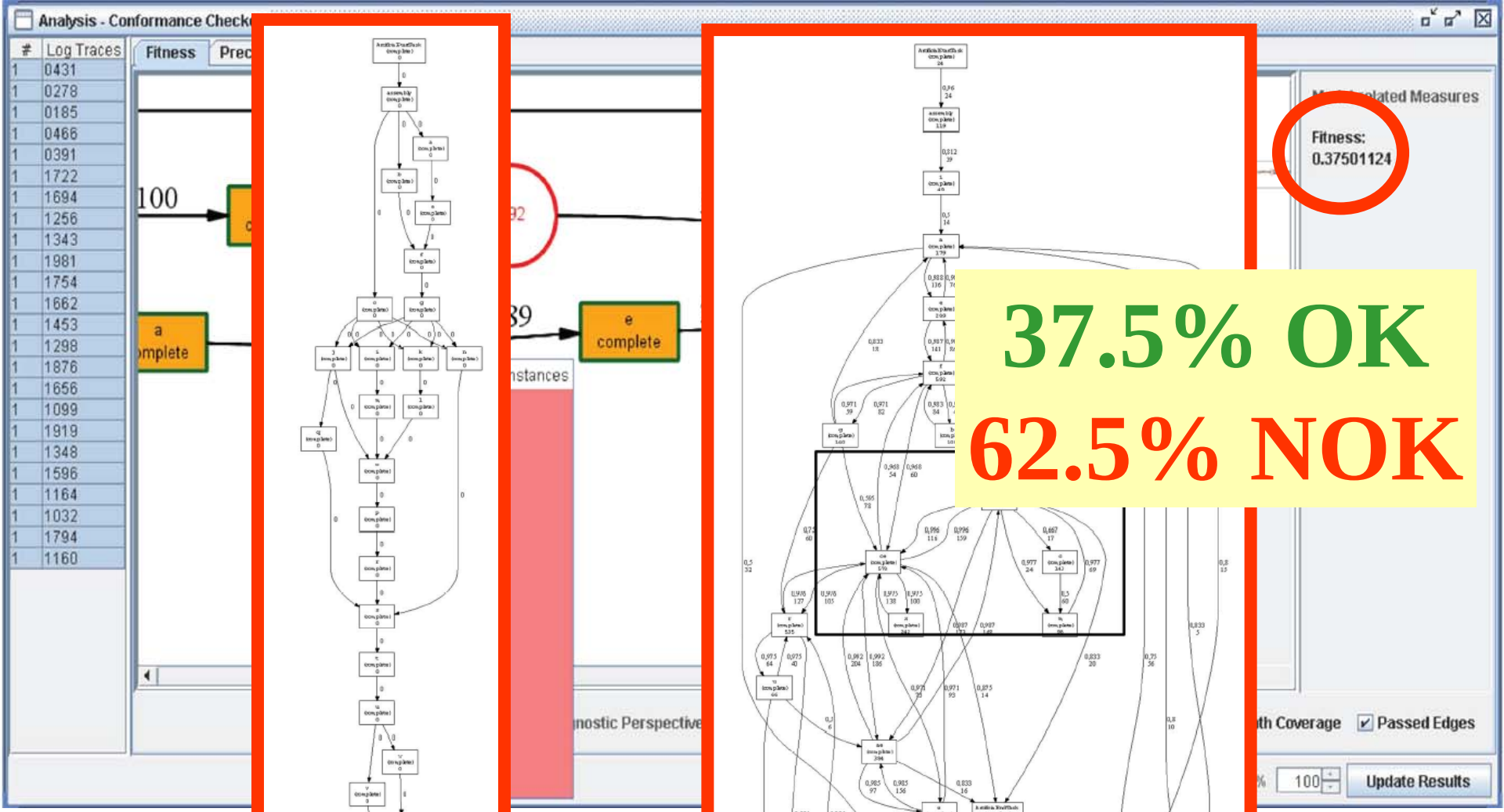


375 houses
18640 events
82 different activities


2712 patients
29258 events
264 different activities


874 patients
10478 events
181 different activities


24 machines
154966 events
360 different activities



37.5% OK

62.5% NOK

design

reality

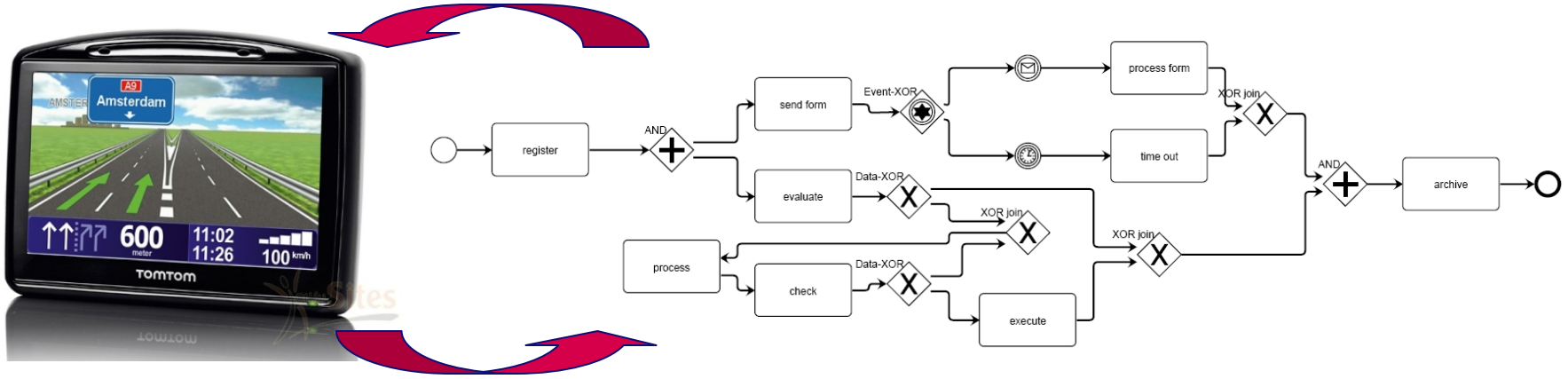
Process Mining: TomTom for Business Processes

TOMTOM

TU/e Technische Universiteit Eindhoven University of Technology

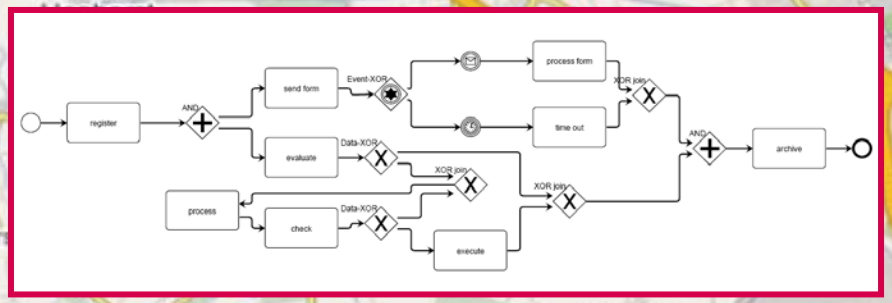
Where innovation starts

Business Process Navigation?



- Often a good process map is missing (incorrect, outdated, no color, ...)
- Process maps inherit the limitations of paper maps (no zoom or views)
- Process maps tend to aim at "controlling the driver"
- Current location unknown
- No traffic information is given
- No recalculation of the route
- No estimated arrival time
- ...

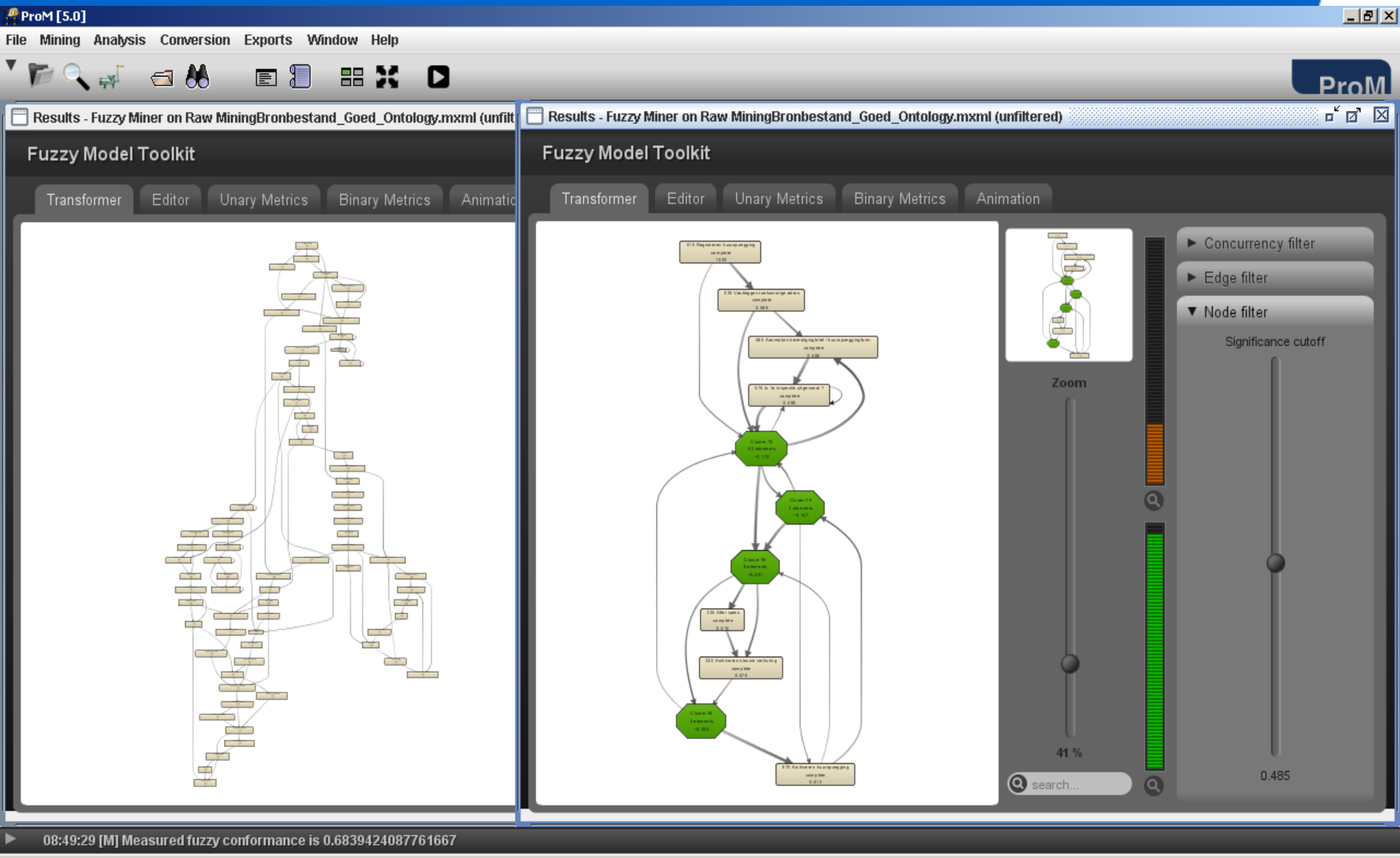
What we can learn from maps ...



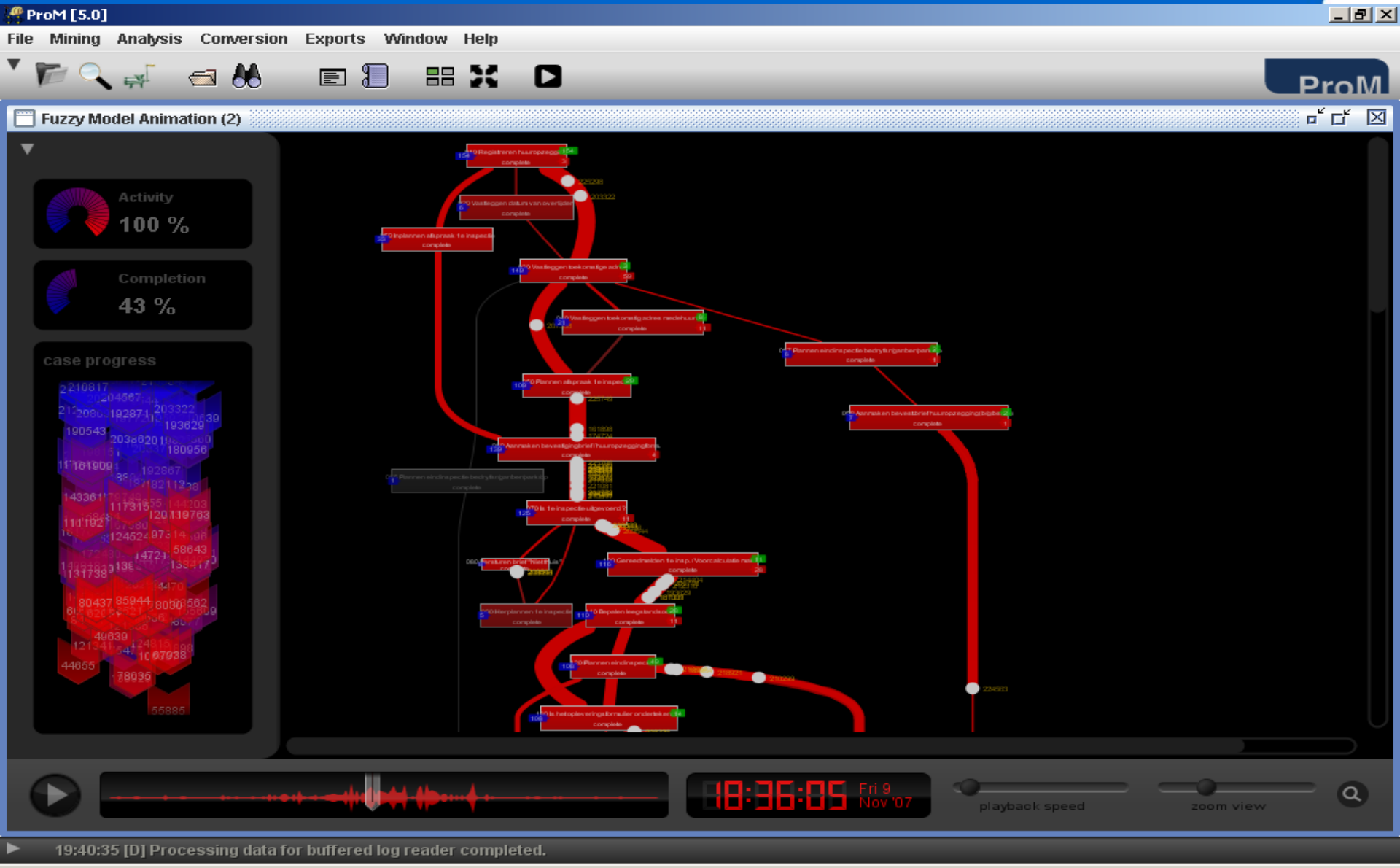
Why imitate paper maps?

- Zoom in - zoom out
- Various views (e.g. show hotels and fuel stations at will)
- Dynamic content!
- Traffic information
- Show current location

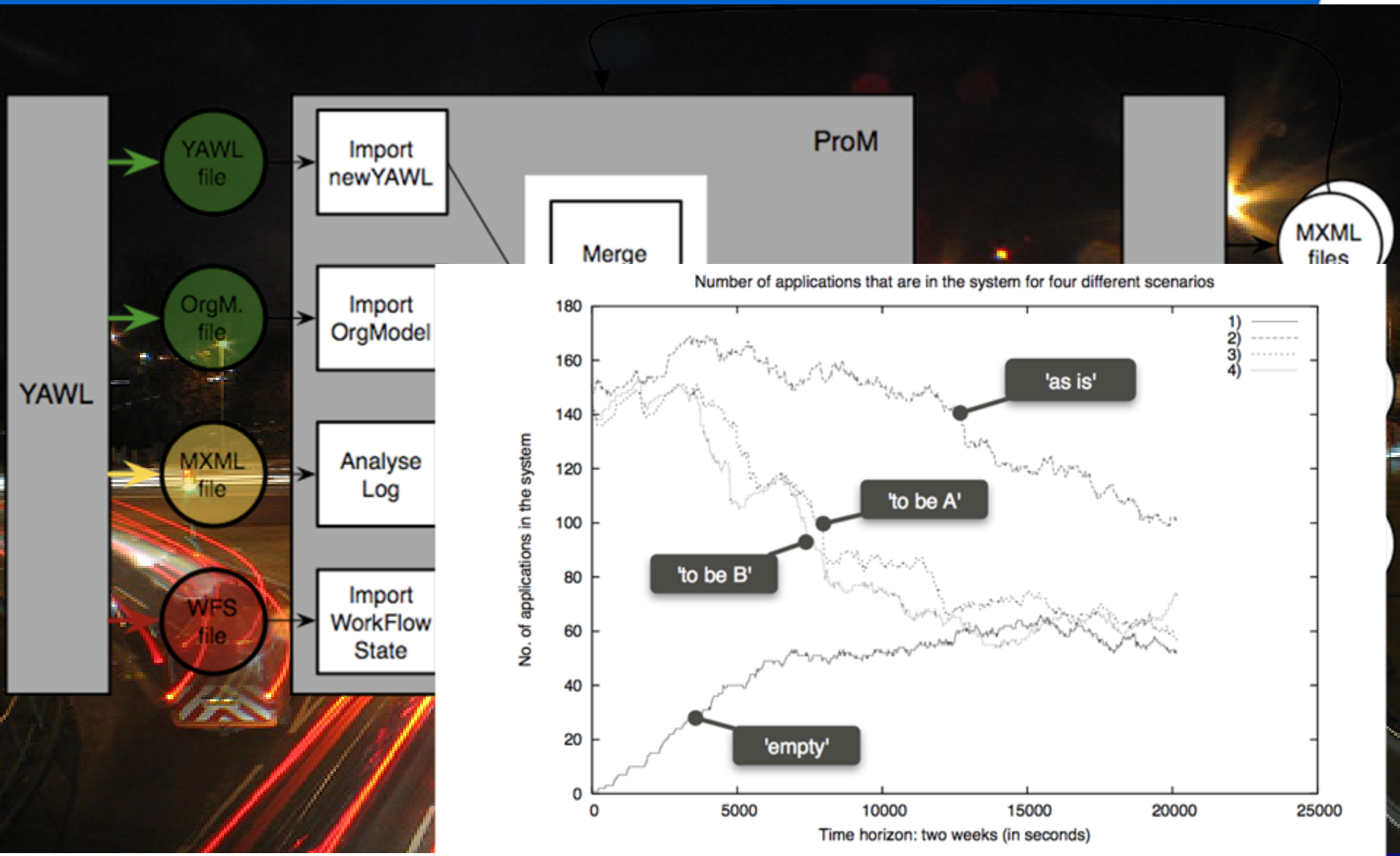
ProM's Fuzzy Miner: Seamless zoom



ProM's "real animation"



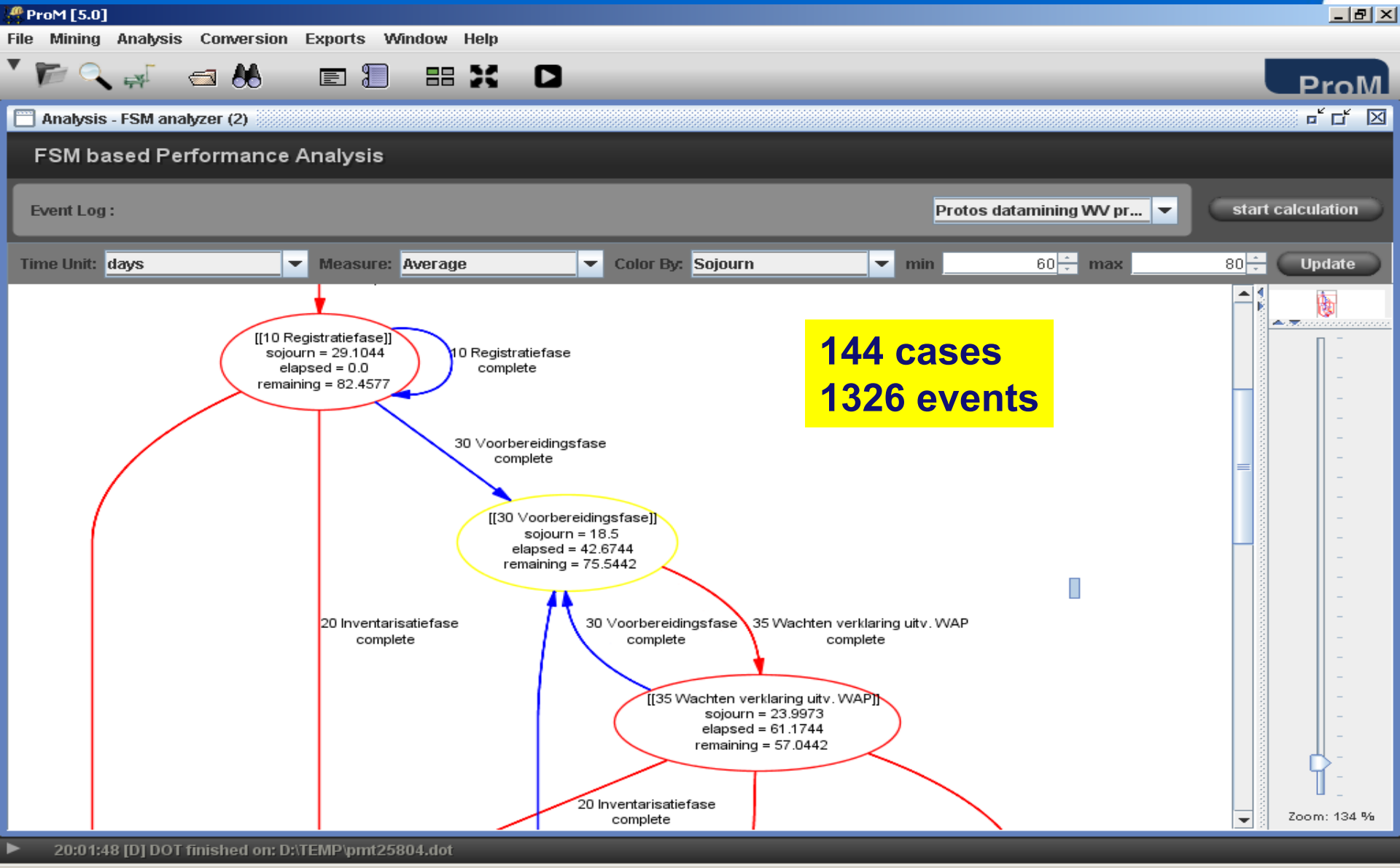
ProM's "real simulation"



Prediction and recommendation

- **Prediction: When are we home?**
- **Recommendation: What should I do next?**
- **Suggestions without force and the willingness to continuously recalculate the route.**

ProM's Case prediction capabilities



Conclusion

Conclusion

- The abundance of event data enables a wide variety of process mining techniques ranging from process discovery to conformance checking.
- This is already possible today!
- Check out ProM with its 250+ plug-ins.
- A reality check for people that are involved in process modeling.
- Demand TomTom functionality!

Thanks! cf. *www.processmining.org*

- Wil van der Aalst
- Peter van den Brand
- Boudewijn van Dongen
- Christian Günther
- Eric Verbeek
- Ana Karla Alves de Medeiros
- Anne Rozinat
- Minseok Song
- Ton Weijters
- Remco Dijkman
- Gianluigi Greco
- Antonella Guzzo
- Kristian Bisgaard Lassen
- Ronny Mans
- Jan Mendling
- Vladimir Rubin
- Kenny van Uden
- Irene Vanderfeesten
- Barbara Weber
- Lijie Wen
- Mercy Amiyo
- Carmen Bratosin
- Toon Calders
- Jorge Cardoso
- Ronald Crooy
- Florian Gottschalk
- Monique Jansen-Vullers
- Peter Khisa Wakholi
- Nicolas Knaak
- Sven Lambrechts
- Joyce Nakatumba
- Mariska Netjes
- Mykola Pechenizkiy
- Maja Pesic
- Hajo Reijers
- Stefanie Rinderle
- Domenico Saccà
- Helen Schonenberg
- Marc Voorhoeve
- Jianmin Wang
- Jan Martijn van der Werf
- Martin van Wingerden
- Jianhong Ye
- Huub de Beer
- Elena Casares
- Alina Chipaila
- Walid Gaaloul
- Martijn van Giessel
- Shaifali Gupta
- Thomas Hoffmann
- Peter Hornix
- René Kerstjens
- Ralf Kramer
- Wouter Kunst
- Laura Maruster
- Andriy Nikolov
- Adarsh Ramesh
- Jo Theunissen
- ...

Relevant WWW sites

- http://www.processmining.org
- **http:// promimport.sourceforge.net**
- **http://prom.sourceforge.net**
- **http://www.workflowpatterns.com**
- **http://www.workflowcourse.com**
- **http://www.vdaalst.com**