Process Mining in Healthcare



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

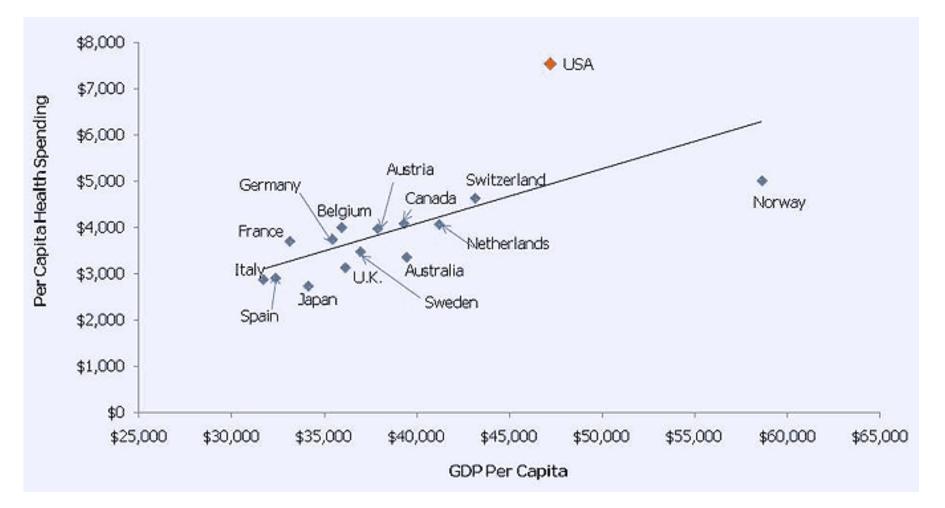
TUe Technische Universiteit Eindhoven University of Technology

Where innovation starts



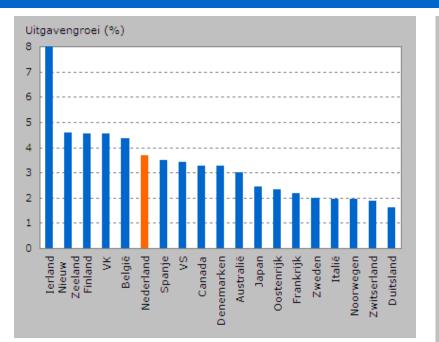
Healthcare Costs

Total Health Expenditure per Capita and GDP per Capita 2008

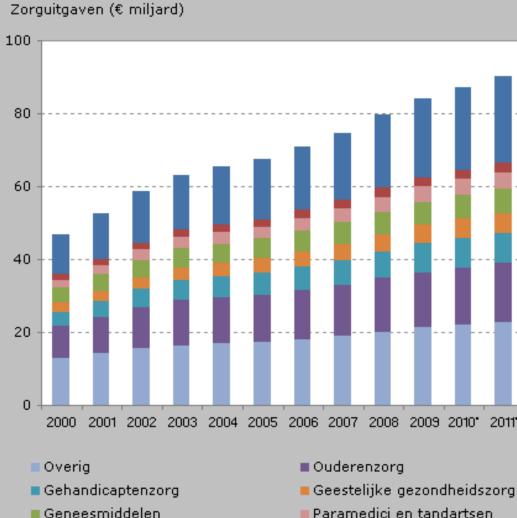


Source: Organisation for Economic Co-operation and Development (2010), "OECD Health Data", *OECD Health Statistics* (database). doi: 10.1787/data-00350-en

Costs are Rising!!



Gemiddelde jaarlijkse reële uitgavengroei per inwoner (%), 2000-2008 (Bron: OECD Health Data, 2010)a

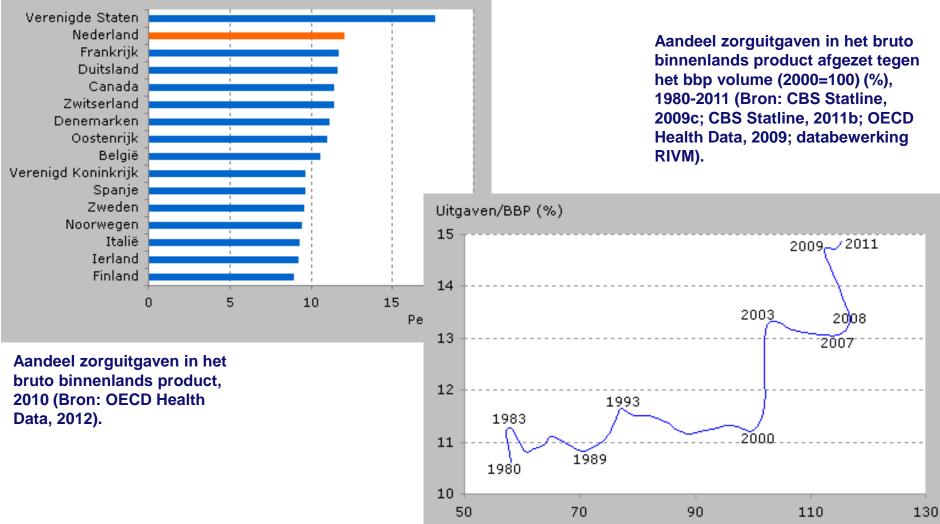


Zorguitgaven per sector, € miljard, 2000-2011 (Bron: CBS Statline, 2011ba; CBS, 2012).

Huisartsen

Ziekenhuizen

Healthcare costs as a percentage of the Gross Domestic Product (GDP)

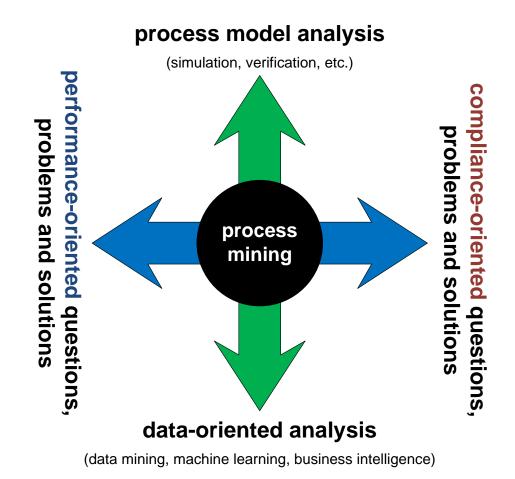


BBP volume index (2000=100)

Process Mining in Healthcare

- Show the actual care processes (rather than idealized or assumed ones) and provide insights
- Uncover bottlenecks in care processes
- Show deviations from the "happy path"
- Performance and compliance analyzed in a unified, evidence-based, manner
- Provide operational support: predicting problems and recommending actions
- From "politics" to "analytics" to avoid "management by PowerPoint" and the usual "IT disasters"
- Improve service and reduce costs in healthcare

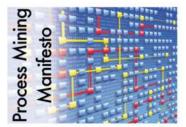
Positioning Process Mining



6

Advances in Process Mining

- Many process discovery and conformance checking algorithms and tools are available (cf. the various **ProM** packages).
- Also commercial software based on these ideas: Disco (Fluxicon), Reflect (Futura/Perceptive), BPMOne (Pallas Athena/Perceptive), ARIS Process Performance Manager (Software AG), Interstage Automated Process Discovery (Fujitsu), QPR ProcessAnalyzer/Analysis (QPR Software), flow (fourspark), Discovery Analyst (StereoLOGIC), etc.
- We applied process mining in over 100 organizations.

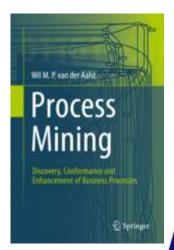


A manifesto is a "public	Protect and
declaration of principles	terrorch die
and intentions" by a group	mining in B
of people. This incridesto is	methoday at himself. The lat
written by members and	discost. m
supporters of the IEEE Task	promiser [
Force on Process Mining.	Ingie results
The good of this task force	(information)
is to promote the research,	Substantian (see
development, education.	mant ligh a
implementation, evolution,	and and
fo and understanding of	second and a
process mining.	Internation

he use held, and process	model actention, model repair, prediction, and fasterphoned micromonodifions.	
Ind analysis on the other service and improves well (in , we'l assessment processed or , we'l assessment processed or produktion in training assesses training assessed assesses module literature deviations for comparison deviations for comparison deviati	Contents: Forein Hong - Source Hon for Source - Contenge University Barrier	3 6 10 13 14

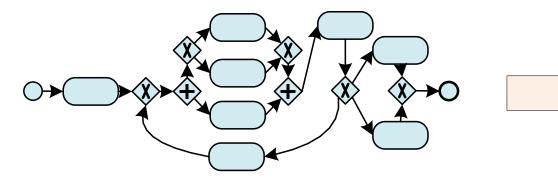
More than 75 people involving more than 50 organizations created the Process Mining Manifesto in the context of the IEEE Task Force on Process Mining.

Available in 13 languages



On the different roles of (process) models ...



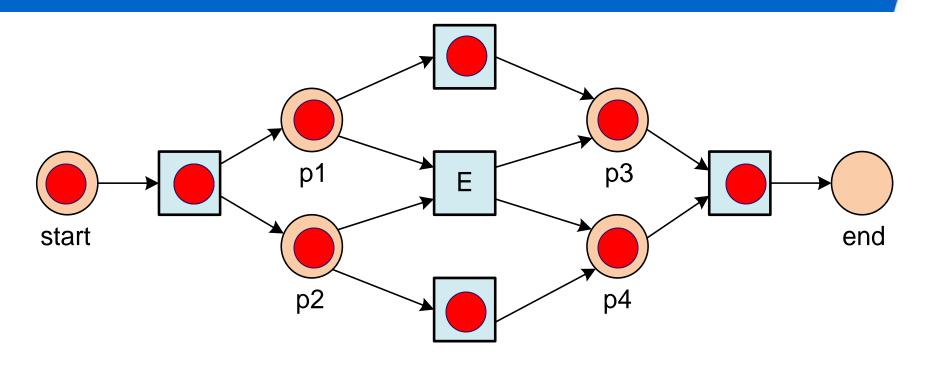




process model

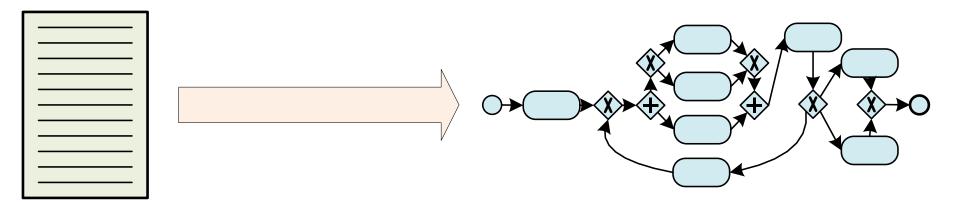
event log

Play-Out (Classical use of models)



A B C D A E DA E DA B C DA B C DA C B DA C B DA E DA C B DA E DA C B DA C B D



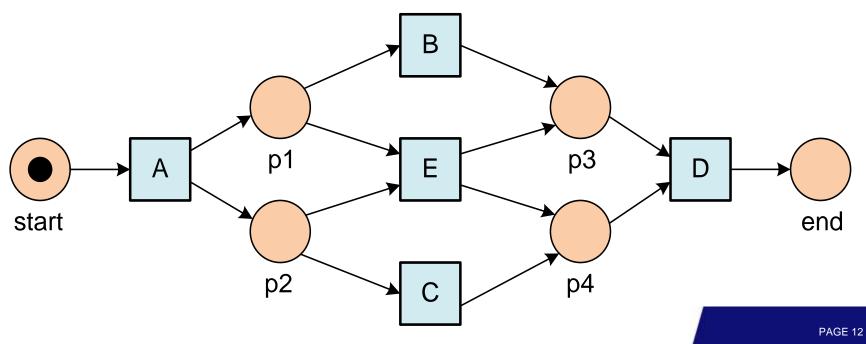


event log

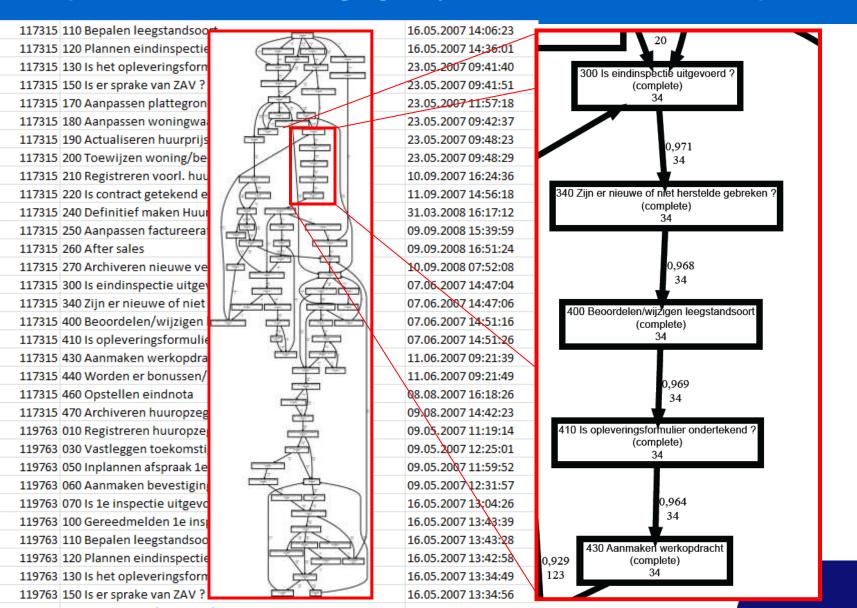
process model



ABCD AED AED ACBD ABCD ACBD AED ACBD

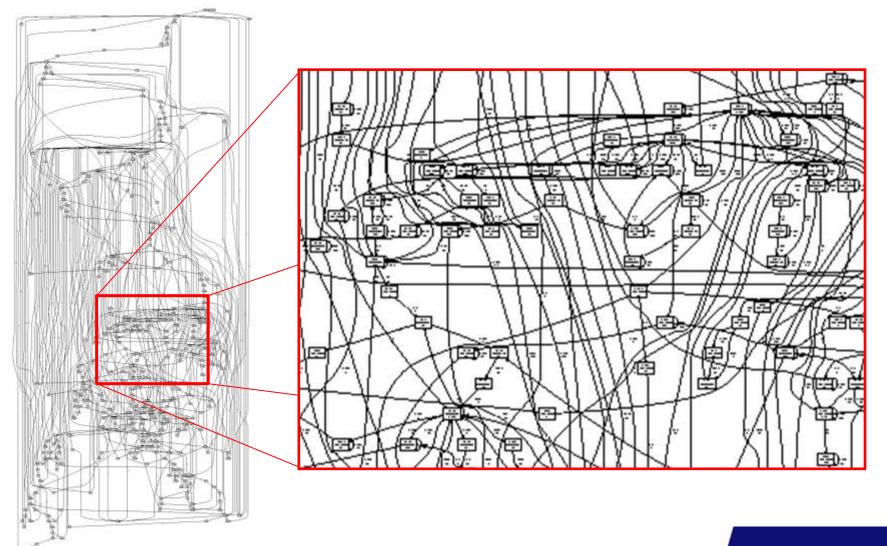


Example Process Discovery (Vestia, Dutch housing agency, 208 cases, 5987 events)

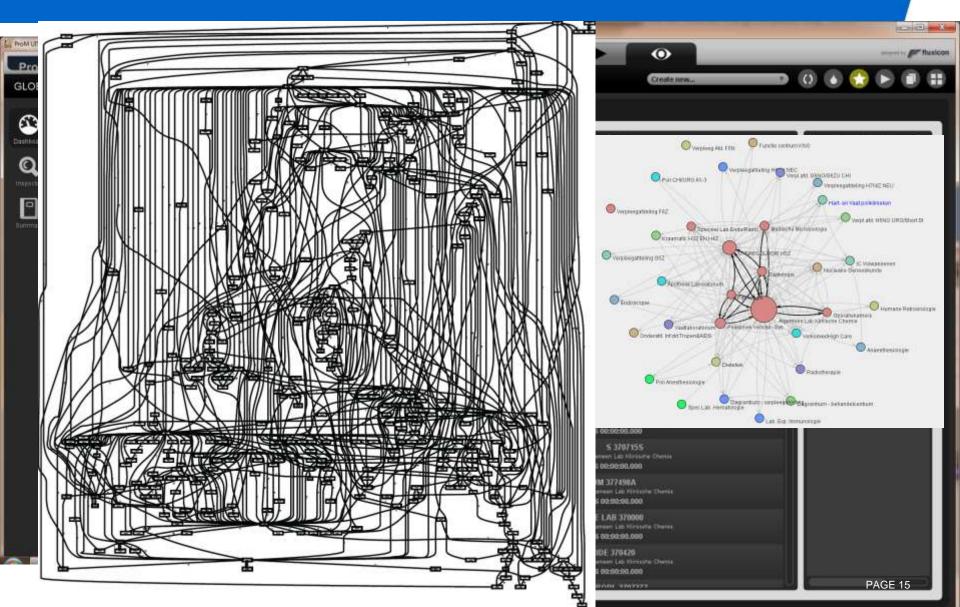


PAGE 13

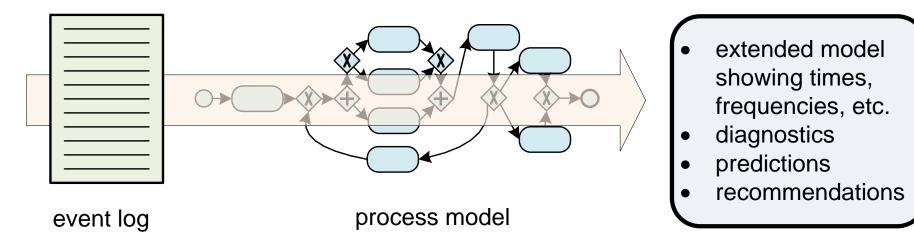
Example Process Discovery (ASML, test process lithography systems, 154966 events)



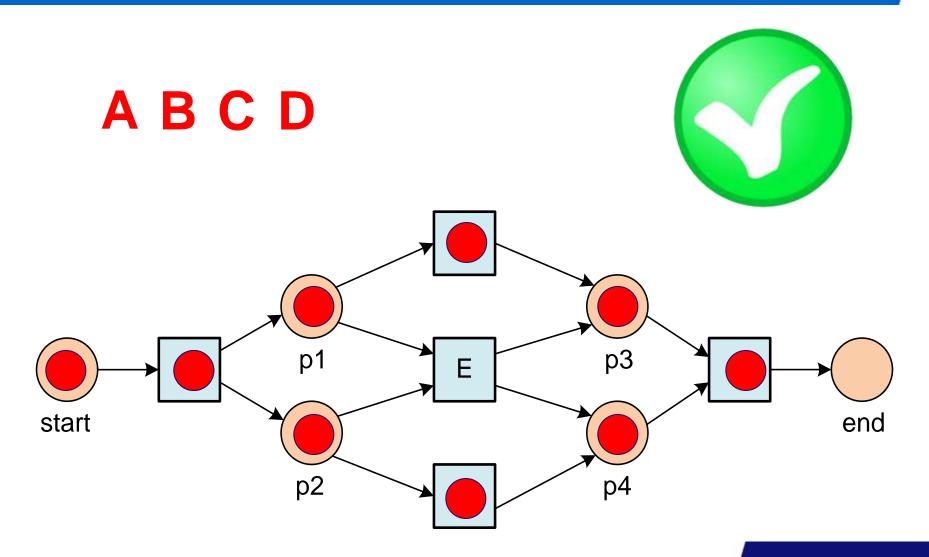
Example Process Discovery (AMC, 627 gynecological oncology patients, 24331 events)



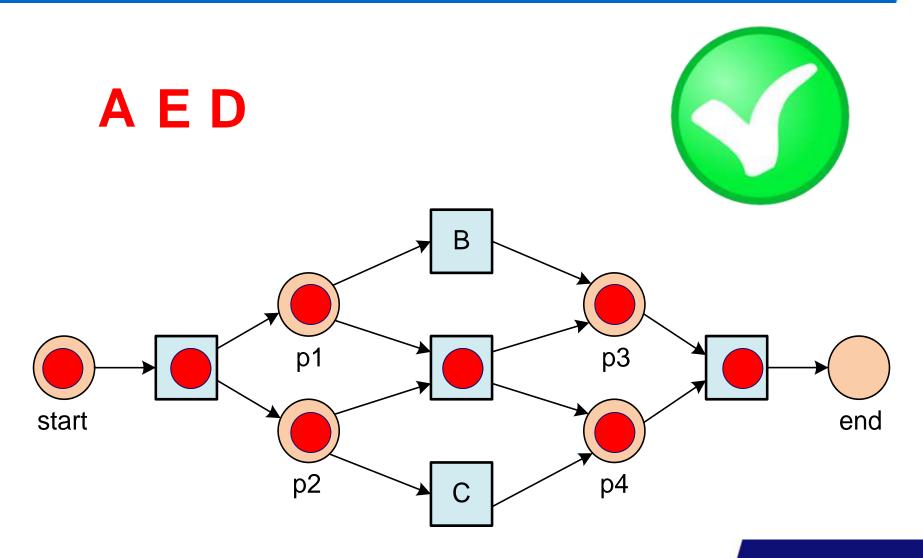




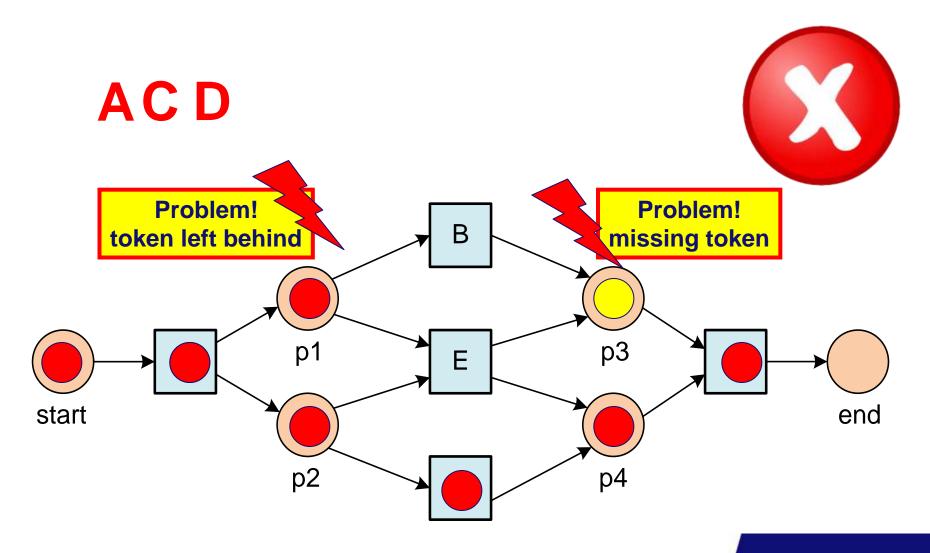


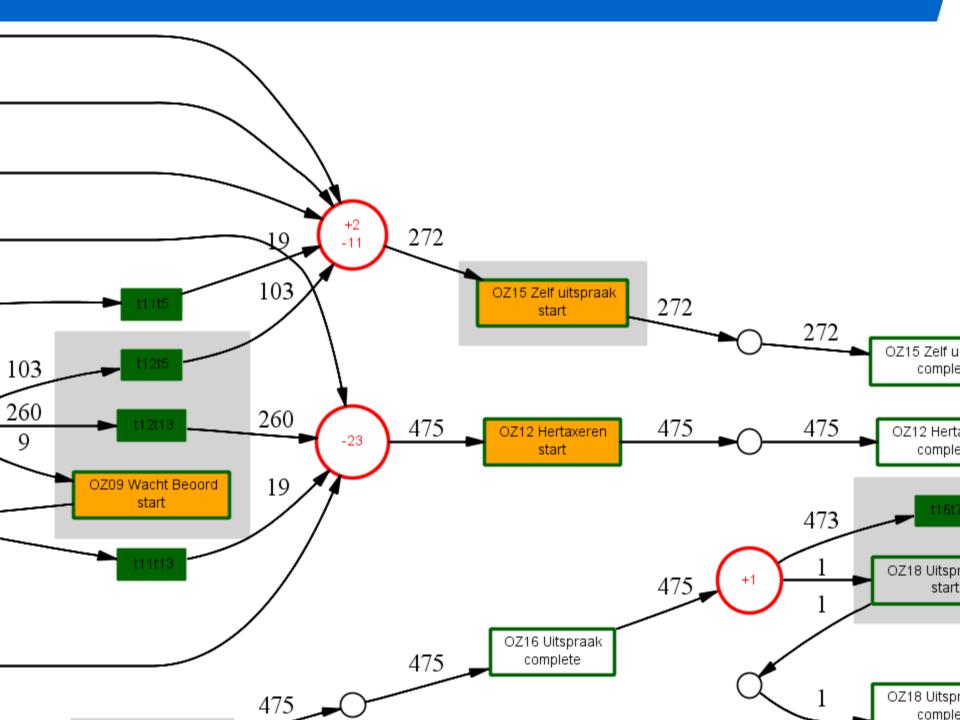




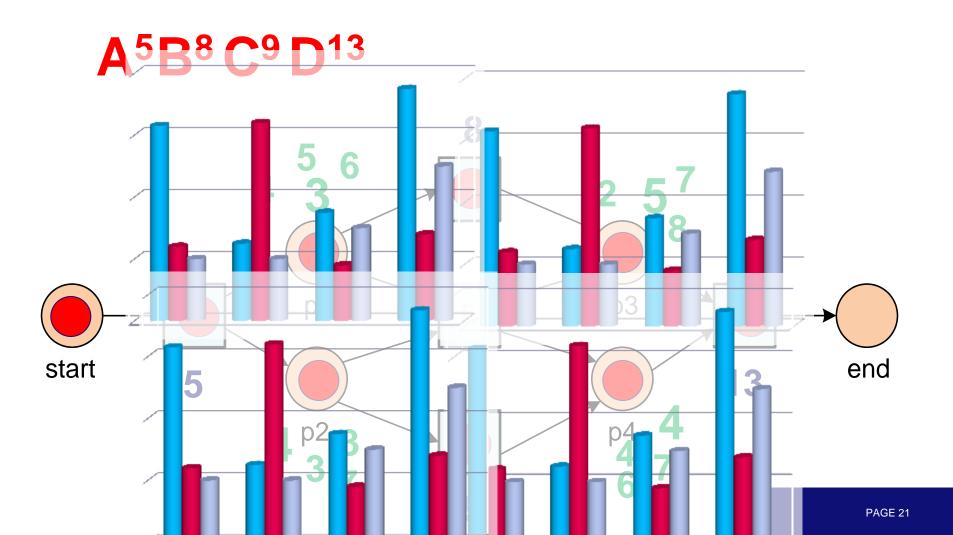


Replay can detect problems

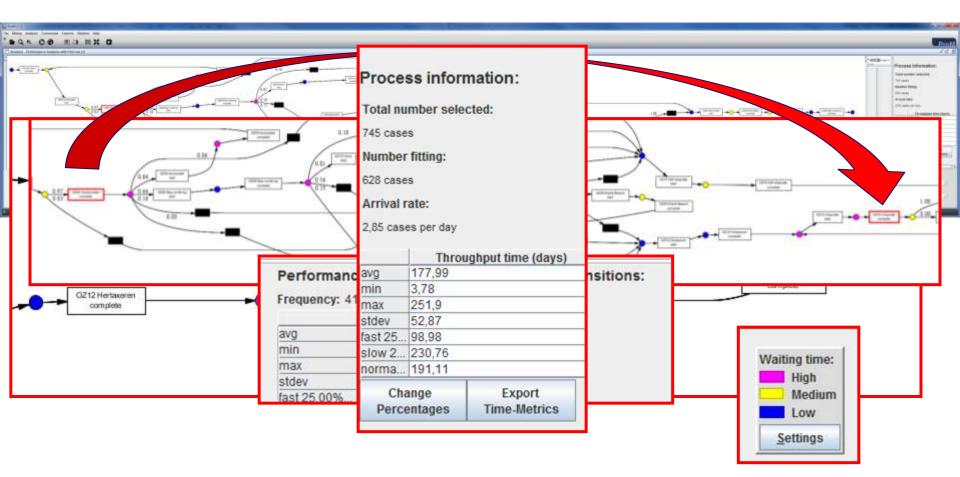




Replay can extract timing information



Performance Analysis Using Replay (WOZ objections Dutch municipality, 745 objections, 9583 event, f= 0.988)



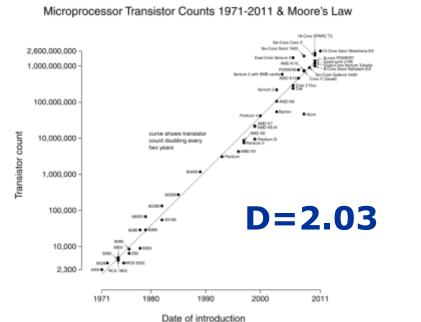
Desire Lines in Big Data

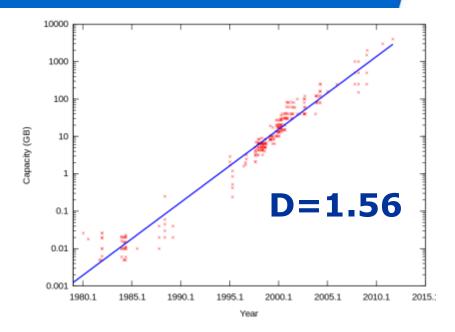






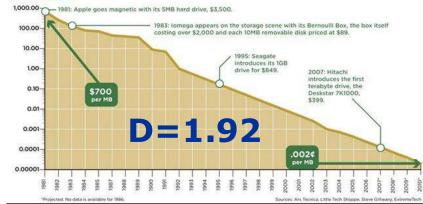
Moore's Law





STORAGE: FROM HIGHWAY ROBBERY TO RUNAWAY BARGAIN

\$ per megabyte



A simple calculation

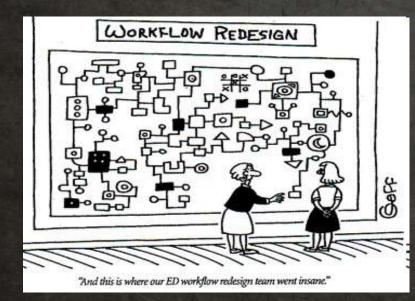


• Starting point 2010:

- Harddisk 1 Terabyte = 10¹² bytes
- **Digital Universe 1.2 Zettabyte = 1.2*10²¹ bytes** (estimate in IDC's annual report, "The Digital Universe Decade Are You Ready?" May 2010)
- Disk needs to grow $2^{30.16} = 1.2^* 10^9 = 1.2^* 10^{21} / 10^{12}$ times its current size.
- Assuming D=1.56 this takes 30.16*1.56 = 47.05 years.
- Hence, in 2060 your laptop can contain all of today's digital universe (internet, computer files, transaction logs, movies, photos, music, books, databases, etc.)!

Evidence-Based Business Process Management

CRIME SCENE DO NOT CROSS

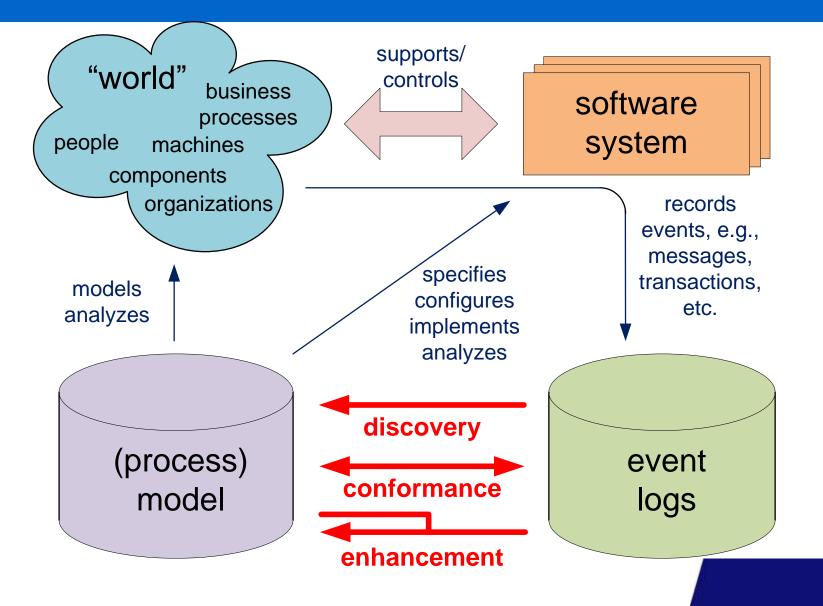


ME SCENE DO NOT CROSS

CRIMA

Process Mining

Process Mining

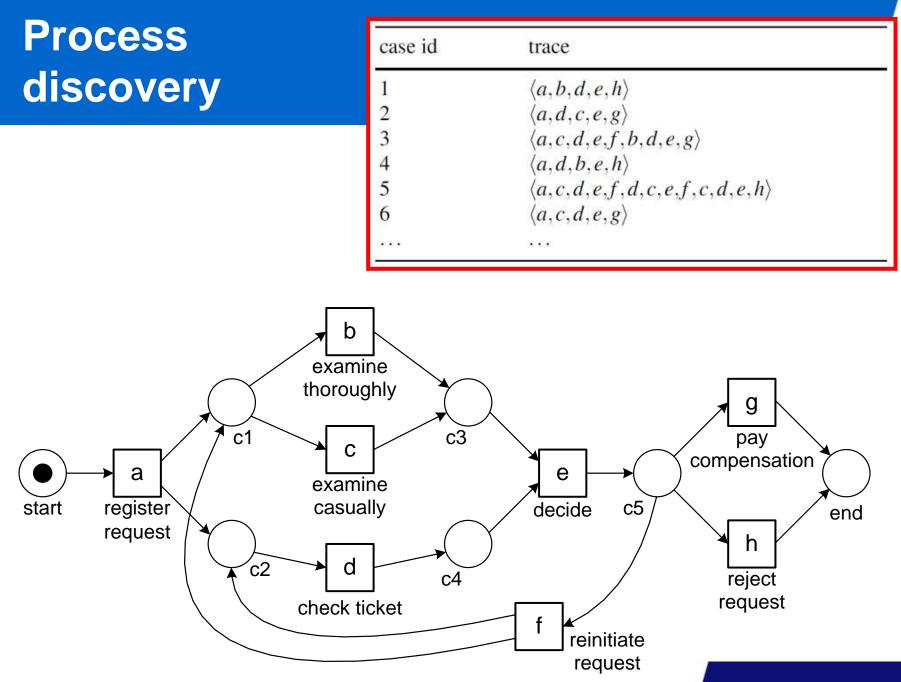


Starting point: event log

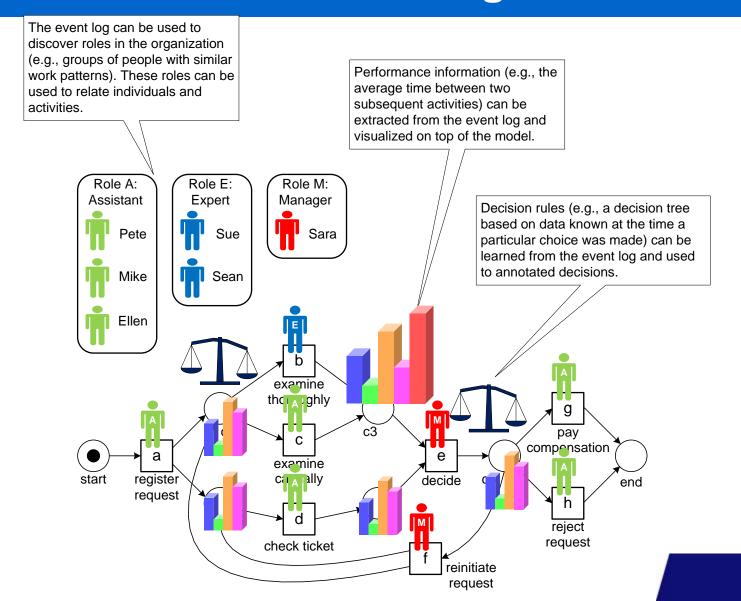
ase id	event id		properties								
		timestamp	activity	resource	cost	111-3					
1	35654424 35654425 35654426	30-12-2010:11.02 31-12-2010:10.06 05-01-2011:15.12 06-01-2011:11.18 07-01-2011:14.24	register request examine thoroughly check ticket decide reject request	Pete Sue Mike Sara Pete	50 400 100 200 200	*** *** *** ***					
2		30-12-2010:11.32 30-12-2010:12.12	register request check ticket	Mike Mike	50	110					
	35654487 30-12-2010:14.16 35654488 05-01-2011:11.22 35654489 08-01-2011:12.05		examine casually decide pay compensation	c	ase id	event id	11	properties			
3	35654522	30-12-2010:14.32 30-12-2010:15.06	register request examine casually	-			timestamp	activity	resource	cost	
	35654525	30-12-2010:16.34 06-01-2011:09.18	check ticket decide			35654423	30-12-2010:11.02	register request	Pete	50	
		06-01-2011:12.18 06-01-2011:13.06	reinitiate request examine thoroughly		1	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	
	35654530	08-01-2011:11.43	check ticket			35654425		check ticket	Mike	100	
		09-01-2011:09.55	decide pay compensation			35654426		decide	Sara	200	
		06-01-2011:15.02	register request	-		35654427		reject request	Pete	200	111
4	35654643	07-01-2011:12.06	check ticket			55054427	07-01-2011.14.24	reject request	rete	200	* * *
		08-01-2011:14.43 09-01-2011:12.02	examine thoroughly decide	(1)		35654483	30-12-2010:11.32	register request	Mike	50	
		12-01-2011:15.44	reject request		2	35654485		check ticket	Mike	100	
8		06-01-2011:09.02	register request		2	35654487	30-12-2010:12.12	examine casually	Pete	400	
5		07-01-2011:10.16 08-01-2011:11.22	examine casually check ticket								6.6.5
		10-01-2011:13.28	decide			35654488		decide	Sara	200	
		11-01-2011:16.18 14-01-2011:14.33	reinitiate request check ticket			35654489	08-01-2011:12.05	pay compensation	Ellen	200	
		16-01-2011:15.50	examine casually	33				14		1.20	
		19-01-2011:11.18	decide	Sara	200	(12)					
		20-01-2011:12.48	reinitiate request	Sara	200	4.1.4					
		21-01-2011:09.06	examine casually	Sue	400	12.53					
		21-01-2011:11.34 23-01-2011:13.12	check ticket decide	Pete Sara	100						
		24-01-2011:14.56	reject request	Mike	200	**************************************					
	35654871	06-01-2011:15.02	register request	Mike	50	***				001	- 1 -
6		06-01-2011:16.06	examine casually	Ellen	400	*** :	XE	ES, MXML, SA [,]		USV.	etc
	35654874	07-01-2011:16.22	check ticket	Mike	100			,, •	- ,	,	
		07-01-2011:16.52	decide	Sara	200						
	35654877	16-01-2011:11.47	pay compensation	Mike	200						
									A.		PAGE 32
		1111							A		AGE 3

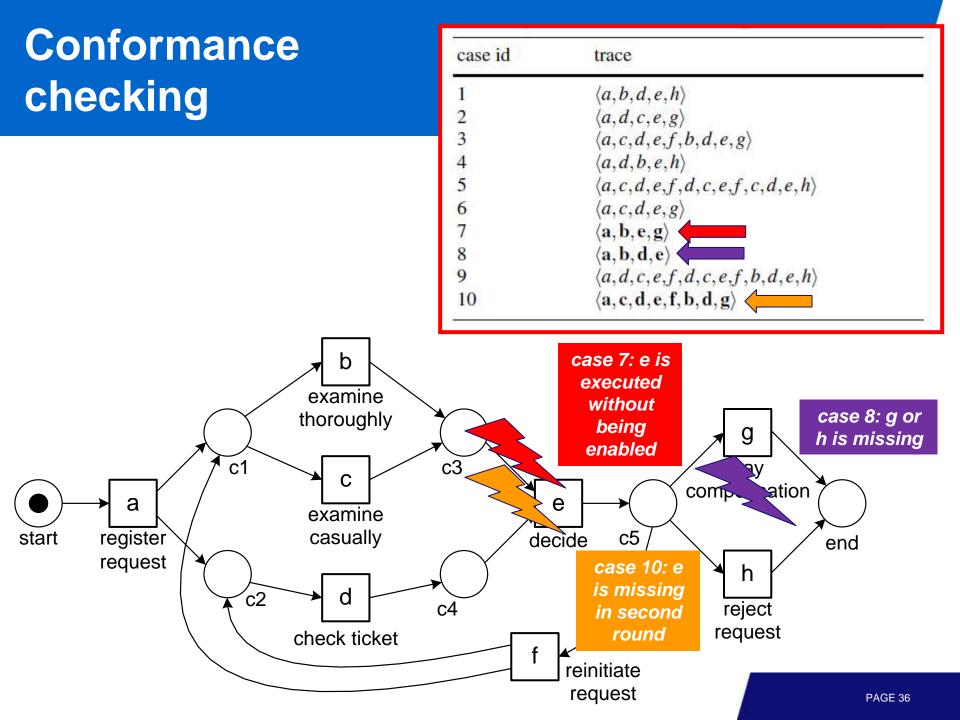
Simplified event log

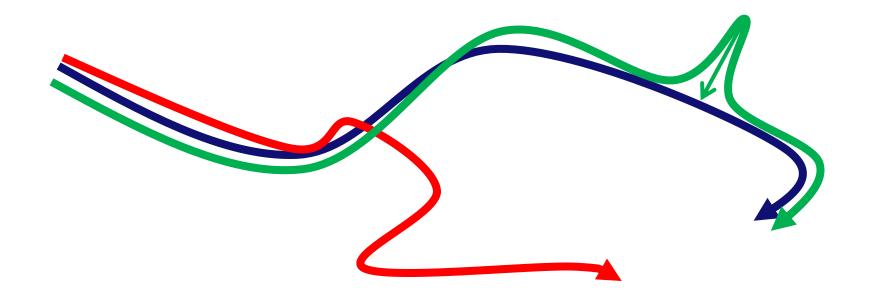
case id	event id		properties			
		timestamp	activity	resource	case i	id trace
I	35654424 35654425 35654426	30-12-2010:11.02 31-12-2010:10.06 05-01-2011:15.12 06-01-2011:11.18 07-01-2011:14.24	register request examine thoroughly check ticket decide reject request	Pete Sue Mike Sara Pete	1	$\langle a, b, d, e, h \rangle$
2	35654485 35654487 35654488	30-12-2010:11.32 30-12-2010:12.12 30-12-2010:14.16 05-01-2011:11.22 08-01-2011:12.05	register request check ticket examine casually decide pay compensation	Mike Mike Pete Sara Ellen	2 3 4	$\begin{array}{l} \langle a,d,c,e,g \rangle \\ \langle a,c,d,e,f,b,d,e,g \rangle \\ \langle a,d,b,e,h \rangle \end{array}$
3	35654522 35654524 35654525 35654526 35654527 35654527 35654530 35654531	$\begin{array}{c} 30\text{-}12\text{-}2010\text{:}14\text{,}32\\ 30\text{-}12\text{-}2010\text{:}15.06\\ 30\text{-}12\text{-}2010\text{:}16.34\\ 06\text{-}01\text{-}2011\text{:}09\text{,}18\\ 06\text{-}01\text{-}2011\text{:}121.18\\ 06\text{-}01\text{-}2011\text{:}13.06\\ 08\text{-}01\text{-}2011\text{:}11\text{,}13\\ 09\text{-}01\text{-}2011\text{:}11\text{,}13\\ 09\text{-}01\text{-}2011\text{:}10\text{,}55\\ 15\text{-}01\text{-}2011\text{:}10\text{,}45 \end{array}$	register request examine casually check ticket decide reinitiate request examine thoroughly check ticket decide pay compensation	Pete Mike Ellen Sara Sean Pete Sara Ellen	5 6 	$\langle a, c, d, e, f, d, c, e, f, c, d, e, h \rangle$ $\langle a, c, d, e, g \rangle$
4	35654643 35654644 35654645	06-01-2011:15.02 07-01-2011:12.06 08-01-2011:14.43 09-01-2011:12.02 12-01-2011:15.44	register request check ticket examine thoroughly decide reject request	Pete Mike Sean Sara Ellen	50 100 400 200 200	a – register request
5	35654712 35654714 35654715 35654716 35654718 35654719 35654720 35654720 35654721 35654722 35654724 35654725 35654726	06-01-2011:09.02 07-01-2011:10.16 08-01-2011:11.22 10-01-2011:11.22 10-01-2011:16.18 14-01-2011:16.18 14-01-2011:11.33 16-01-2011:11.55 09-01-2011:11.18 20-01-2011:01.12.48 21-01-2011:01.14 23-01-2011:11.31 24-01-2011:11.31 24-01-2011:13.12	register request examine casually check ticket decide reinitiate request check ticket examine casually decide reinitiate request examine casually check ticket decide reject request	Ellen Mike Pete Sara Sara Ellen Mike Sara Sara Sue Pete Sara Mike	50 100 200 200 200 200 400 200 200 200 200 200 200 200 200 200 50	a = register request, b = examine thoroughly, c = examine casually, d = check ticket, e = decide, f = reinitiate request,
6	35654873 35654874 35654875	06-01-2011:15.02 06-01-2011:16.06 07-01-2011:16.22 07-01-2011:16.52 16-01-2011:11.47	register request examine casually check ticket decide	Mike Ellen Mike Sara Mike	50 400 100 200 200	g = pay compensation, and h = reject request
	330,940//	10-01-2011:11:47	pay compensation	MINE	200	PAGE 33



Extension: Adding perspectives to model based on event log



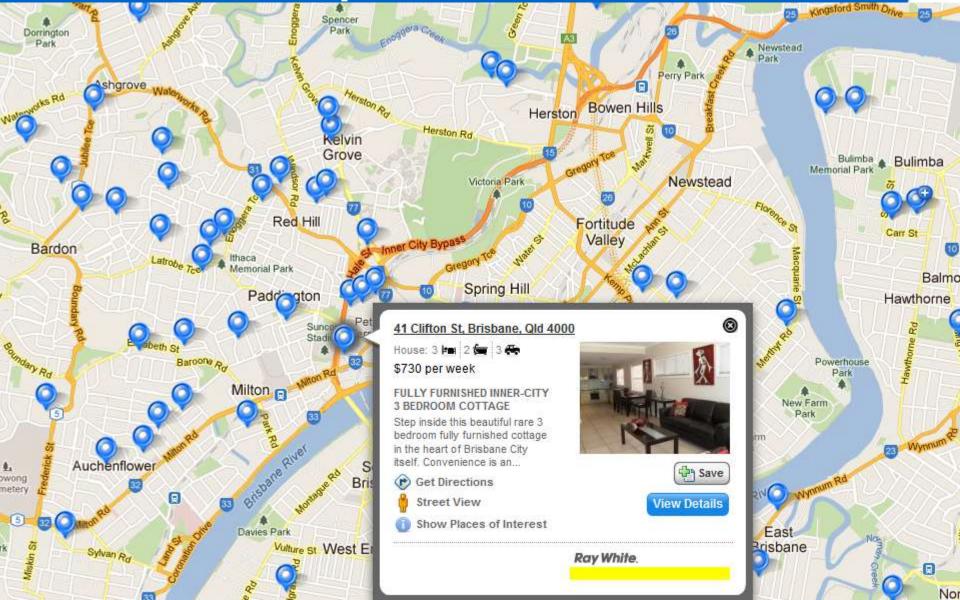




Alignments are essential!

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

Process models should be treated as electronic maps



Food for Thought

We applied ProM in >100 organizations

- 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, azM hospital, Isala clinics, GGzE)
- Multinationals (e.g., DSM, Deloitte)
- High-tech system manufacturers and their customers (e.g., Philips Healthcare, ASML, Ricoh, Thales)
- Media companies (e.g. Winkwaves)

How can process mining help?

- Uncover bottlenecks
- Detect deviations
- Performance
 measurement

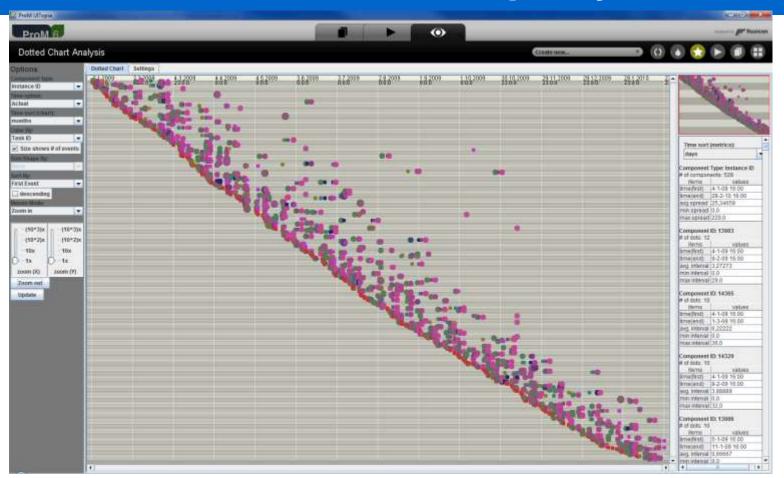
prediction)

- Auditing/compliance
 - Business Process Redesign (BPR) Continuous improvement (Six Sigma) Operational support (e.g., recommendation and

- Provide new insights
- Highlight important
 problems
- An organization's mirror (in two ways)
 Helps to avoid ICT
 - failures
 - Avoid "management by PowerPoint"
 - From "politics" to "analytics"



Example of a Lasagna process: WMO process of a Dutch municipality

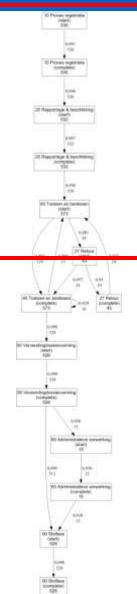


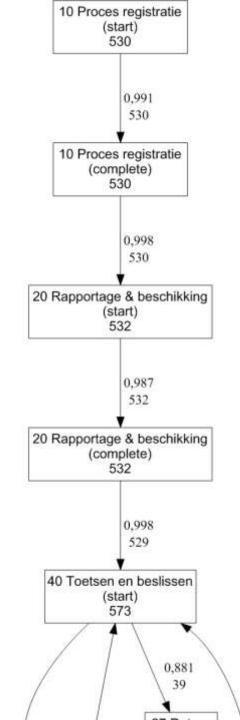
Each line corresponds to one of the 528 requests that were handled in the period from 4-1-2009 until 28-2-2010. In total there are 5498 events represented as dots. The mean time needed to handled a case is approximately 25 days.

WMO process (Wet Maatschappelijke Ondersteuning)

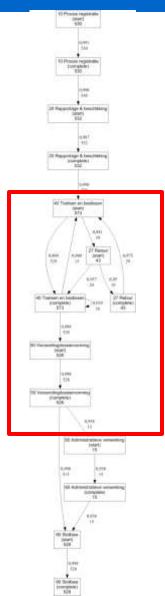
- WMO refers to the social support act that came into force in The Netherlands on January 1st, 2007.
- The aim of this act is to assist people with disabilities and impairments. Under the act, local authorities are required to give support to those who need it, e.g., household help, providing wheelchairs and scootmobiles, and adaptations to homes.
- There are different processes for the different kinds of help. We focus on the process for handling requests for household help.
- In a period of about one year, 528 requests for household WMO support were received.
- These 528 requests generated 5498 events.

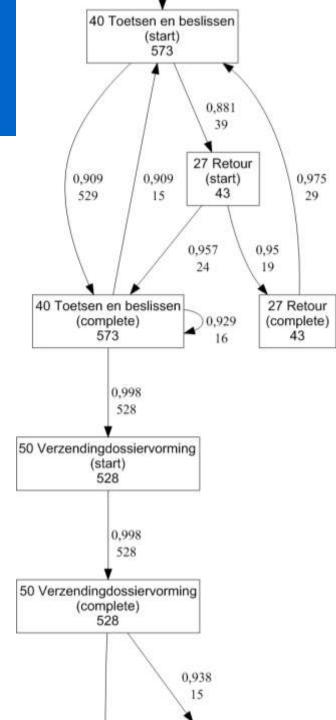
C-net discovered using heuristic miner (1/3)



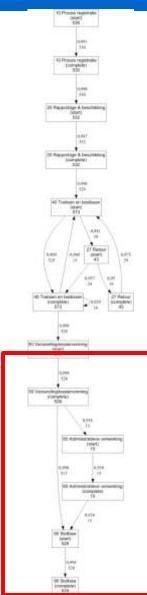


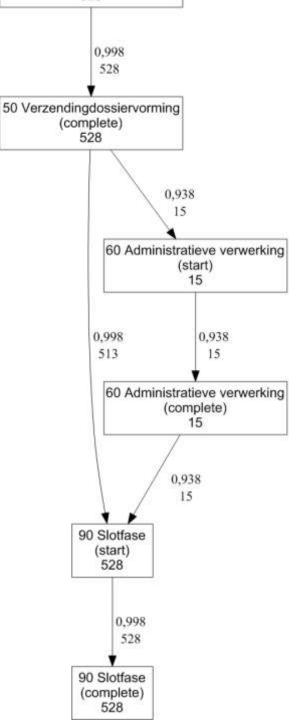
C-net discovered using heuristic miner (2/3)



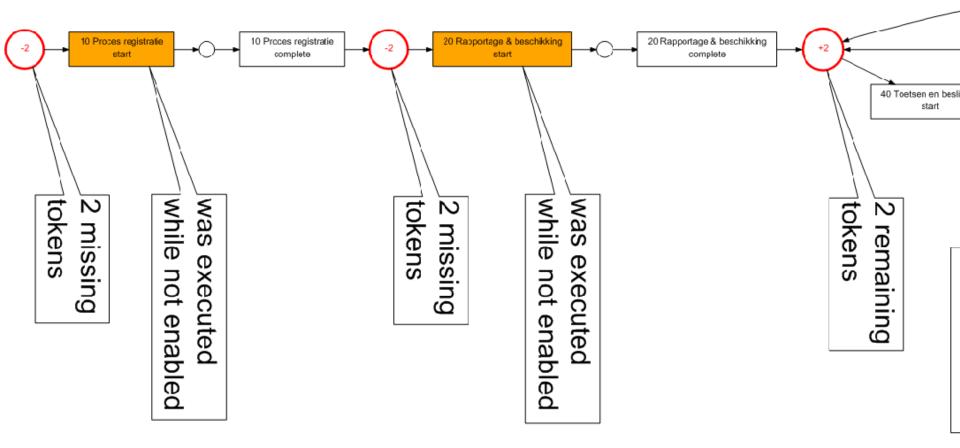


C-net discovered using heuristic miner (3/3)

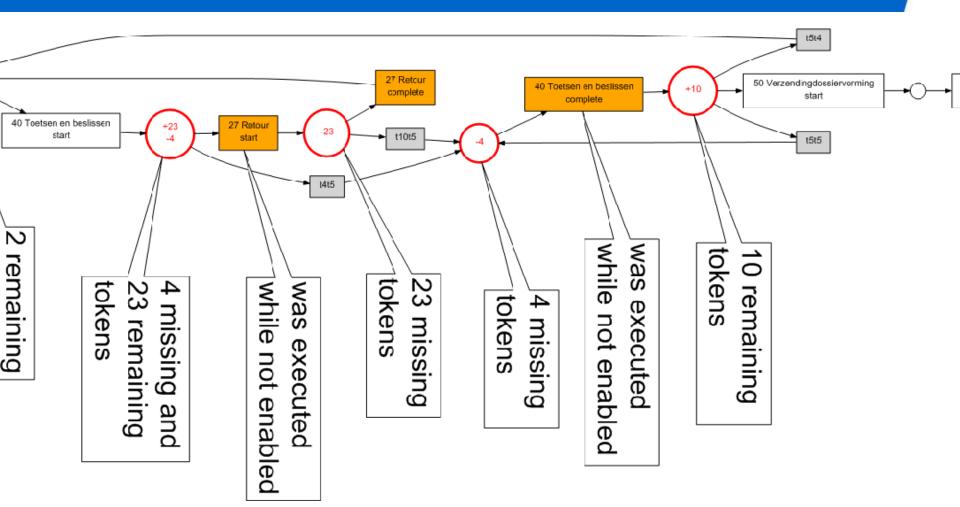




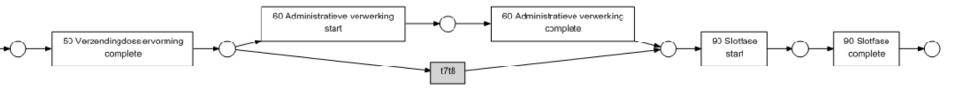
Conformance check WMO process (1/3)



Conformance check WMO process (2/3)

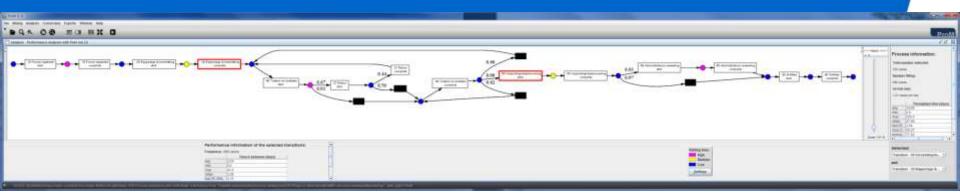


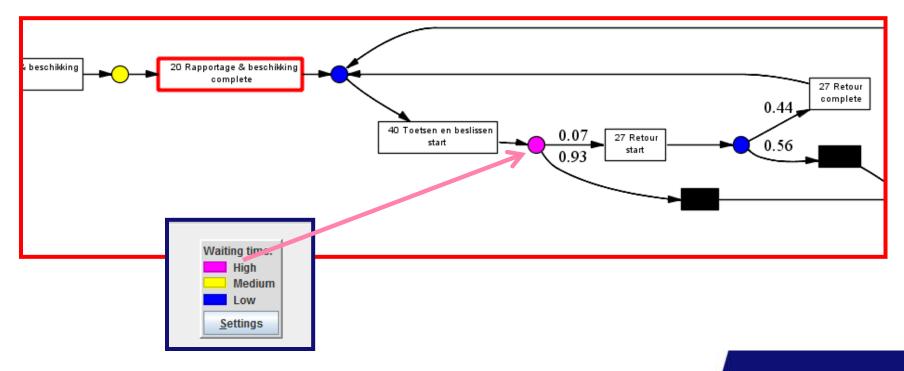
Conformance check WMO process (3/3)



The fitness of the discovered process is 0.99521667. Of the 528 cases, 496 cases fit perfectly whereas for 32 cases there are missing or remaining tokens.

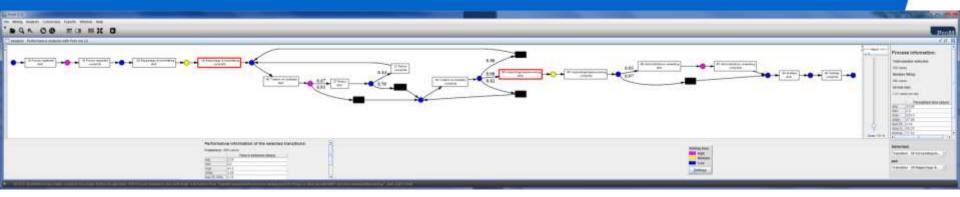
Bottleneck analysis WMO process (1/3)

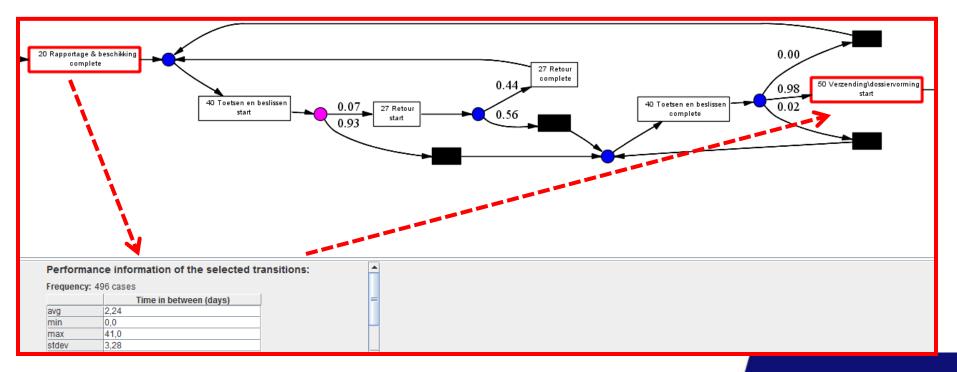




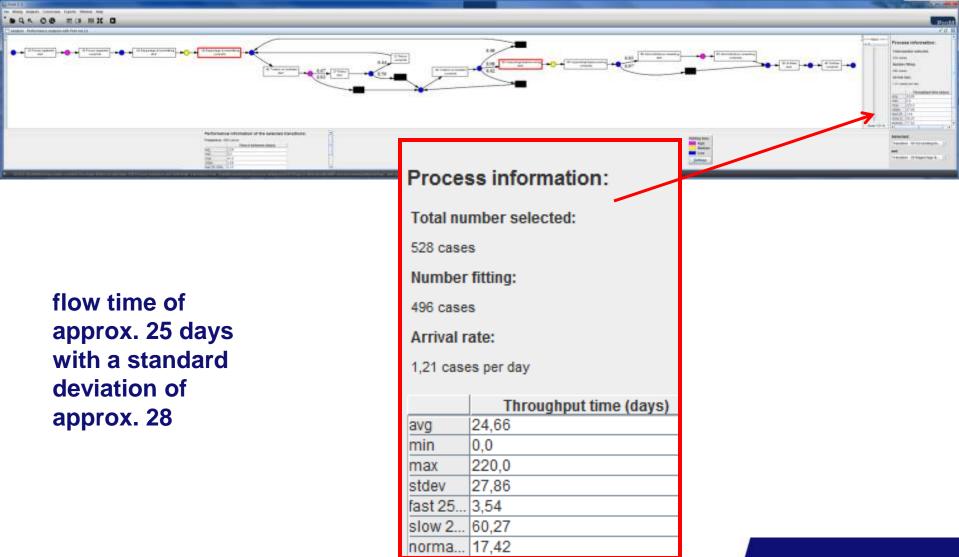
PAGE 51

Bottleneck analysis WMO process (2/3)





Bottleneck analysis WMO process (3/3)



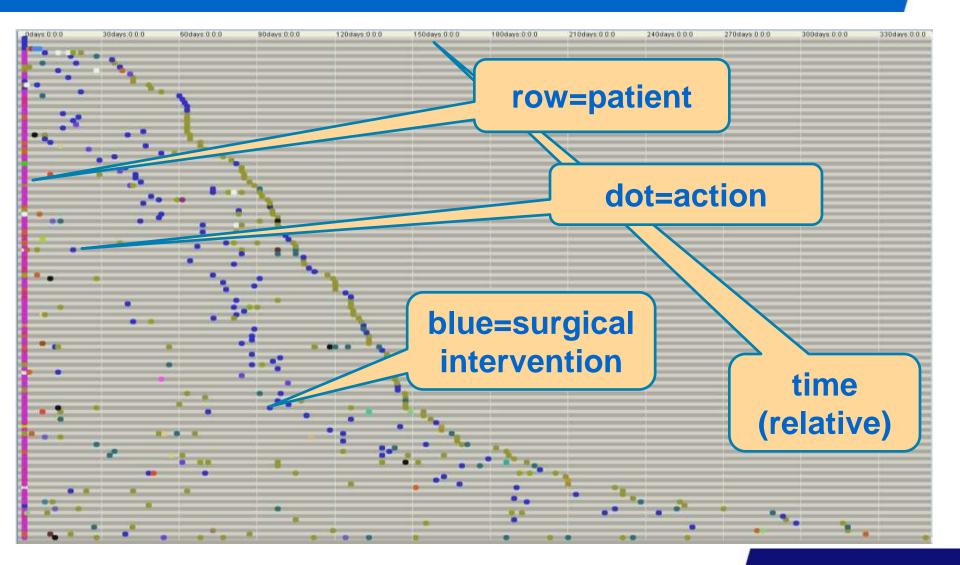
Academic Medical Center (AMC)

- University hospital, Amsterdam
 - >1000 beds
 - 25.000 patients admitted
 - 35.000 day admissions
 - 350.000 outpatient clinic visits
- Process surgery department
 - Diagnosis
 - Surgery
 - Aftercare
- Patient groups considered
 - General complaints stomach (algemene buikklachten n.n.o.)
 - Inguinal hernia (liesbreuk)
 - Varicose veins (spataderen)
 - Hemorroids (aambeien)
 - Lipoma and sebaceous cyst (lipoom (goedaardig gezwel van vetweefselcellen) en atheroom (verstopte talgklier))

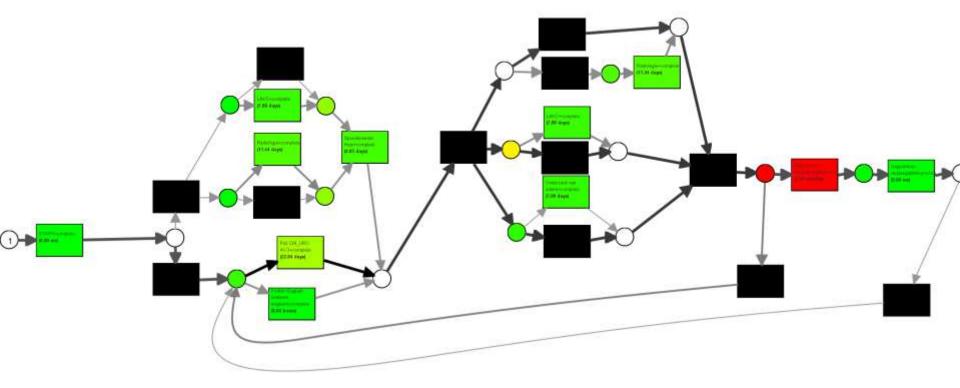




Dotted chart showing flow times hernia inguinalis (liesbreuk) with day admission



Discovered process model with bottlenecks hernia inguinalis (liesbreuk) with day admission

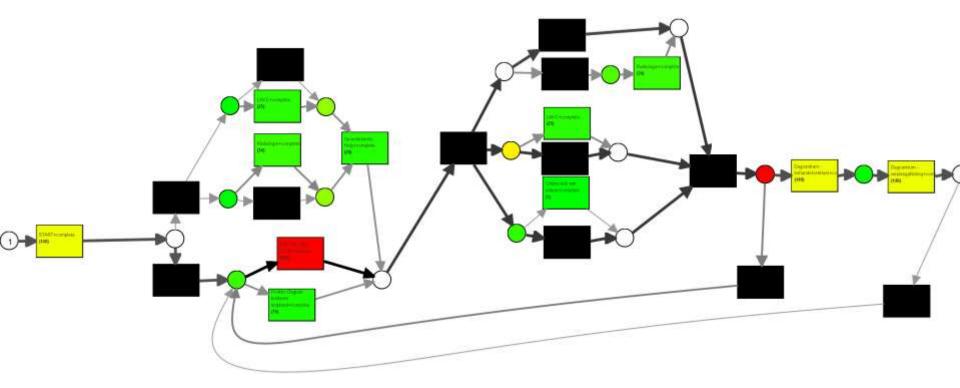


color indicates time

fitness=0.97

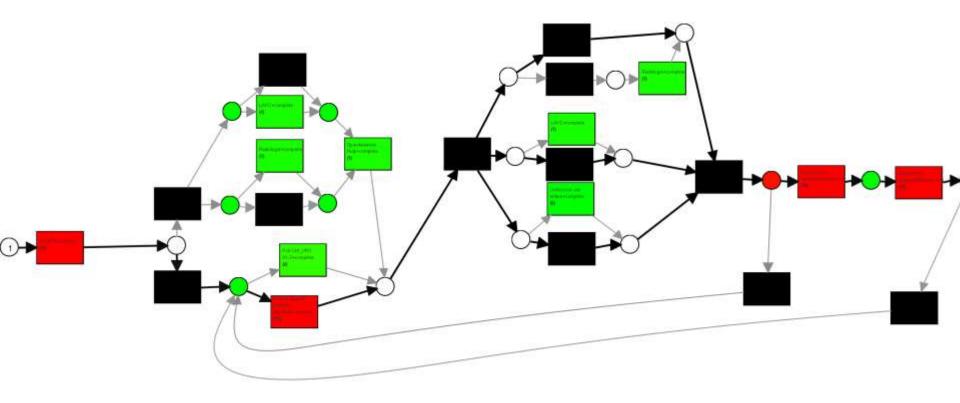
Discovered process model showing frequencies

hernia inguinalis (liesbreuk) with day admission

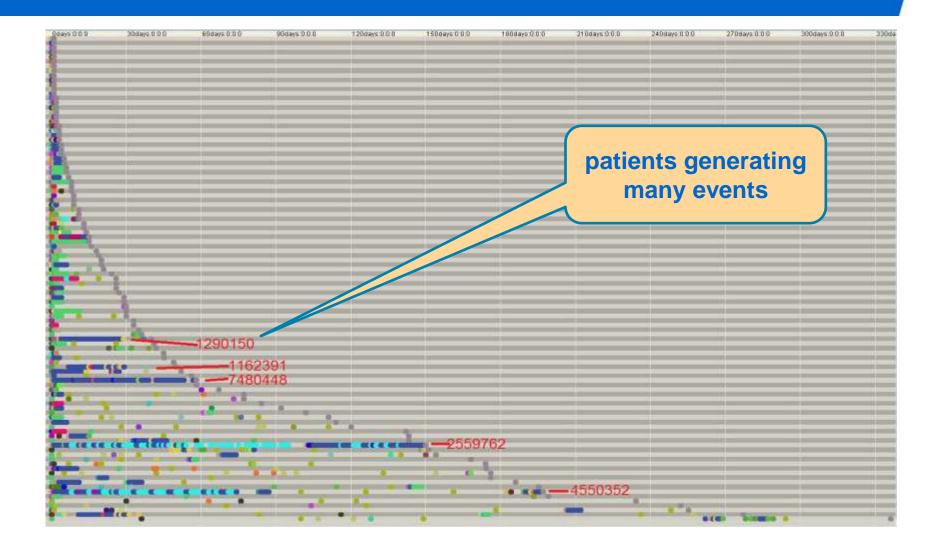


color indicates frequency

Same process but now for only children

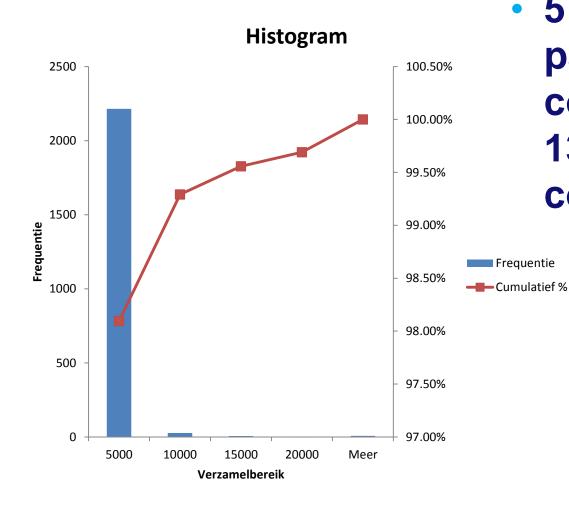


Uneven distribution in number of actions



Number of actions

Total costs (histogram)



 5 of the 2260 patients consume 13% of the costs!

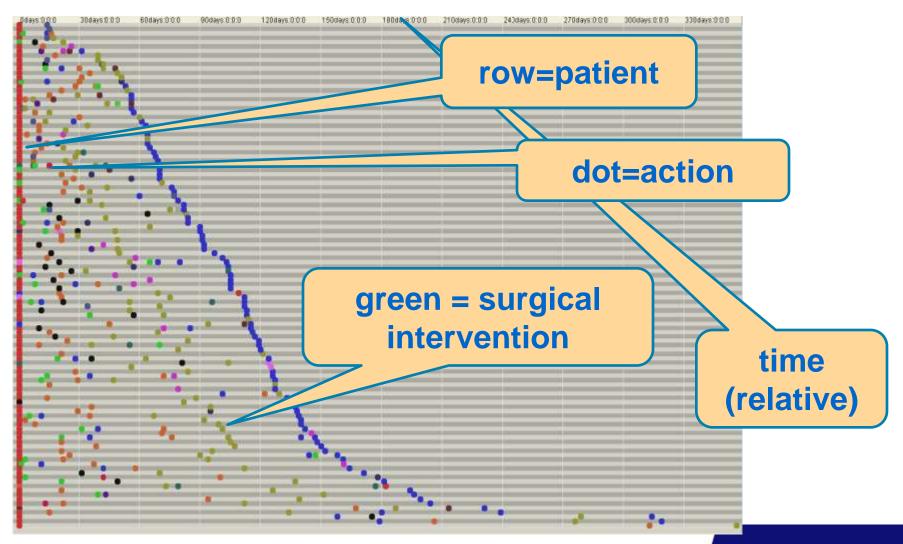
Isala Hospital



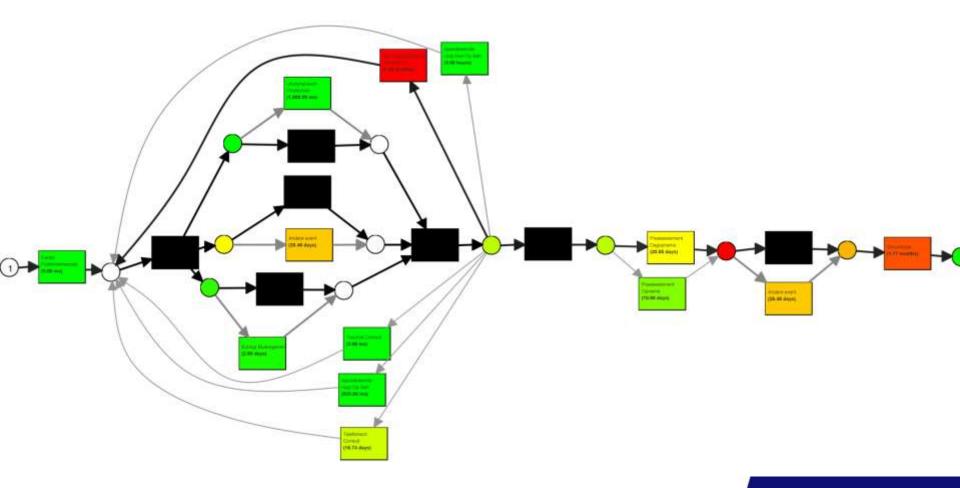
- Top clinical care hospital Zwolle
 - >900 beds
 - > 5.000 employees
- Process urology department
 - Diagnosis
 - Surgery
 - Aftercare
- Patient groups considered
 - Phimosis (voorhuidsvernauwing)
 - Hydrocele (goedaardige zwelling uitgaande van de balzak)
 - Undescended testis (niet ingedaalde testikel)
 - Bladder cancer (blaastumor)
 - Ureter stones (steen in de urineleider)



Dotted chart Phimosis



Process model showing where most time is spent by Phimosis patients





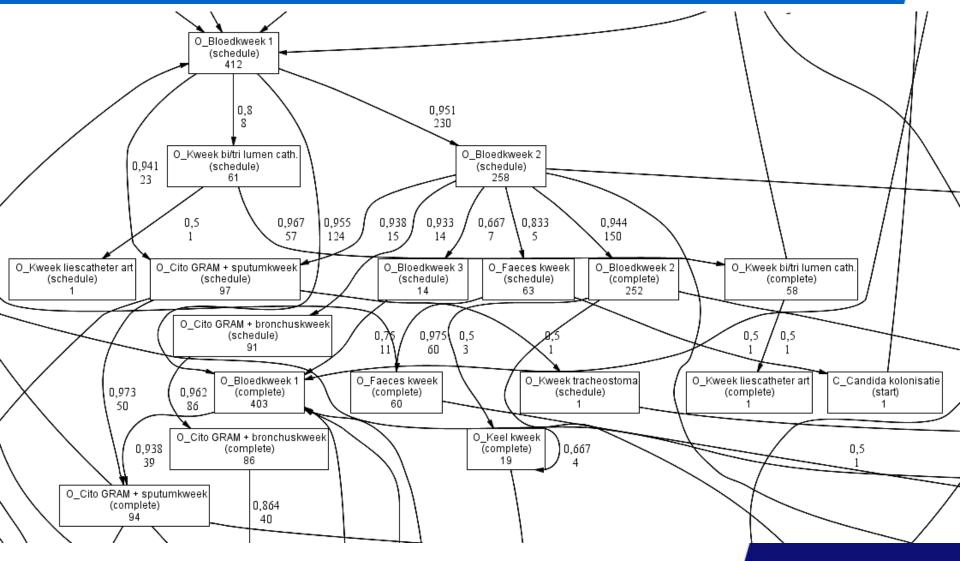
Example of a Spaghetti process





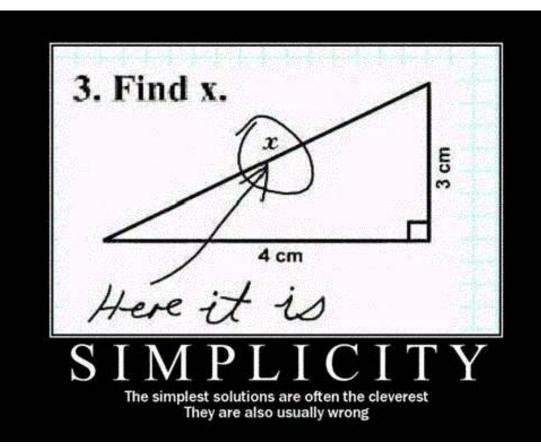
Spaghetti process describing the diagnosis and treatment of 2765 patients in a Dutch hospital. The process model was constructed based on an event log containing 114,592 events. There are 619 different activities (taking event types into account) executed by 266 different individuals (doctors, nurses, etc.).

Fragment 18 activities of the 619 activities (2.9%)



Don't oversimplify

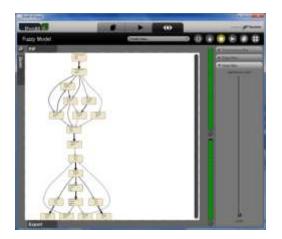
- Reality cannot adequately be described using a simple PowerPoint
- Healthcare cannot be supported by simple logistical principles and mechanistic IT solutions



How to get started?

Hundreds of plug-ins available covering the whole process mining spectrum







open-source (L-GPL)



Download from: www.processmining.org

How to Get Started?

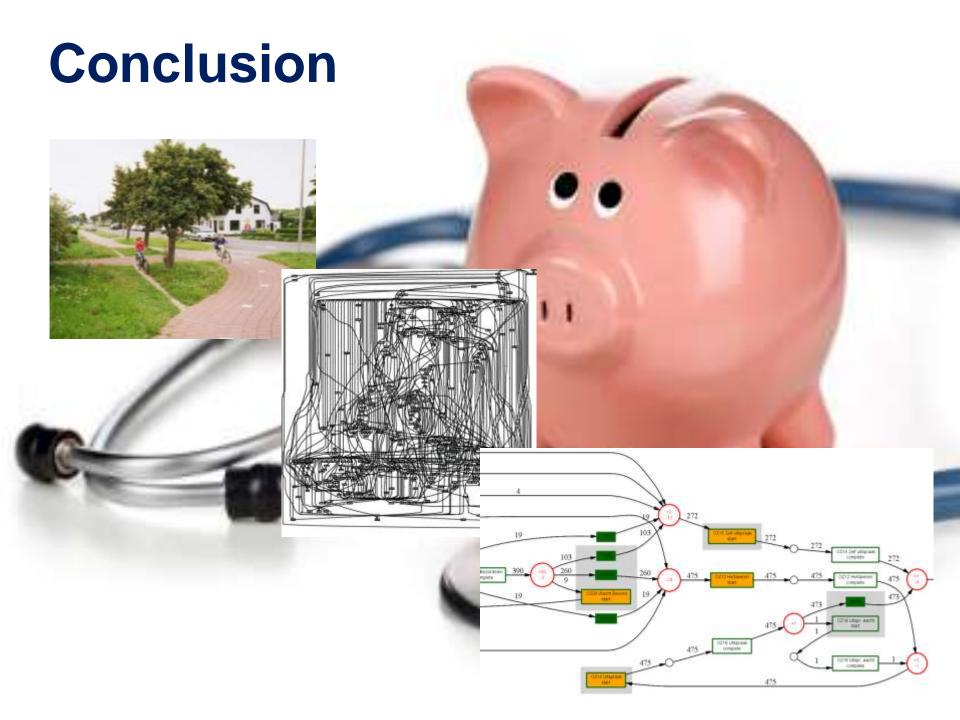
Collect event data

- Minimal requirement: events referring to an activity name and a process instance.
- Good to have: timestamps, resource information, additional data elements.
- Challenges: scoping and sometimes correlation.

Collect questions

- What kind problems would you like to address (cost, time, risk, compliance, service, etc.)?
- Related to discovery, conformance, enhancement?
- Iterative process: can be "curiosity driven" initially.

Conclusion



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.



Process Mining

Wil M. P. van der Aalst

Process Mining

Discovery, Conformance and Enhancement of Business Processes

www.processmining.org

Computer Science ISBN 978-3-642-19344-6 9 7/808574293404

www.win.tue.nl/ieeetfpm/

