Process Mining & Predictive Process Monitoring

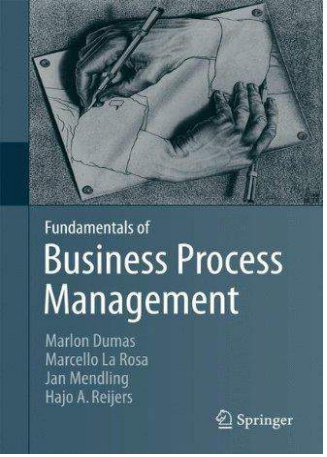
Marlon Dumas

University of Tartu, Estonia

marlon.dumas@ut.ee

UNIVERSITY OF TARTU
INSTITUTE OF COMPUTER SCIENCE

2do Foro BPM, Universidad de los Andes, 22/06/2017



Identification

architecture

Discovery

"as is" models

Analysis

waste, issues,
risks,opportunities

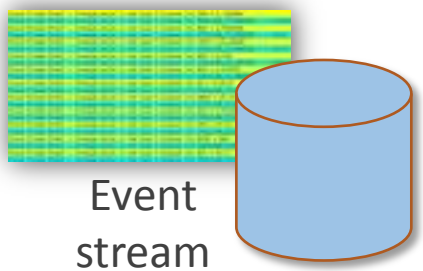
Redesign

"to be" model

Implementation

Business Process
Monitoring

Business Process Monitoring



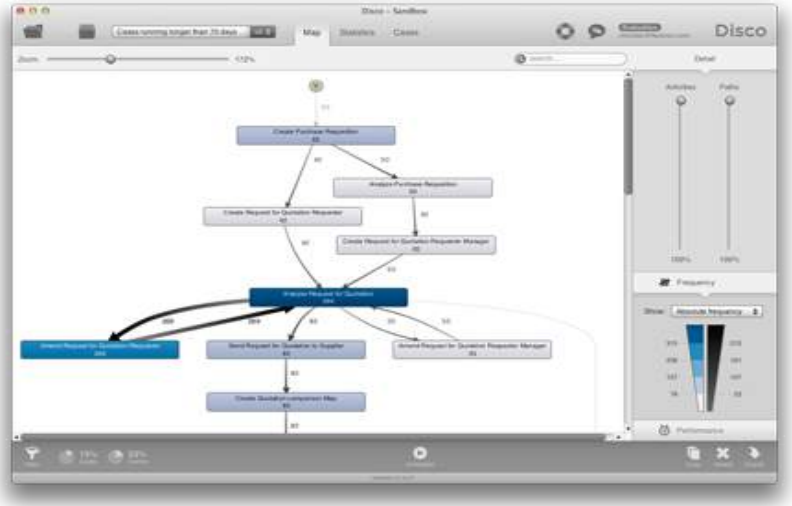
Event stream

DB logs

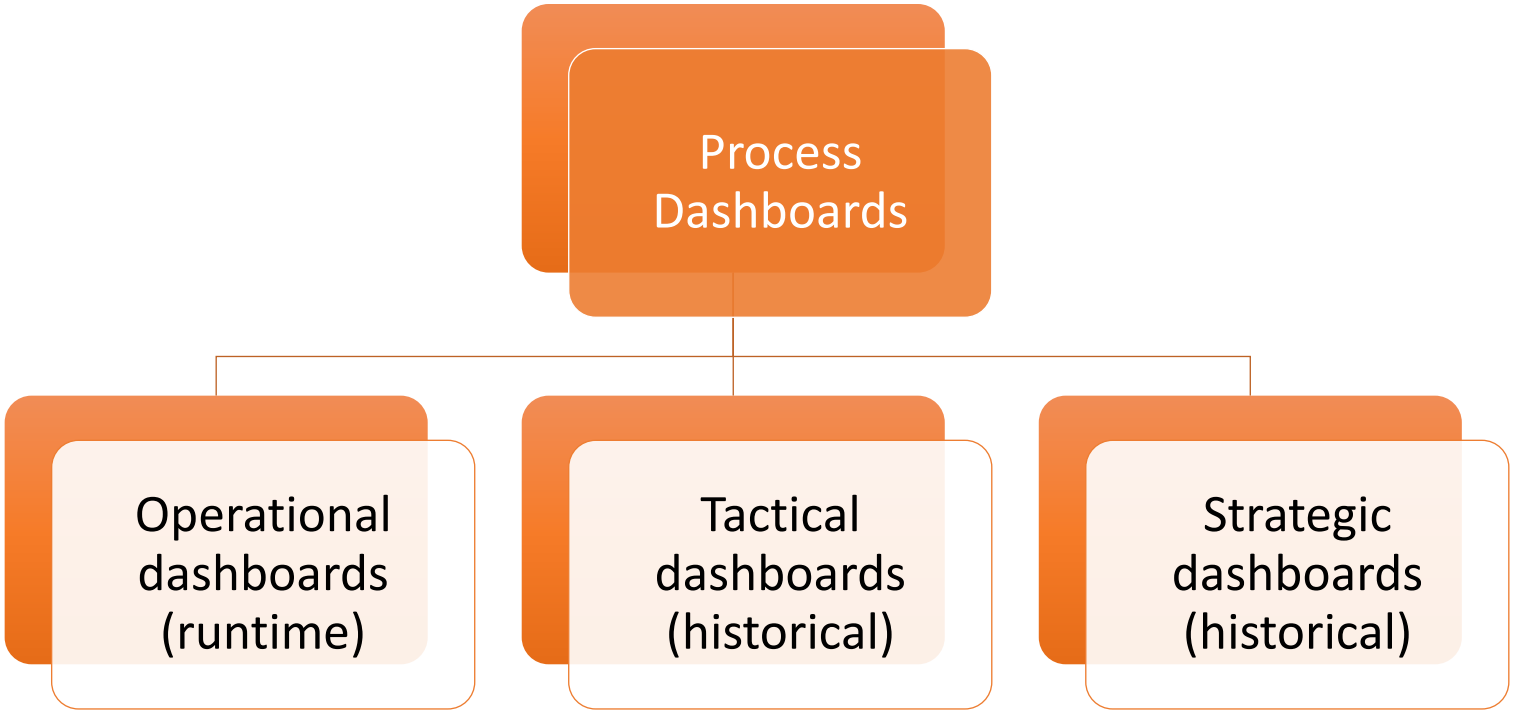
Event log

Dashboards & reports

Process mining



Types of process dashboards

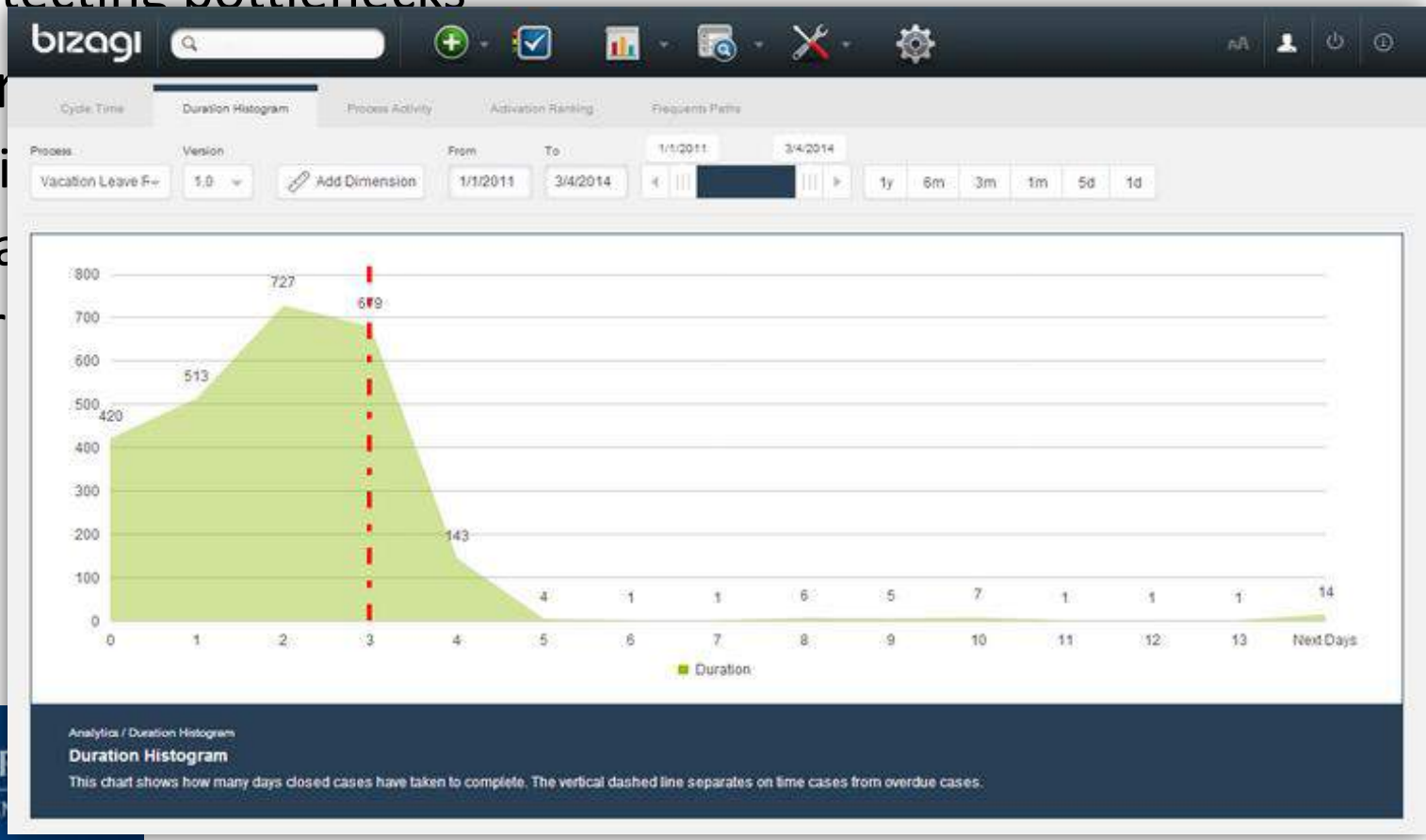


Operational process dashboards

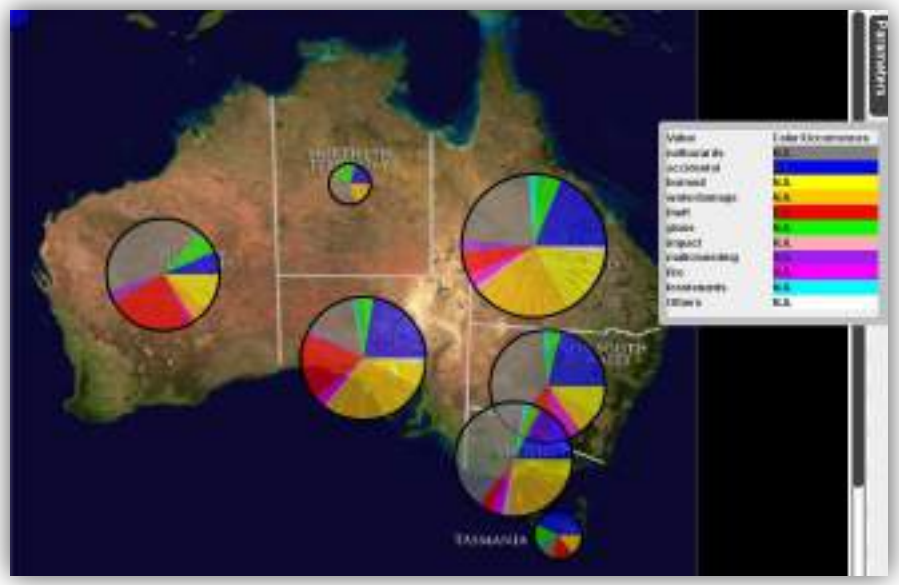
- Aimed at process workers & operational managers
- Emphasis on monitoring (detect-and-respond), e.g.:
  - Work-in-progress
  - Problema
  - Resource

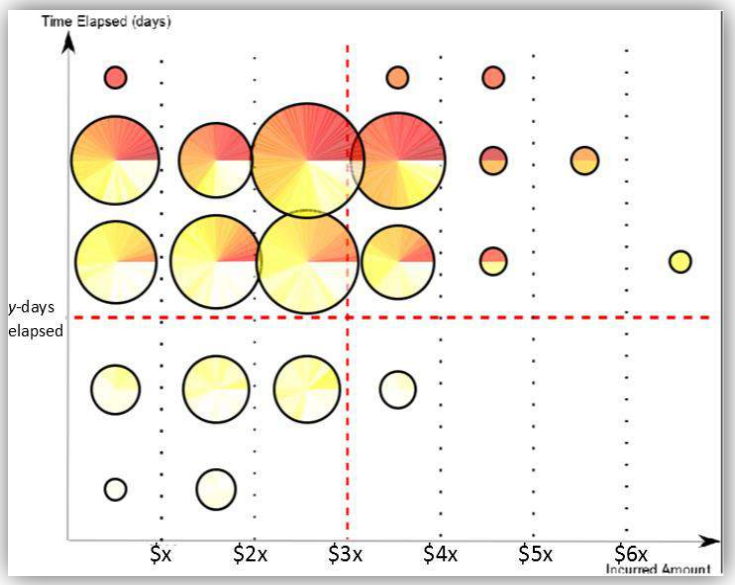
Tactical dashboards

- Aimed at process owners / managers
- Emphasis on analysis and management
  - E.g. detecting bottlenecks
- Typical pr
  - Cycle ti
  - Error ra
  - Resour



Tactical Performance Dashboard @ Australian Insurer



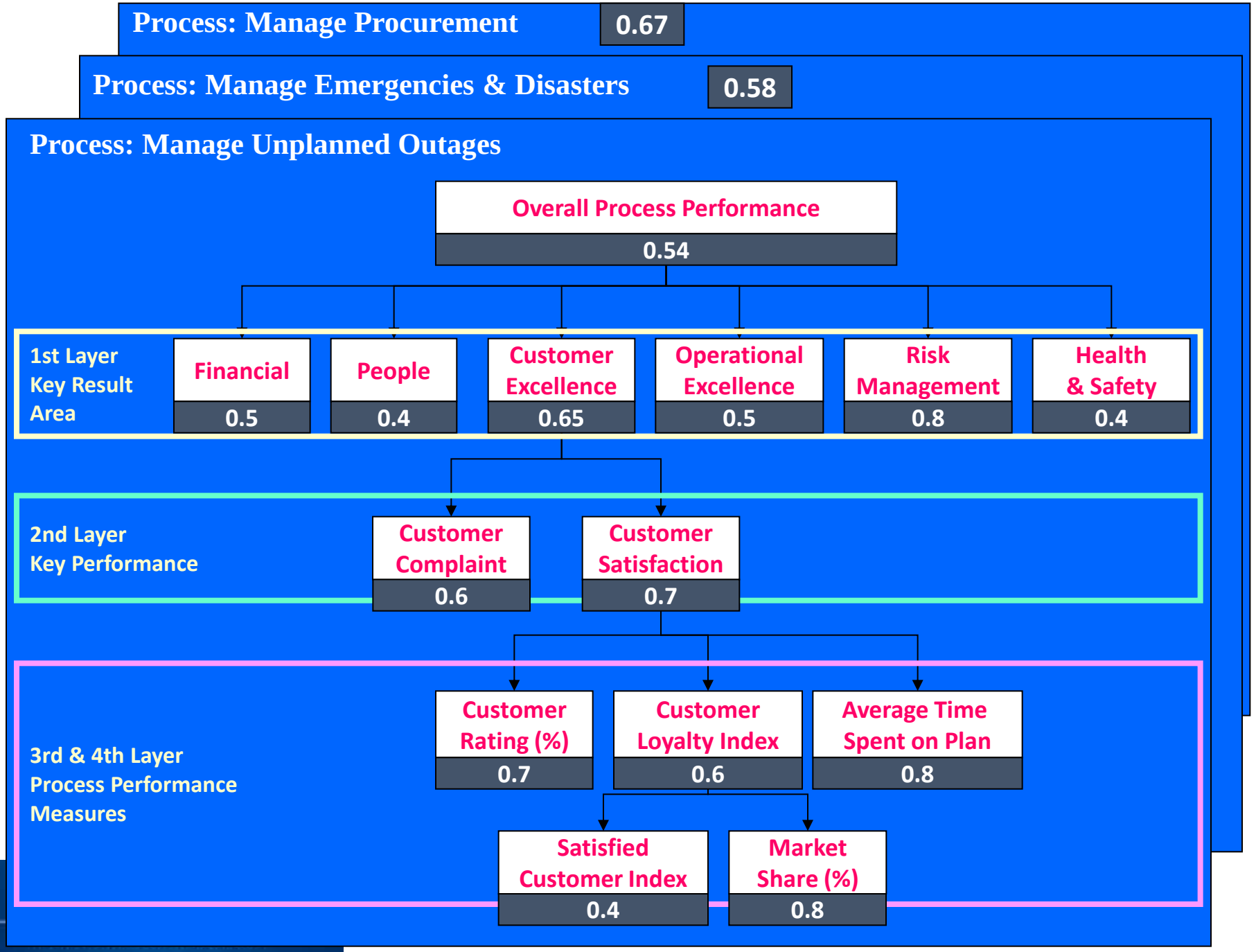


Strategic dashboards

- Aimed at executives & managers
- Emphasis on linking process performance to strategic objectives

Strategic Performance Dashboard @ Australian Utilities Provider

Process / Key Performance	Manage Unplanned Outages	Manage Emergencies & Disasters	Manage Work Programming & Resourcing	Manage Procurement
Customer Satisfaction	0.5	0.55	-	**0.2**
Customer Complaint	0.6	-	-	**0.5**
Customer Feedback	**0.4**	-	-	0.8
Connection Less Than Agreed Time	**0.3**	0.6	0.7	-


Process: Manage Procurement
0.67
Process: Manage Emergencies & Disasters
0.58
Process: Manage Unplanned Outages
Overall Process Performance
0.54
1st Layer Key Result Area
Financial
0.5
People
0.4
Customer Excellence
0.65
Operational Excellence
0.5
Risk Management
0.8
Health & Safety
0.4
2nd Layer Key Performance
Customer Complaint
0.6
Customer Satisfaction
0.7
3rd & 4th Layer Process Performance Measures
Customer Rating (%)
0.7
Customer Loyalty Index
0.6
Average Time Spent on Plan
0.8
Satisfied Customer Index
0.4
Market Share (%)
0.8

PERFORMANCE DASHBOARDS

MEASURING, MONITORING, AND MANAGING YOUR BUSINESS

SECOND EDITION

WAYNE ECKERSON

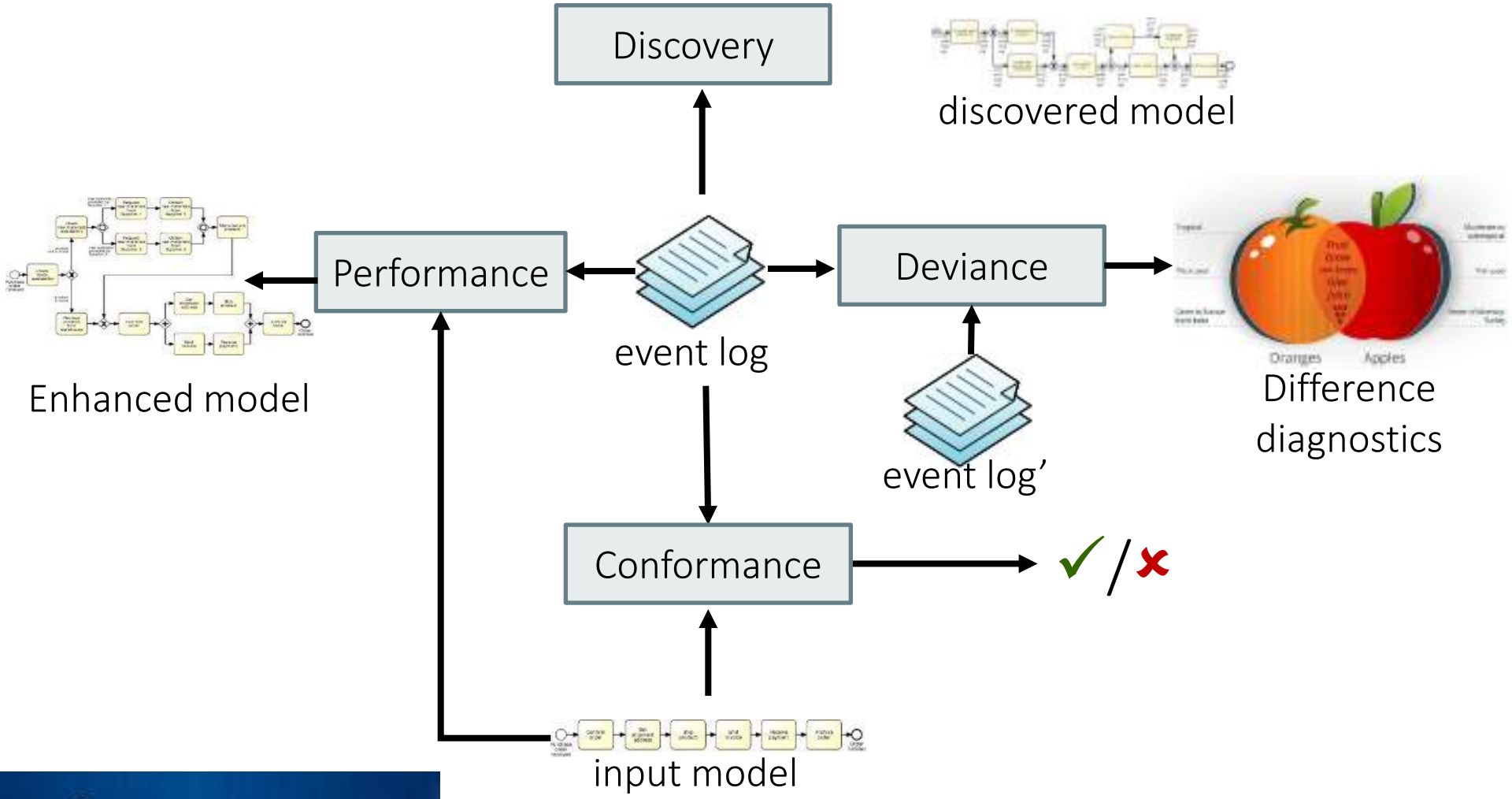
cole nussbaumer knaflic

storytelling with data

a data visualization guide for business professionals

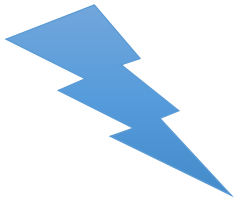
WILEY

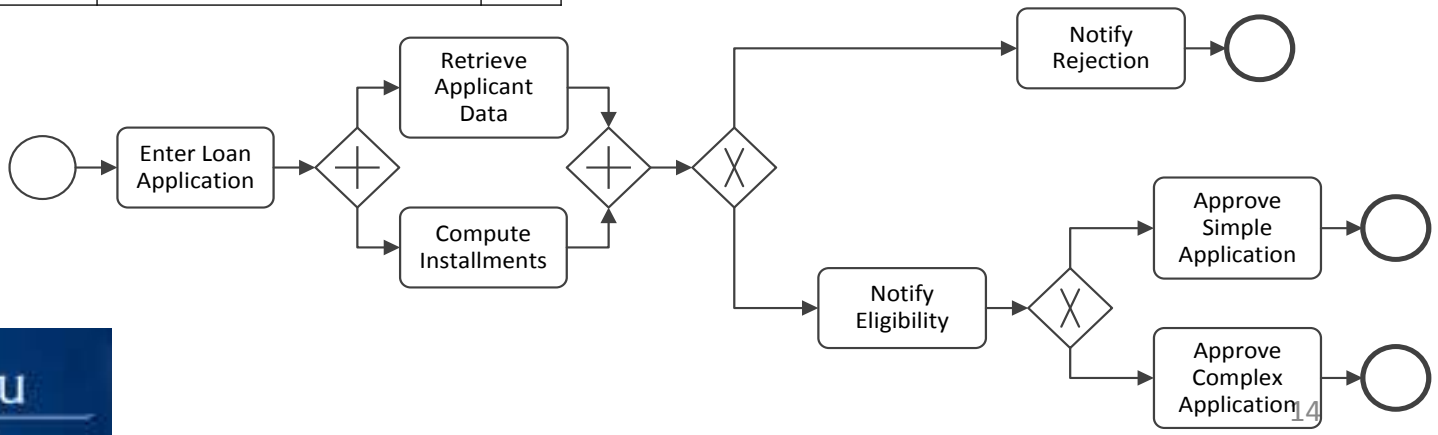
Process Mining



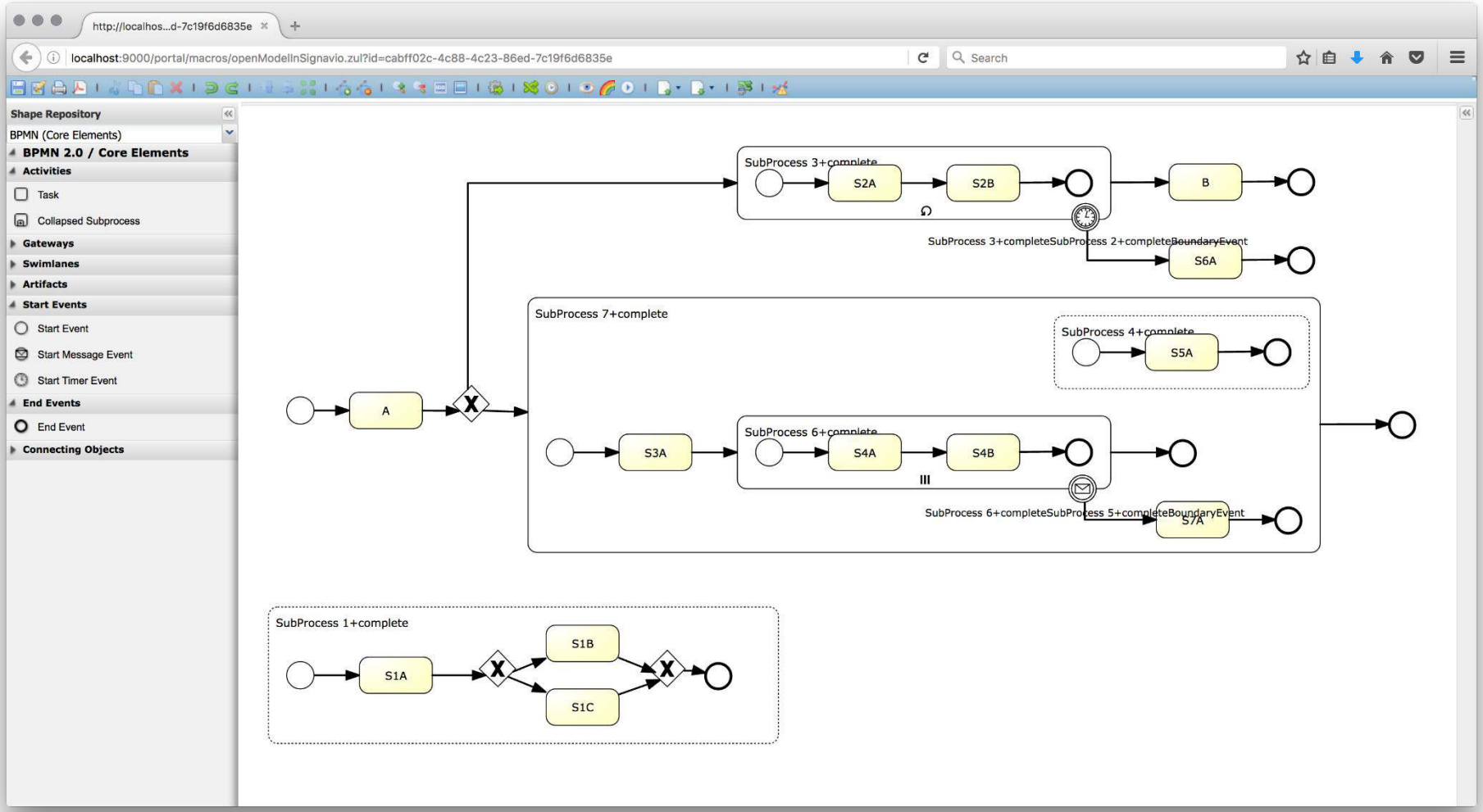
Automated Process Discovery

CID	Task	Time Stamp	...
13219	Enter Loan Application	2007-11-09 T 11:20:10	-
13219	Retrieve Applicant Data	2007-11-09 T 11:22:15	-
13220	Enter Loan Application	2007-11-09 T 11:22:40	-
13219	Compute Installments	2007-11-09 T 11:22:45	-
13219	Notify Eligibility	2007-11-09 T 11:23:00	-
13219	Approve Simple Application	2007-11-09 T 11:24:30	-
13220	Compute Installements	2007-11-09 T 11:24:35	-
...	...	...	...



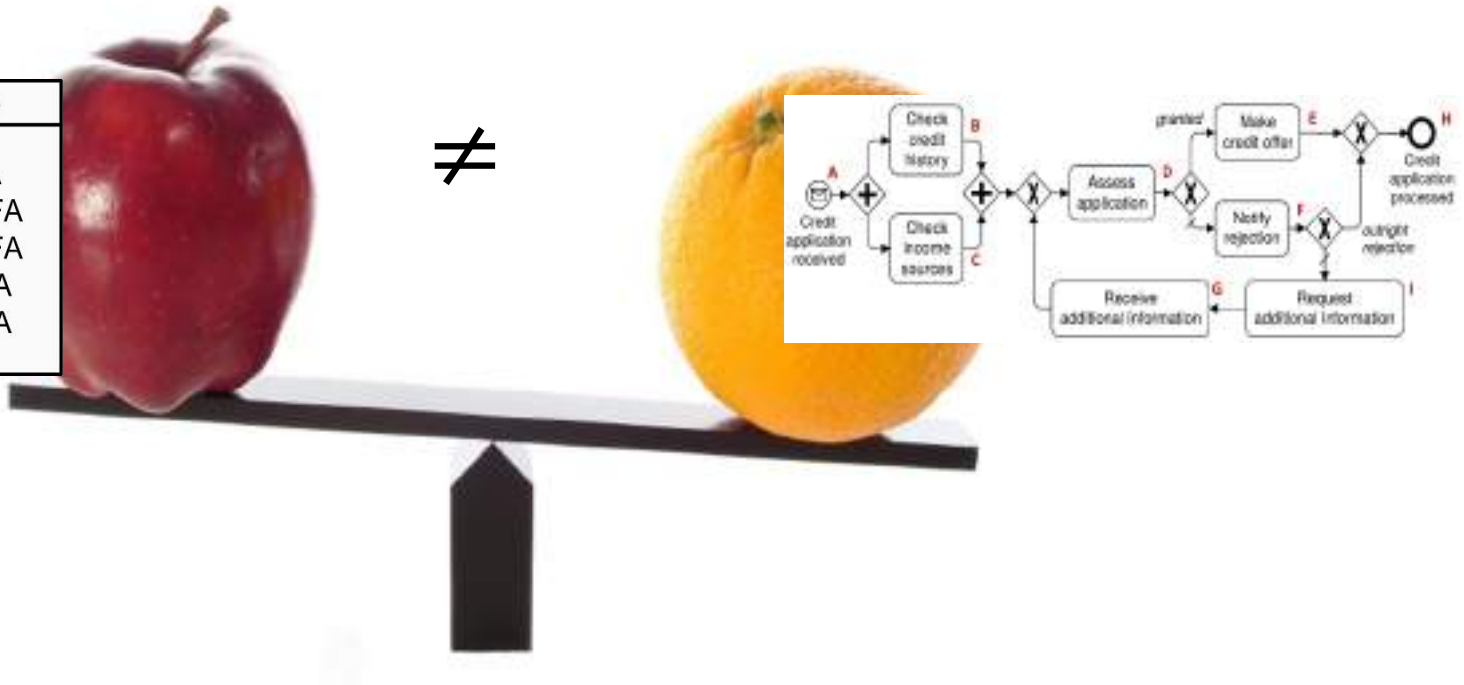


Automated Process Discovery in Action

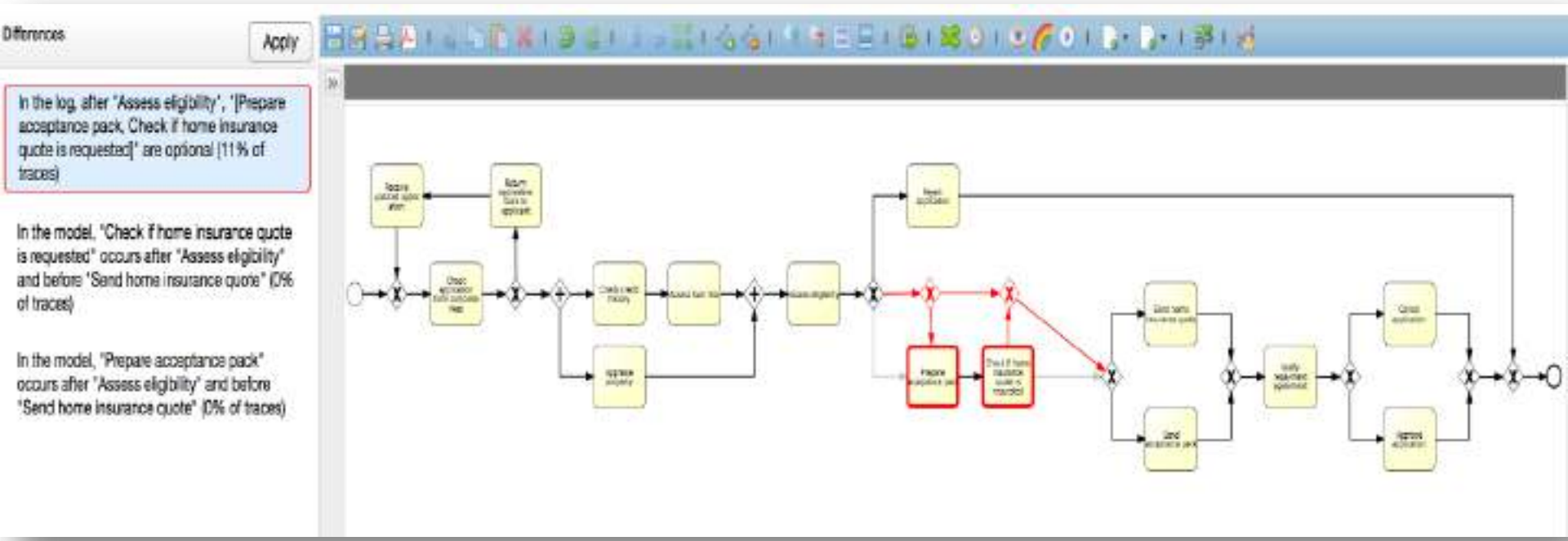


Conformance Checking

No. of Instances	Log Traces
1207	ABDEA
145	ACDGHFA
56	ACGDHFA
23	ACHDFA
28	ACDHFA

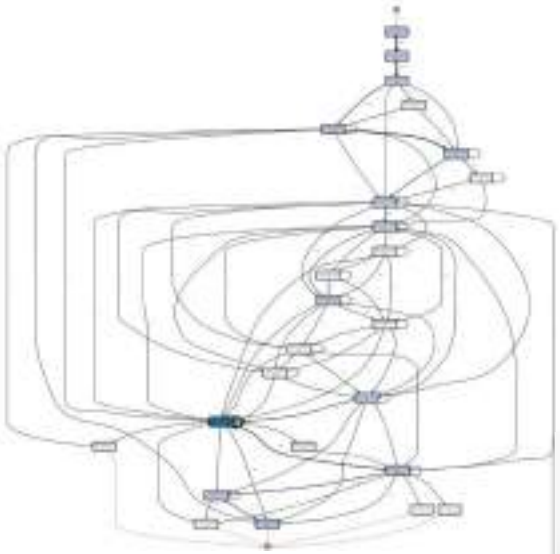


Conformance Checking in Action



Full demo at:
https://www.youtube.com/watch?v=3d00pORc9X8

Deviance & Variance Mining



L1 - Short stay

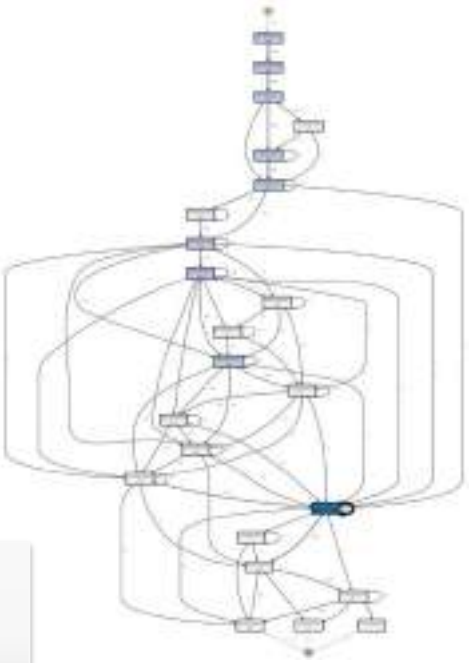
448 cases

7329 events

Delta analysis

In L1, “Nursing Primary Assessment” is repeated after “Medical Assign” and “Triage Request”, while in L2 it is not

…



L2 - Long stay

363 cases

7496 events

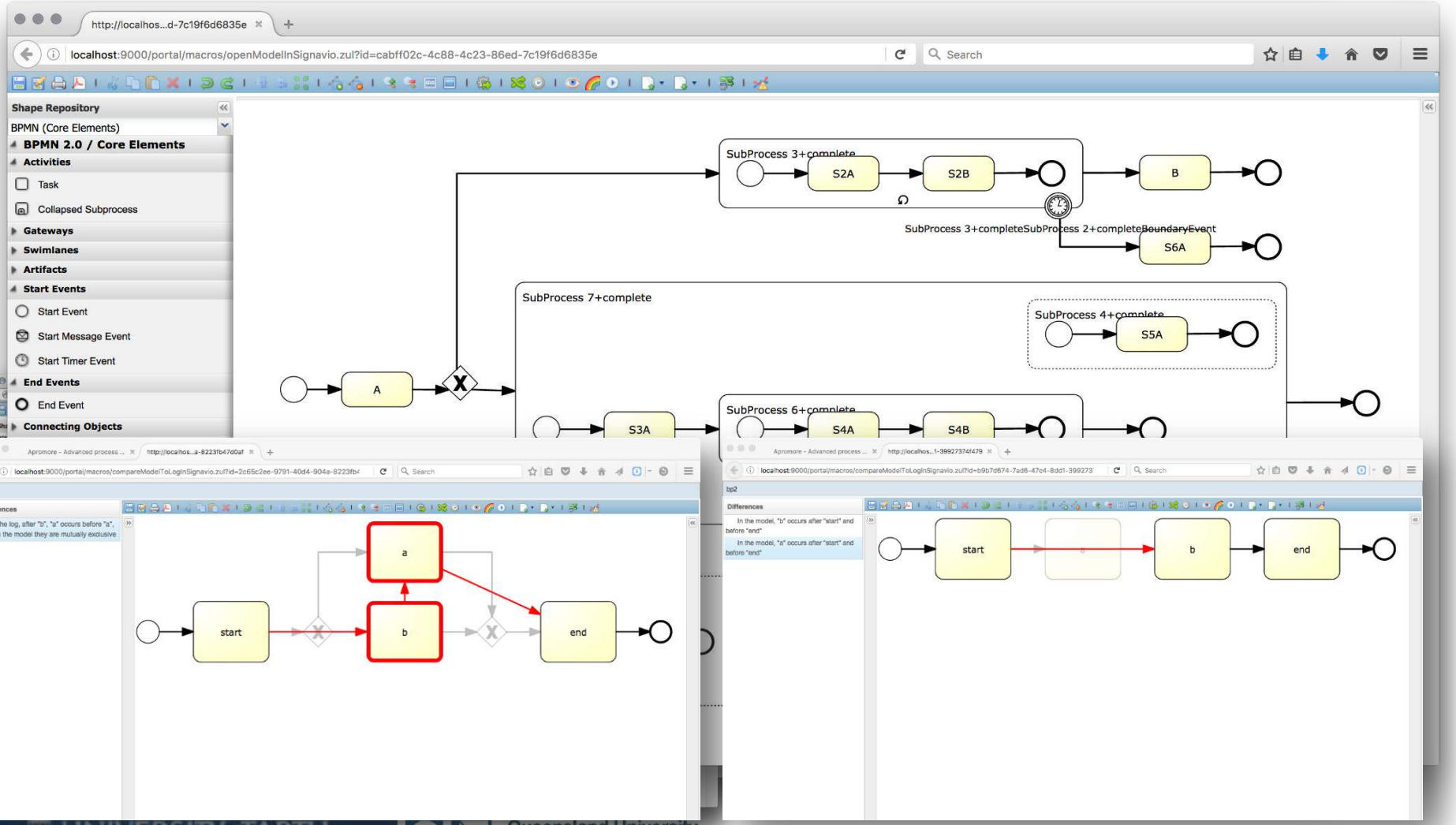
Queensland University of Technology

N.R. van Beest, L. Garcia-Banuelos, M. Dumas, M. La Rosa, Log Delta Analysis: Interpretable Differencing of Business Process Event Logs. BPM 2015: 386-405

Process Mining Tools

Open-source	Lightweight	Mid-range	Heavyweight
• **Apromore** • ProM	• Disco	• Minit • myInvenio • QPR Process Analyzer • Signavio Process Intelligence • StereoLOGIC Discovery Analyst	• ARIS Process Performance Manager • Celonis Process Mining • Perceptive Process Mining (Lexmark) • Interstage Process Discovery (Fujitsu)

Apromore.org



Process Mining: Where is it used?

- Insurance
  - **Suncorp, Australia**
- Government
  - Qld Treasury & Trade, Australia
- Health
  - AMC Hospital, The Netherlands
  - São Sebastião Hospital, Portugal
  - Chania Hospital, Greece
  - EHR Workflow Inc., USA
- Transport
  - ANA Airports, Portugal
  - Busan Port, South Korea
  - Kuehne + Nagel, Switzerland-Germany
- Electronics
  - Phillips, The Netherlands
- Banking, construction… etc.

Case Study: Suncorp Group

- General & life insurance, banking, superannuation and investments management
- 9M customers
- 16K employees
- $85 billion in assets

Suncorp Insurance

One Company Many Brands — SUNCORP GROUP

End to end insurance process

500 tasks

Source: Guidewire reference models

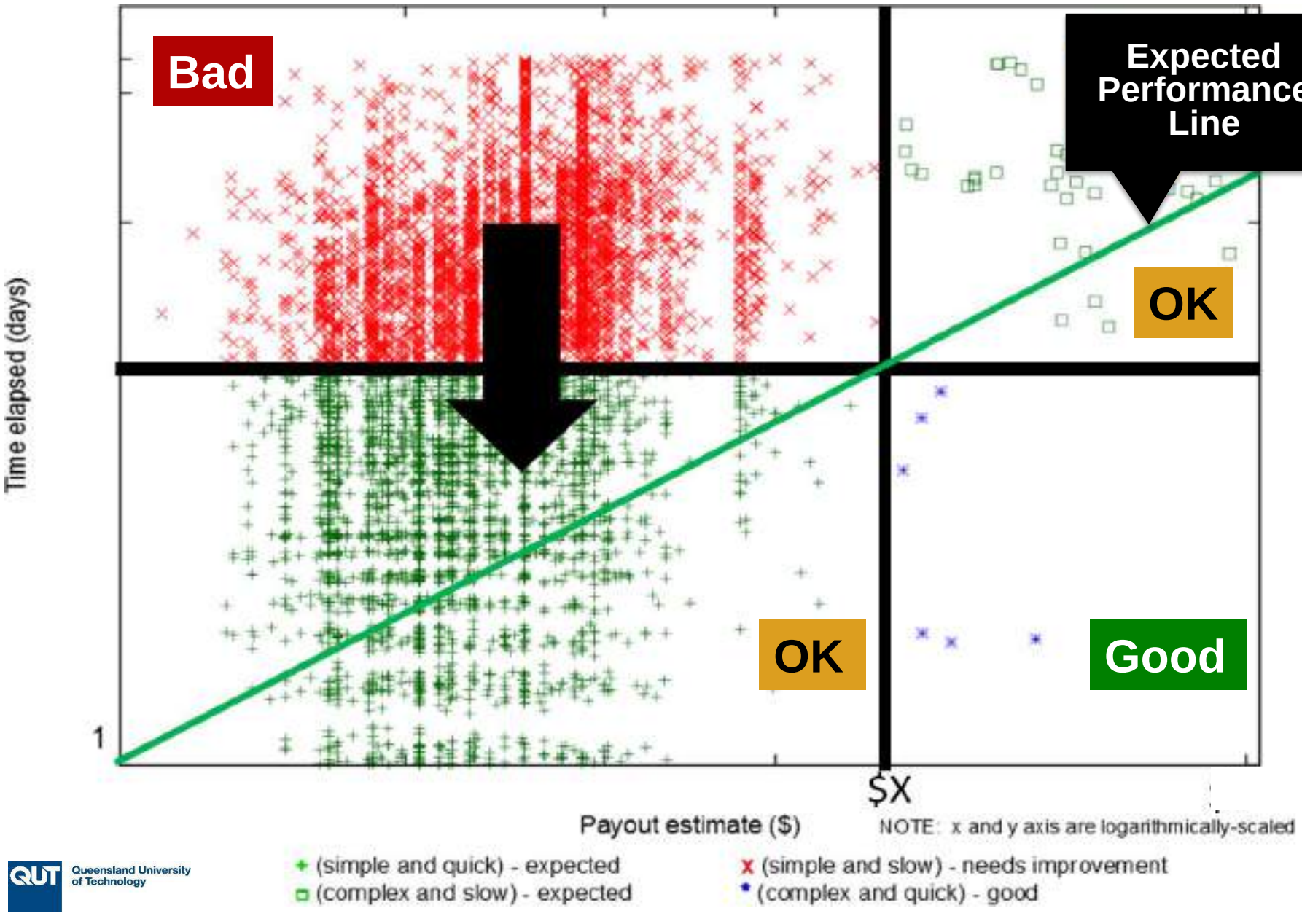
Each process is varied by product & brand…

	AAMI	Apia	bingle	CIL	GIO	JUST·CAR		SUNCORP	vero
Home	☑	☑			☑		☑	☑	☑
Motor	☑	☑	☑		☑	☑	☑	☑	☑
Commercial	☑			☑	☑			☑	☑
Liability	☑	☑			☑			☑	☑
CTP / WC	☑	☑			☑ ☑		☑	☑	

30 variations

Total process variants: **3,000+**

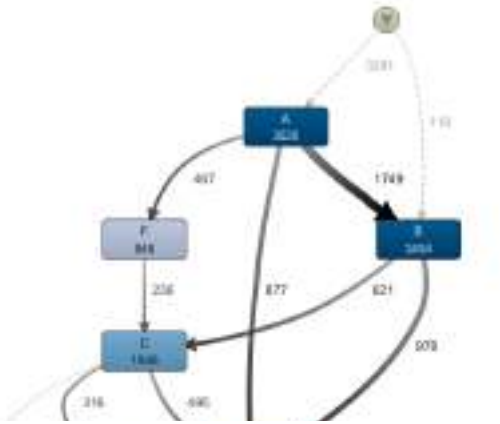
Processing Problem



Deviance & Variance Mining

Discover and analyse *actual* organisational processes from data

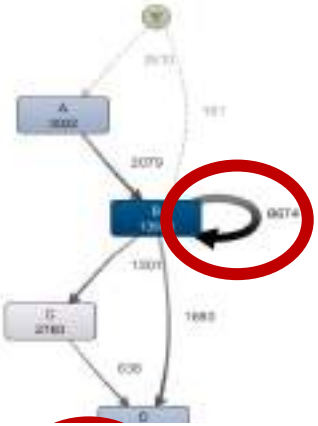
Simple Claim and Quick



MODEL

DISCRIMINATION

Simple Claim and Slow

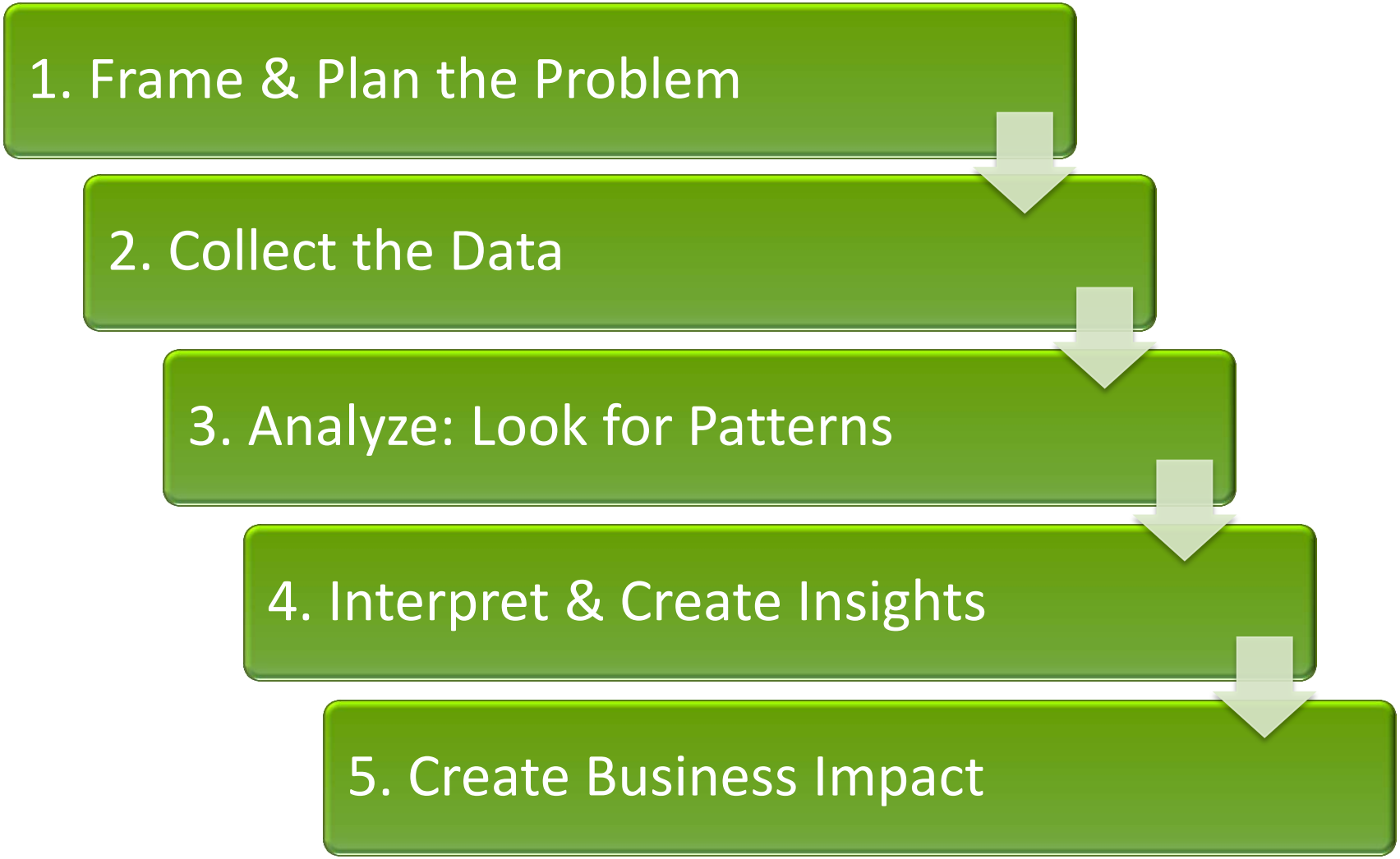


Suncorp is banking on data mining to slash its insurance claims processing times as the company scours for new avenues to boost its business.

Claims channelled through Suncorp's commercial insurance arm typically took 30 to 60 days to process, but a project with Queensland University of Technology has drastically reduced the window to between one and five days.

It comes after Suncorp's business cover division found "low value" claims, such as glass repairs or stolen laptops took much longer than they should to finalise.

Process Mining Methodology



Wil van der Aalst, 2012

1. Plan & Frame the Problem

- Frame a top-level question or phenomenon:
  - How and why does customer experiences with our order-to-cash processes diverge (geographically, product-wise, temporally)?
  - Why does the process perform poorly (bottlenecks, slow handovers)?
  - Why do we have frequent defects or performance deviance?
- Refine problem into:
  - Sub-questions
  - Identify success criteria and metrics
- Identify needed resources, get buy-in, plan remaining phases

1. Plan & Frame the Problem – Suncorp

- Often “simple” claims take an unexpectedly long time to complete:
  - What distinguishes the processing of simple claims completed on-time, and simple claims not completed on time?
  - What *early predictors* can be used to determine that a given “simple” claim will not be completed on time?
- Define what a “simple” claim is
- Create awareness of the extent of the problem

Resources:

- 2 part-time Business Analysts, 1 DB Administrator, 1 Executive Manager (sponsor)
- 1 full-time data scientist

Timeframe: 4 months

2. Collect the data

- Find relevant data sources
  - Information systems, SAP, Oracle, BPM Systems…
  - Identify process-related entities and their identifiers and map entities to relevant processes in the process architecture
- Extract traces
  - Collect records associated with process entities
  - Group records by process identifier to produce “traces”
  - Export traces into standard format (XES or MXML)
- Clean
  - Filter irrelevant events
  - Combine equivalent events
  - Filter out traces of infrequent variants if not relevant

2. Collect the data: minimum requirements

case id	event id	properties				
		timestamp	activity	resource	cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	reject request	Pete	200	...
2	35654483	30-12-2010:11.32	register request	Mike	50	...
	35654485	30-12-2010:12.12	check ticket	Mike	100	...
	35654487	30-12-2010:14.16	examine casually	Pete	400	...
	35654488	05-01-2011:11.22	decide	Sara	200	...
	35654489	08-01-2011:12.05	pay compensation	Ellen	200	...

3. Analyze – look for patterns

- Discover the **real** process from the logs
- Calculate process metrics
  - Cycle times, waiting times, error rates…
- Explore frequent paths
- Discover types of cases (good vs bad)
- Identify *process deviances* and *early predictors*
- …

Beyond Deviance Mining: Predictive Process Monitoring

How likely is it that a running process will become "deviant"?

- Will it end up in a negative outcome?
- Will it fail to meet its SLAs in the next 24 hours?
- Will it generate abnormal effort, costs or rework?

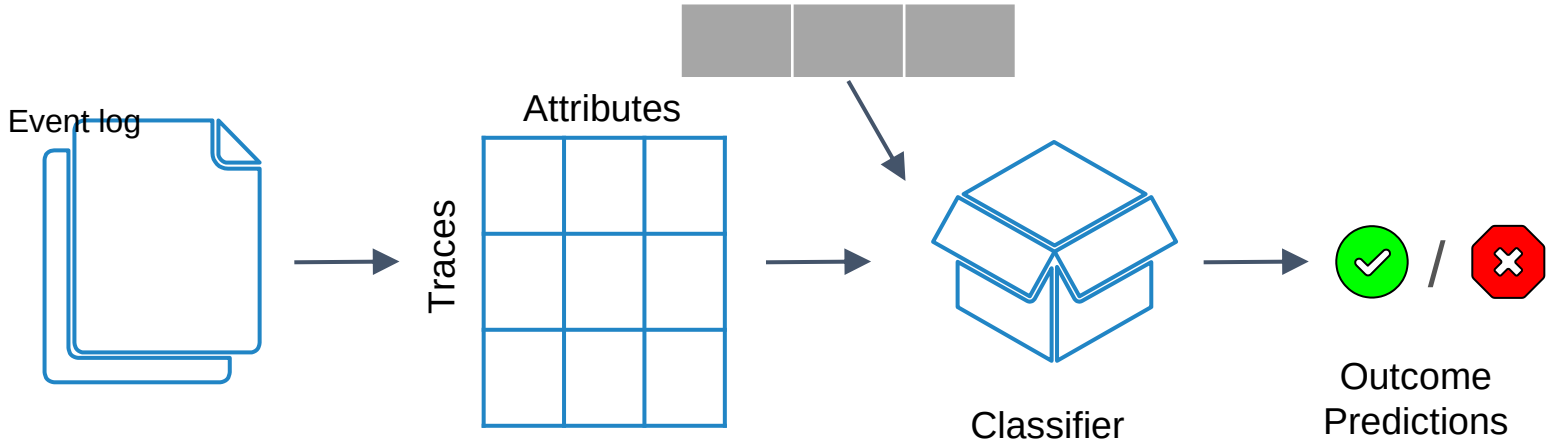
Predictive Process Monitoring – Detailed View

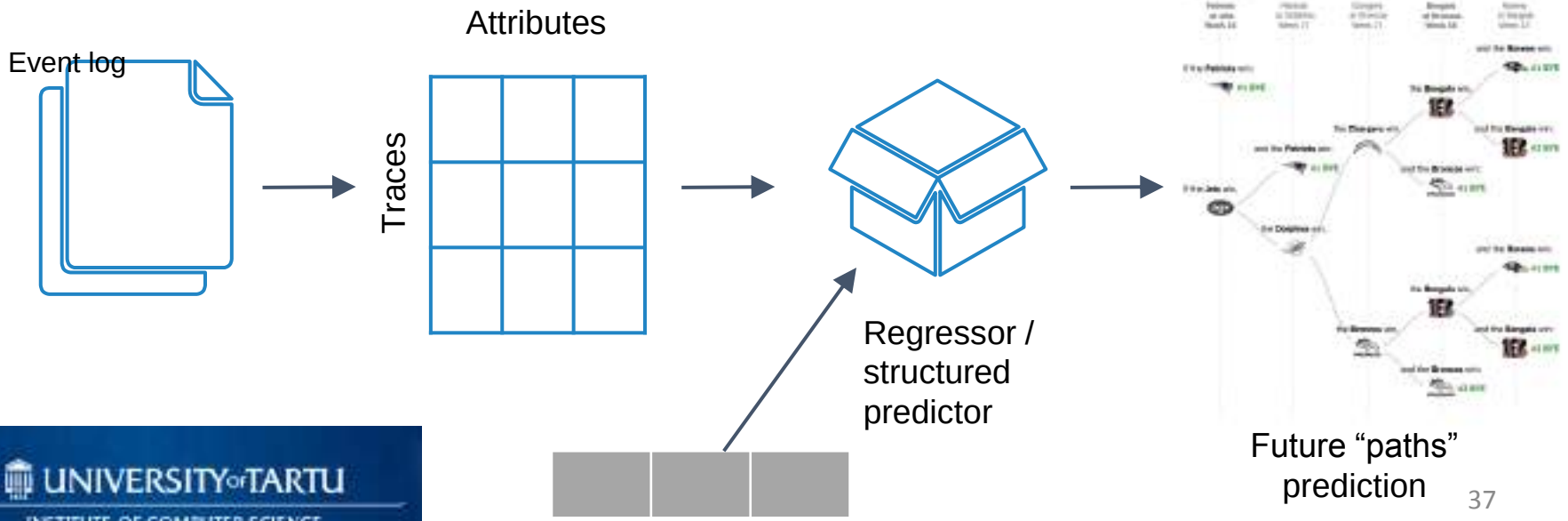
Case id	Event id	Properties				
		Timestamp	Activity	Resource	Cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	reject request	Pete	200	...
2	35654483	30-12-2010:11.32	register request	Mike	50	...
	35654485	30-12-2010:12.12	check ticket	Mike	100	...
	35654487	30-12-2010:14.16	examine casually	Pete	400	...
	35654488	05-01-2011:11.22	decide	Sara	200	...
	35654489	08-01-2011:12.05	pay compensation	Ellen	200	...
3	35654521	30-12-2010:14.32	register request	Pete	50	...
	35654522	30-12-2010:15.06	examine casually	Mike	400	...
	35654524	30-12-2010:16.34	check ticket	Ellen	100	...
	35654525	06-01-2011:09.18	decide	Sara	200	...
	35654526	06-01-2011:12.18	reinitiate request	Sara	200	...
	35654527	06-01-2011:13.06	examine thoroughly	Sean	400	...
	35654530	08-01-2011:11.43	check ticket	Pete	100	...
	35654531	09-01-2011:09.55	decide	Sara	200	...
	35654533	15-01-2011:10.45	pay compensation	Ellen	200	...

- What is the next activity for this case?
- When is this next activity going to take place?
- How long is this case still going to take until it is finished?
- What is the outcome of this case? Is the compensation going to be paid? Or rejected?

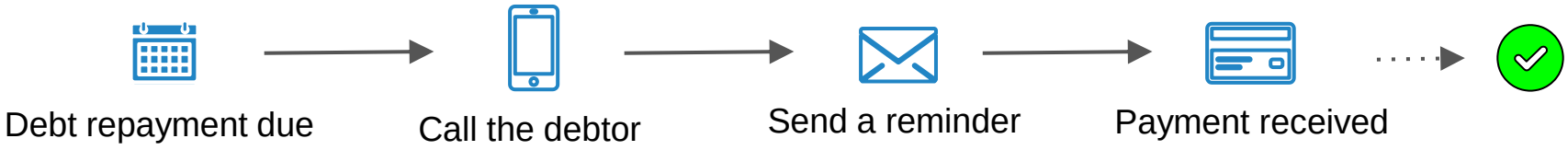
Current situation

Predictive Process Monitoring: General Approach

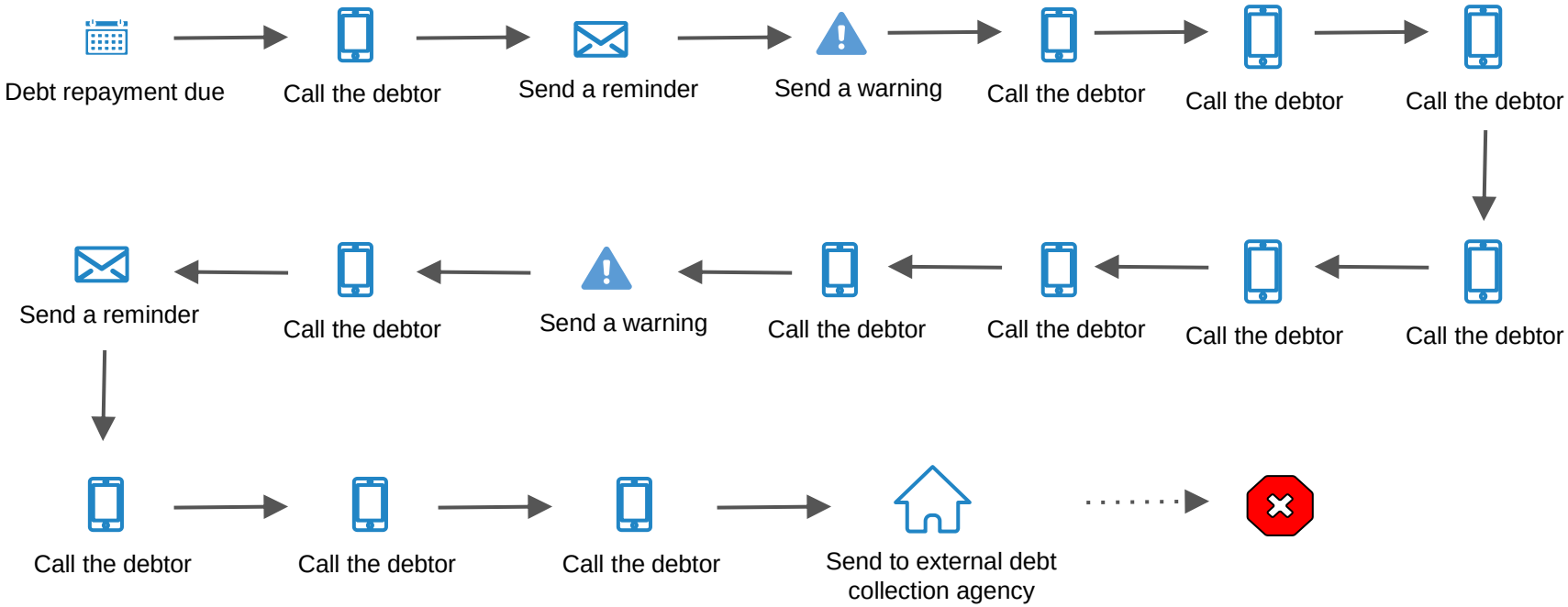




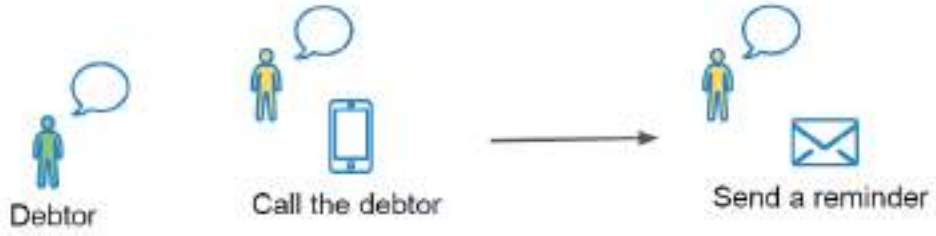
Predictive Monitoring Example: Debt Recovery Process



Predictive Monitoring Example: Debt Recovery Process

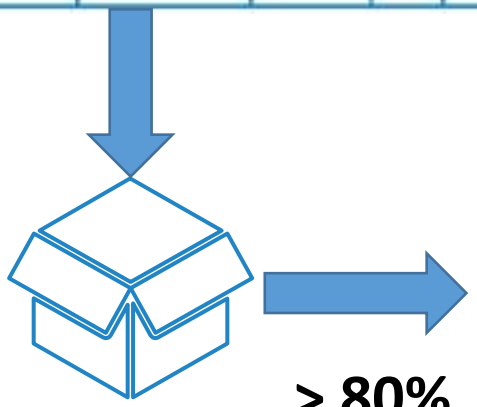


Predictive Process Monitoring for Debt Collection



Events **Case attributes** Encoding of textual data

	Event1	Event2	Resource1	Resource2	Debtor	t_{11}	...	t_{1n}	t_{21}	...	t_{2n}
Trace1	Call the debtor	Send a reminder	Sue	Bob	Mark	0.2	...	0.1	0.4	...	0.4



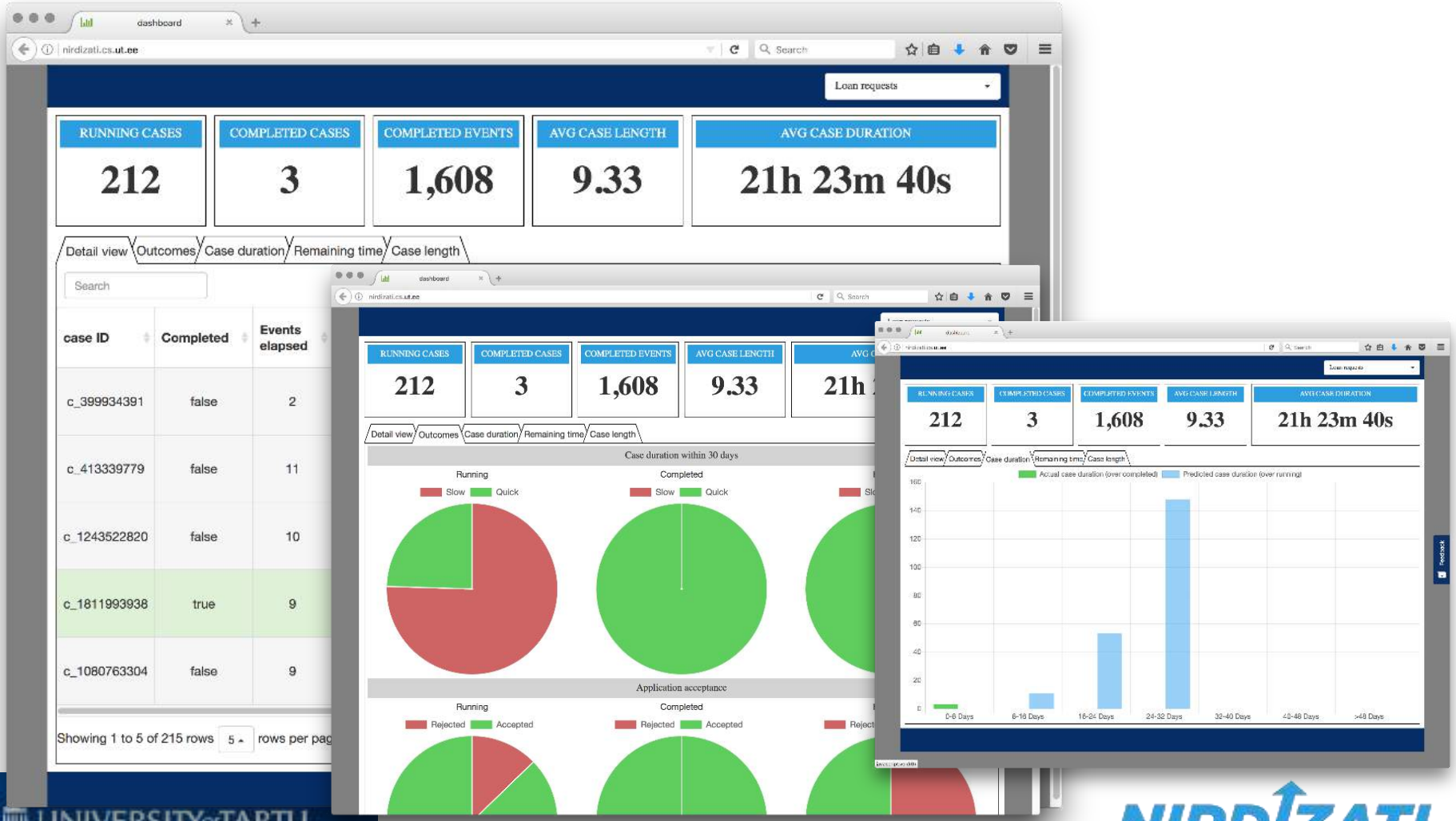
Classifier

> 80% accuracy

Will repay in 60 days or not?

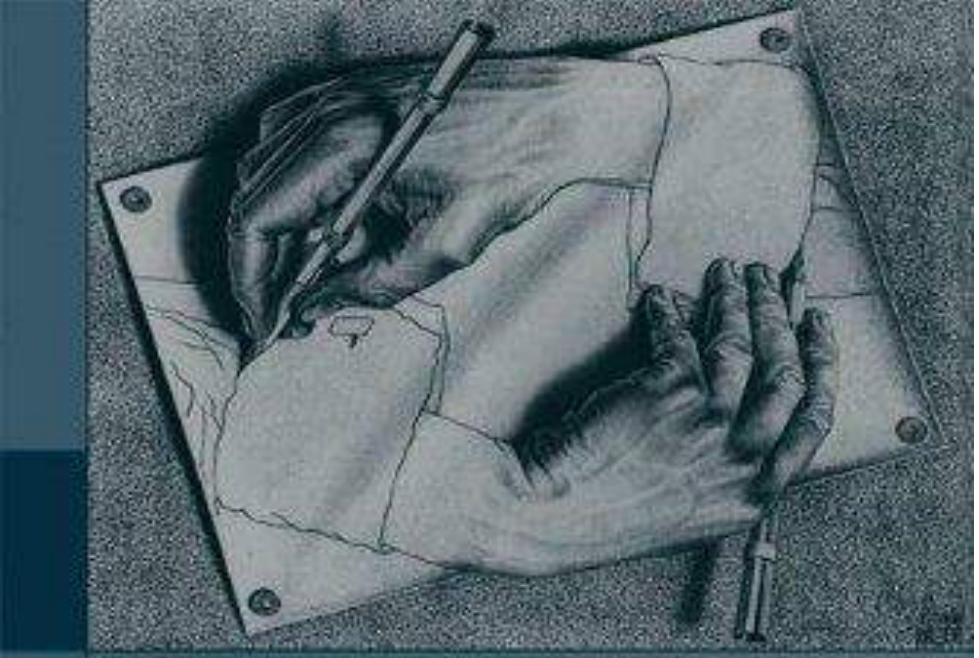
Nirdizati.com

Open-Source Predictive Process Monitoring



UNIVERSITY OF TARTU
INSTITUTE OF COMPUTER SCIENCE

NIRDIZATI



Fundamentals of

Business Process Management

Marlon Dumas
Marcello La Rosa
Jan Mendling
Hajo A. Reijers

Springer

Wil M. P. van der Aalst

Process Mining

Discovery, Conformance and Enhancement of Business Processes

Springer