Do you want to do a process mining project?

What are the requirements?
How do I start?

prof.dr.ir. Wil van der Aalst
RWTH Aachen University
W: vdaalst.com T:@wvdaalst

Many people and organizations contact me to apply process mining. They all have data and processes. However, often the prerequisites of process mining are unclear.

On the one hand, process mining is super generic and can be applied in any domain, just like spreadsheets are used in any organization. Spreadsheets can do anything with numbers. Process mining can do anything with events.

On the other hand, event data are not just any type of data and the notion of process is very broad.

These slides aim to clarify this. You need to check:
1. Do my events have a case id, activity name, and timestamp?
2. Can I sketch the expected process model in terms of the activities in the event log?
   (The process will be very different, but you should have some expectations, otherwise it is pointless to talks about processes.)
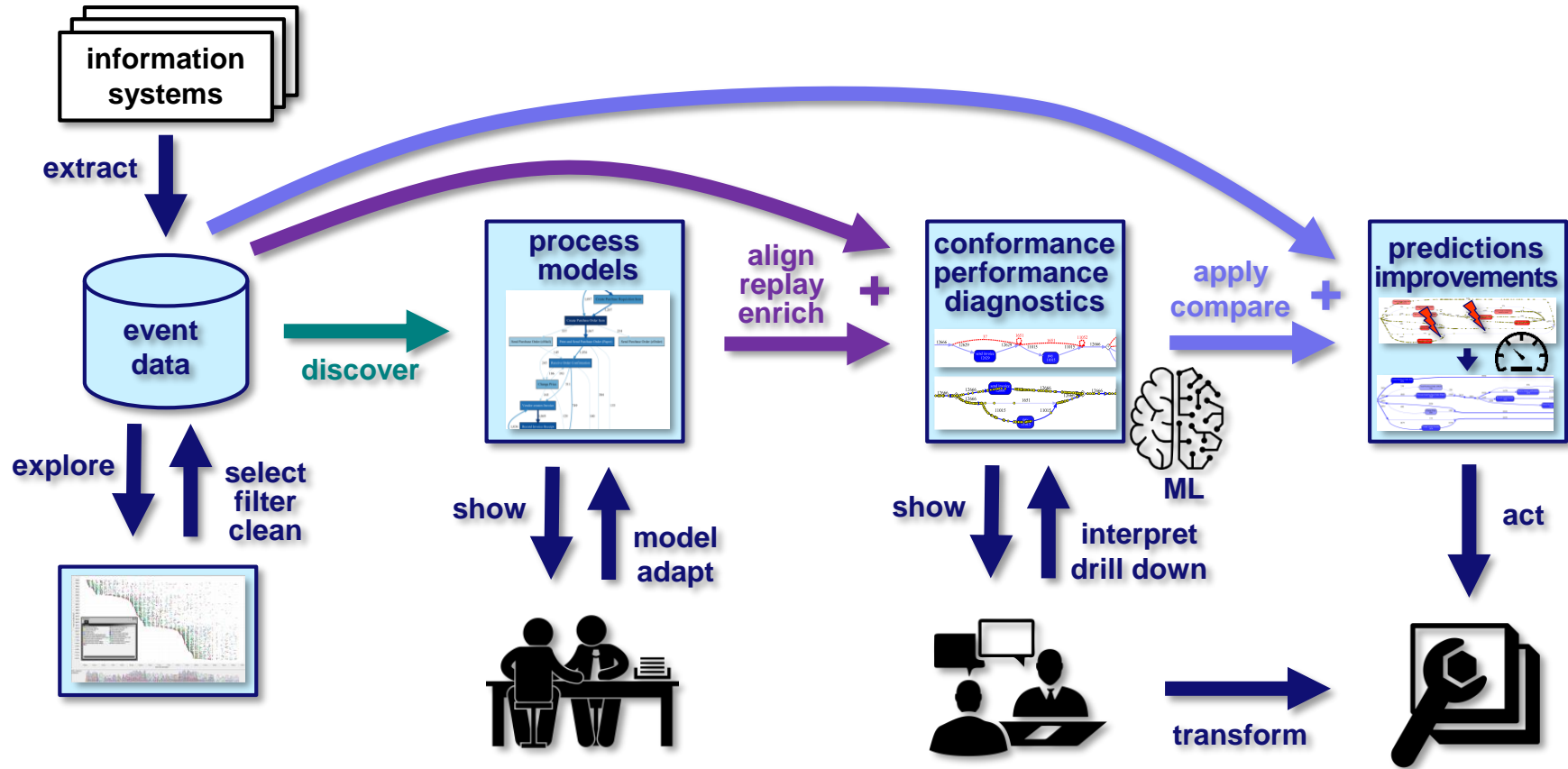
Wil van der Aalst

PADS Chair of Process and Data Science

RWTHAACHEN UNIVERSITY

Did your read the previous slide?

Please do if you want to work with us. If you cannot or do not want to answer these questions, then do <u>not</u> ask for a meeting to collaborate!

PADS | Chair of Process and Data Science

RWTH AACHEN UNIVERSITY

What is it?

"event data are everywhere"

PADS
Chair of Process
and Data Science

RWTH AACHEN
UNIVERSITY

information systems

extract

event data

explore

select filter clean

discover

process models

align replay enrich

+

conformance performance diagnostics

apply compare

+

predictions improvements

show

model adapt

show

interpret drill down

ML

act

transform

PADS

Chair of Process and Data Science

Starting point: Event data

Case ID	Activity	Resource	Timestamp	product	prod-price	quantity	address
...	...	....	...	....	...	...	...
6350	place order	Aiden	2018/02/13 14:29:45.000	APPLE iPhone 6 16 GB	639,00 €	5	NL-7751DG-21
6283	pay	Lily	2018/02/13 14:39:25.000	SAMSUNG Galaxy S6 32 GB	543.99	3	NL-7828AM-11a
6253	prepare delivery	Sophia	2018/02/13 15:01:33.000	APPLE iPhone 6 16 GB	639,00 €	3	NL-7887AC-13
6257	prepare delivery	Aiden	2018/02/13 15:03:43.000	SAMSUNG Galaxy S6 32 GB	543.99	1	NL-9521KJ-34
6185	confirm payment	Emily	2018/02/13 15:05:36.000	SAMSUNG Galaxy S4	329,00 €	1	NL-9521GC-32
6218	confirm payment	Emily	2018/02/13 15:08:11.000	APPLE iPhone 6s Plus 64 GB	969,00 €	2	NL-7948BX-10
6245	make delivery	Michael	2018/02/13 15:14:04.000	APPLE iPhone 6 16 GB	639,00 €	3	NL-7905AX-38
6272	pay	Emily	2018/02/13 15:20:36.000	APPLE iPhone 6 16 GB	639,00 €	1	NL-7821AC-3
6269	pay	Charlotte	2018/02/13 15:25:21.000	SAMSUNG Galaxy S4	329,00 €	1	NL-7907EJ-42
6212	prepare delivery	Sophia	2018/02/13 15:43:39.000	HUAWEI P8 Lite	234,00 €	1	NL-7905AX-38
6323	send invoice	Alexander	2018/02/13 15:46:08.000	APPLE iPhone 6 16 GB	639,00 €	1	NL-7833HT-15
6246	confirm payment	Jack	2018/02/13 15:56:03.000	SAMSUNG Galaxy S4	329,00 €	3	NL-7833HT-15
6347	send invoice	Jack	2018/02/13 15:57:42.000	SAMSUNG Galaxy S4	329,00 €	3	NL-7905AX-38
6351	place order	Zoe	2018/02/13 16:17:37.000	APPLE iPhone 5s 16 GB	449,00 €	3	NL-9521GC-32
6204	prepare delivery	Sophia	2018/02/13 16:31:28.000	SAMSUNG Core Prime G361	135,00 €	1	NL-7828AM-11a
6204	make delivery	Kaylee	2018/02/13 16:51:54.000	SAMSUNG Core Prime G361	135,00 €	1	NL-7828AM-11a
6265	confirm payment	Lily	2018/02/13 16:55:55.000	SAMSUNG Galaxy S4	329,00 €	4	NL-9521GC-32
6250	confirm payment	Jack	2018/02/13 17:03:26.000	MOTOROLA Moto G	199,00 €	4	NL-7942GT-2
6328	send invoice	Lily	2018/02/13 17:30:16.000	APPLE iPhone 6s 64 GB	858,00 €	4	NL-9514BV-16
6352	place order	Aiden	2018/02/13 17:53:22.000	APPLE iPhone 6 16 GB	639,00 €	2	NL-9514BV-16
6317	send invoice	Jack	2018/02/13 18:45:30.000	APPLE iPhone 6s 64 GB	858,00 €	5	NL-7907EJ-42
6353	place order	Sophia	2018/02/13 20:16:20.000	APPLE iPhone 5s 16 GB	449,00 €	4	NL-7751AR-19
...	...	....	...	...	...	...	...

event

**71,043 events
12,666 cases
7 activities**

Starting point: Event data

Case ID	Activity	Resource	Timestamp	product	prod-price	quantity	address
...	...	....	...	....	...	...	...
6350	place order	Aiden	2018/02/13 14:29:45.000	APPLE iPhone 6 16 GB	639,00 €	5	NL-7751DG-21
6283	pay	Lily	2018/02/13 14:39:25.000	SAMSUNG Galaxy S6 32 GB	543.99	3	NL-7828AM-11a
6253	prepare delivery	Sophia	2018/02/13 15:01:33.000	APPLE iPhone 6	639.00 €	3	NL-7887AC-13
6257	prepare delivery	Aiden	2018/02/13 15:03:43.000	SAMSUNG Galaxy S6 32 GB	543.99	1	NL-9521KJ-34
6185	confirm payment	Emily	2018/02/13 15:05:36.000	SAMSUNG Galaxy S4	329,00 €	1	NL-9521GC-32
6218	confirm payment	Emily	2018/02/13 15:08:11.000	APPLE iPhone 6s Plus 64 GB	969,00 €	2	NL-7948BX-10
6245	make delivery	Michael	2018/02/13 15:14:04.000	APPLE iPhone	49,00 €	3	NL-7905AX-38
6272	pay	Emily	2018/02/13 15:20:36.000	APPLE iPhone 6	639,00 €	1	NL-7821AC-3
6269	pay	Charlotte	2018/02/13 15:25:21.000	SAMSUNG Galaxy S4	329,00 €	1	NL-7907EJ-42
6212	prepare delivery	Sophia	2018/02/13 15:43:39.000	HUAWEI P8 Lite	234,00 €	1	NL-7905AX-38
6323	send invoice	Alexander	2018/02/13 15:46:08.000	APPLE iPhone		1	NL-7833HT-15
6246	confirm payment	Jack	2018/02/13 15:56:03.000	SA	99 €	4	NL-7833HT-15
6347	send invoice	Jack	2018/02/13 15:57:42.000	SAMSUNG Galaxy S4	329,00 €	3	NL-7905AX-38
6351	place order	Zoe	2018/02/13 16:17:37.000	APPLE iPhone 5s 16 GB	449,00 €	3	NL-9521GC-32
6204	prepare delivery	Sophia	2018/02/13 16:31:28.000	SAMS			NL-7828AM-11a
6204	make delivery	Kaylee	2018/02/13 16:51:54.000	SAMS			28AM-11a
6265	confirm payment	Lily	2018/02/13 16:55:55.000	SAMSUNG Galaxy S4	329,00 €	4	NL-9521GC-32
6250	confirm payment	Jack	2018/02/13 17:03:26.000	MOTOROLA Moto G	199,00 €	4	NL-7942GT-2
6328	send invoice	Lily	2018/02/13 17:30:16.000	APPLE iPhone 6s 64 GB	858,00 €	4	NL-9514BV-16
6352	place order	Aiden	2018/02/13 17:53:22.000	APPLE iPhone 6 16 GB	639,00 €	2	NL-9514BV-16
6317	send invoice	Jack	2018/02/13 18:45:30.000	APPLE iPhone 6s 64 GB	858,00 €	5	NL-7907EJ-42
6353	place order	Sophia	2018/02/13 20:16:20.000	APPLE iPhone 5s 16 GB	449,00 €	4	NL-7751AR-19
...	...	....	...	...	...	...	...

**event =
case +
activity +
timestamp +
...**

Chair of Process and Data Science

Let's look at orders 6350, 6351, and 6352

Case ID	Activity	Timestamp
6350	place order	2018/02/13 14:29:45.000
6351	place order	2018/02/13 16:17:37.000
6352	place order	2018/02/13 17:53:22.000
6352	send invoice	2018/02/19 09:20:28.000
6351	send invoice	2018/02/19 16:08:07.000
6350	send invoice	2018/02/21 09:38:16.000
6350	pay	2018/03/02 12:39:37.000
6352	pay	2018/03/05 15:46:47.000
6351	cancel order	2018/03/06 10:17:01.000
6350	prepare delivery	2018/03/07 13:50:35.000
6350	make delivery	2018/03/07 16:41:01.000
6350	confirm payment	2018/03/07 16:53:00.000
6352	prepare delivery	2018/03/07 17:05:59.000
6352	confirm payment	2018/03/07 17:59:55.000
6352	make delivery	2018/03/08 09:54:36.000

Chair of Process
and Data Science

Let's look at orders 6350, 6351, and 6352

Case ID	Activity	Timestamp
6350	place order	2018/02/13 14:29:45.000
6351	place order	2018/02/13 16:17:37.000
6352	place order	2018/02/13 17:53:22.000
6352	send invoice	2018/02/19 09:20:28.000
6351	send invoice	2018/02/19 16:08:07.000
6350	send invoice	2018/02/21 09:38:16.000
6350	pay	2018/03/02 12:39:37.000
6352	pay	2018/03/05 15:46:47.000
6351	cancel order	2018/03/06 10:17:01.000
6350	prepare delivery	2018/03/07 13:50:35.000
6350	make delivery	2018/03/07 16:41:01.000
6350	confirm payment	2018/03/07 16:53:00.000
6352	prepare delivery	2018/03/07 17:05:59.000
6352	confirm payment	2018/03/07 17:59:55.000
6352	make delivery	2018/03/08 09:54:36.000

order 6350

place order

send invoice
pay

prepare delivery
make delivery
confirm payment

order 6351

place order

send invoice

cancel order

order 6352

place order
send invoice

pay

prepare delivery
confirm payment
make delivery

Using the whole event log

Case ID	Activity	Timestamp
6350	place order	2018/02/13 14:29:45.000
6351	place order	2018/02/13 16:17:37.000
6352	place order	2018/02/13 17:53:22.000
6352	send invoice	2018/02/19 09:20:28.000
6351	send invoice	2018/02/19 16:08:07.000
6350	send invoice	2018/02/21 09:38:16.000
6350	pay	2018/03/02 12:39:37.000
6352	pay	2018/03/05 15:46:47.000
6351	cancel order	2018/03/06 10:17:01.000
6350	prepare delivery	2018/03/07 13:50:35.000
6350	make delivery	2018/03/07 16:41:01.000
6350	confirm payment	2018/03/07 16:53:00.000
6352	prepare delivery	2018/03/07 17:05:59.000
6352	confirm payment	2018/03/07 17:59:55.000
6352	make delivery	2018/03/08 09:54:36.000

place order
send invoice
pay
prepare delivery
make delivery
confirm payment

8016 x

place order
send invoice
cancel order

1651 x

place order
send invoice
pay
prepare delivery
confirm payment
make delivery

2962 x

place order
pay
send invoice
prepare delivery
make delivery
confirm payment

30 x

place order
pay
send invoice
prepare delivery
confirm payment
make delivery

7 x

Using the whole event log
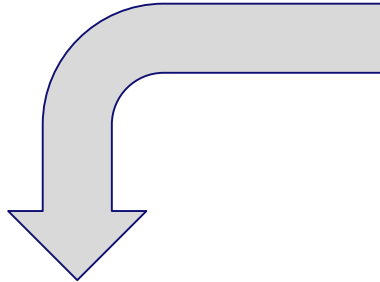
place order
send invoice
pay
prepare delivery
make delivery
confirm payment

8016 x

place order
send invoice
cancel order

1651 x

place order
send invoice
pay
prepare delivery
confirm payment
make delivery

2962 x

place order
pay
send invoice
prepare delivery
make delivery
confirm payment

30 x

place order
pay
send invoice
prepare delivery
confirm payment
make delivery
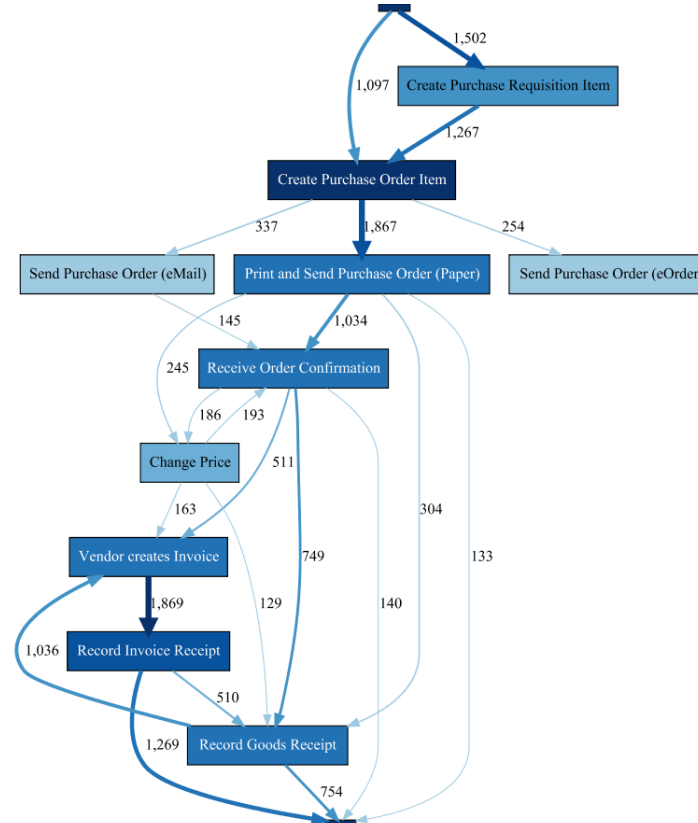
7 x
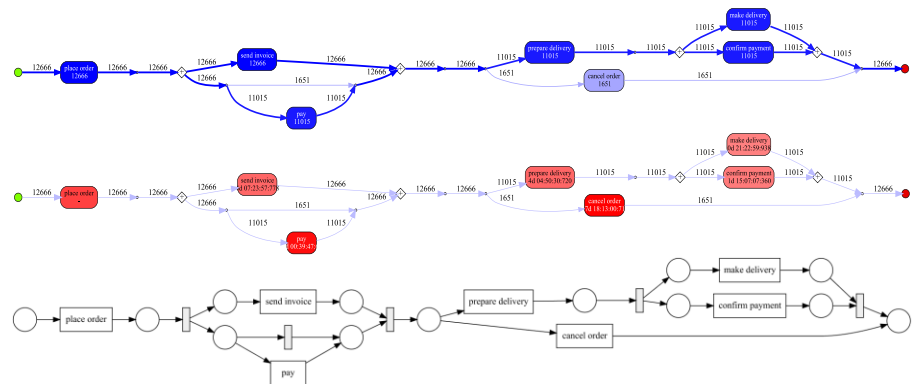
Performance and Compliance
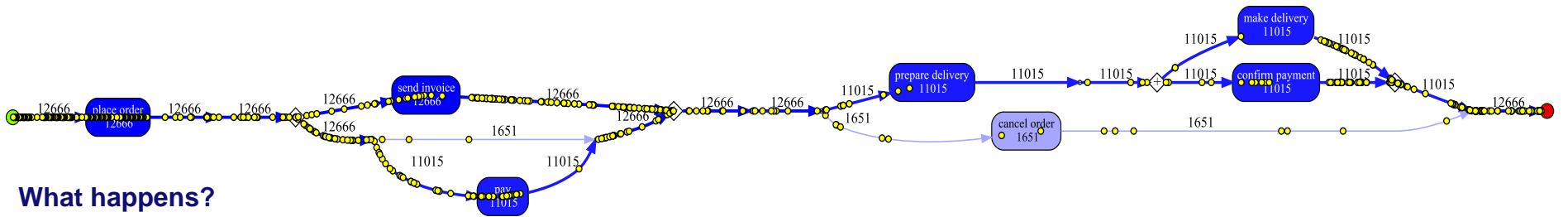
What happens?
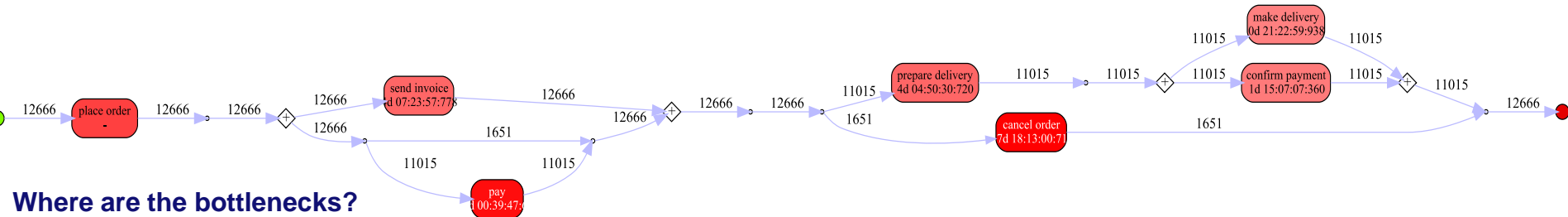
Where are the bottlenecks?
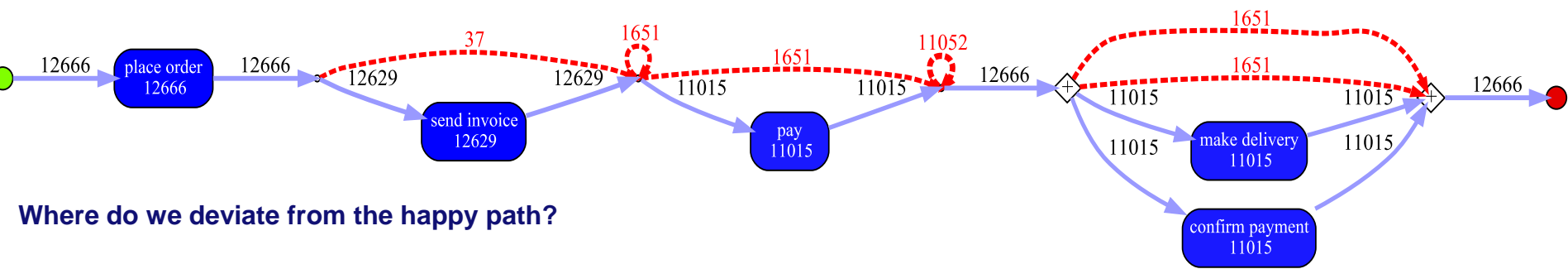
Where do we deviate from the happy path?

Reality is not so simple
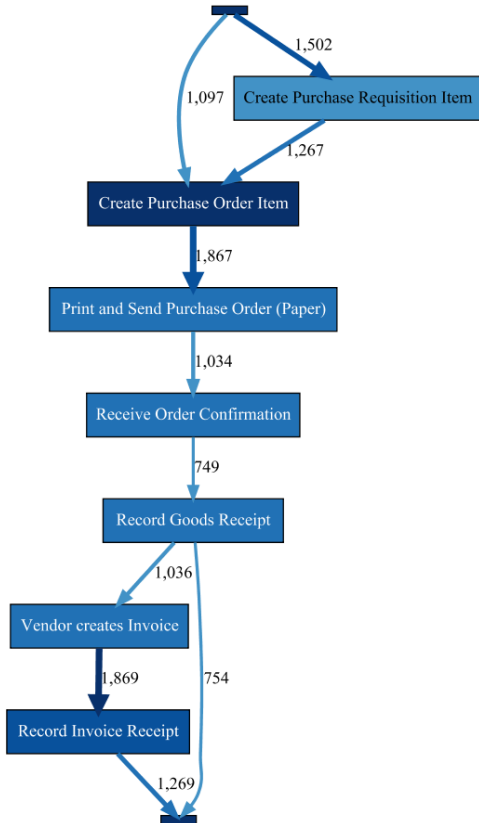
Reality is not so simple

It is common to find thousands of different variants for simple core processes like P2P and O2C!

Caused by hand-offs, rework, duplication, ineffective communication, etc.

Process mining helps organizations to address these problems (to actually realize the economies of scale promised)

Reveal performance and conformance issues and suggest actions.
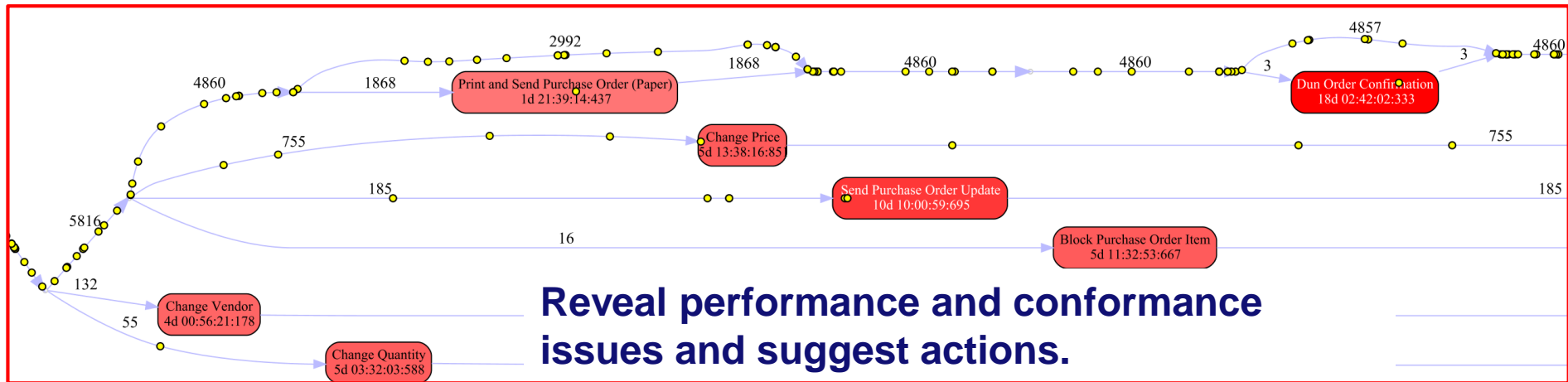
More on event data

"are your data really event data"

PADS
Chair of Process
and Data Science

RWTH AACHEN
UNIVERSITY

Event log

- **We assume the existence of an event log where each event refers to a case, an activity, and a point in time.**

- **An event log can be seen as a collection of cases.**

- **A case can be seen as a trace/sequence of events.**

Event data may come from …

- a database system (e.g., patient data in a hospital),
- a comma-separated values (CSV) file or spreadsheet,
- a transaction log (e.g., a trading system),
- a business suite/ERP system (SAP, Oracle, etc.),
- a message log (e.g., from IBM middleware),
- an open API providing data from websites or social media, …

An example log

student name	course name	exam date	mark
Peter Jones	Business Information systems	16-1-2014	8
Sandy Scott	Business Information systems	16-1-2014	5
Bridget White	Business Information systems	16-1-2014	9
John Anderson	Business Information systems	16-1-2014	8
Sandy Scott	BPM Systems	17-1-2014	7
Bridget White	BPM Systems	17-1-2014	8
Sandy Scott	Process Mining	20-1-2014	5
Bridget White	Process Mining	20-1-2014	9
John Anderson	Process Mining	20-1-2014	8
…	…	…	…

case id   **activity name**   **timestamp**   **other data**

Chair of Process and Data Science

Another event log: order handling

order number	activity	timestamp	user	product	quantity
9901	register order	22-1-2014@09.15	Sara Jones	iPhone5S	1
9902	register order	22-1-2014@09.18	Sara Jones	iPhone5S	2
9903	register order	22-1-2014@09.27	Sara Jones	iPhone4S	1
9901	check stock	22-1-2014@09.49	Pete Scott	iPhone5S	1
9901	ship order	22-1-2014@10.11	Sue Fox	iPhone5S	1
9903	check stock	22-1-2014@10.34	Pete Scott	iPhone4S	1
9901	handle payment	22-1-2014@10.41	Carol Hope	iPhone5S	1
9902	check stock	22-1-2014@10.57	Pete Scott	iPhone5S	2
9902	cancel order	22-1-2014@11.08	Carol Hope	iPhone5S	2
…	…	…	…	…	…

case id **activity name** **timestamp** **resource** **other data**

Another event log: patient treatment

patient	activity	timestamp	doctor	age	cost
5781	make X-ray	23-1-2014@10.30	Dr. Jones	45	70.00
5541	blood test	23-1-2014@10.18	Dr. Scott	61	40.00
5833	blood test	23-1-2014@10.27	Dr. Scott	24	40.00
5781	blood test	23-1-2014@10.49	Dr. Scott	45	40.00
5781	CT scan	23-1-2014@11.10	Dr. Fox	45	1200.00
5833	surgery	23-1-2014@12.34	Dr. Scott	24	2300.00
5781	handle payment	23-1-2014@12.41	Carol Hope	45	0.00
5541	radiation therapy	23-1-2014@13.57	Dr. Jones	61	140.00
5541	radiation therapy	23-1-2014@13.08	Dr. Jones	61	140.00
…	…	…	…	…	…

case id   **activity name**   **timestamp**   **resource**   **other data**

Chair of Process
and Data Science

Minimal requirements in terms of a CSV file

- **Each row corresponds to an event.**
- **There are at least three columns:**
  - **Case id (patient id, order number, claim number, …)**
  - **Activity name (approve, reject, request, send, …)**
  - **Timestamp (2015-08-18T06:36:40, …)**
- **There may be many other (optional) columns: resource, transaction type, age, costs, etc.**

order number	activity	timestamp	user	product	quantity
9901	register order	22-1-2014@09.15	Sara Jones	iPhone5S	1
9902	register order	22-1-2014@09.18	Sara Jones	iPhone5S	2

Extensions

- **Transactional information** on activity instances:
  An event can represent
  a **start**, **complete**,
  suspend, **resume**,
  abort, etc.
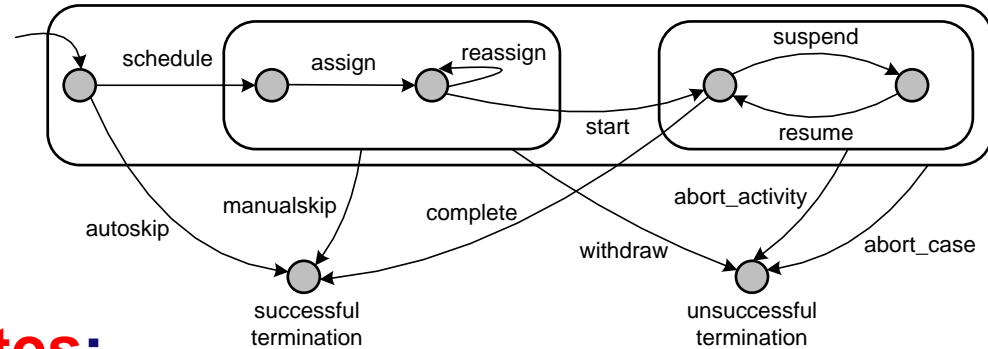
- **Case** versus **event** **attributes**:

  - case attributes do not change, e.g., the birth date or gender of a patient,

  - event attributes are related to a particular step in the process.

XES (eXtensible Event Stream)

- **Adopted by the IEEE Task Force on Process Mining.**
- **The format is supported by tools such as ProM and Disco (used in this course).**
- **Predecessors: MXML and SA-MXML.**
- **Conversion from other formats (CSV) is easy if the right data are available.**
- **XML syntax and OpenXES library available.**
- **See www.xes-standard.org.**

Simplistic view on event data

order number	activity	timestamp	user	product	quantity
9901	register order	22-1-2014@09.15	Sara Jones	iPhone5S	1
9902	register order	22-1-2014@09.18	Sara Jones	iPhone5S	2
9903	register order	22-1-2014@09.27	Sara Jones	iPhone4S	1
9901	check stock	22-1-2014@09.49	Pete Scott	iPhone5S	1
9901	ship order	22-1-2014@10.11	Sue Fox	iPhone5S	1
9903	check stock	22-1-2014@10.34	Pete Scott	iPhone4S	1
9901	handle payment	22-1-2014@10.41	Carol Hope	iPhone5S	1
9902	check stock	22-1-2014@10.57	Pete Scott	iPhone5S	2
9902	cancel order	22-1-2014@11.08	Carol Hope	iPhone5S	2
…	…	…	…	…	…

case id **activity name** **timestamp** **resource** **other data**

Chair of Process and Data Science

A more refined view

Transactional model for activities
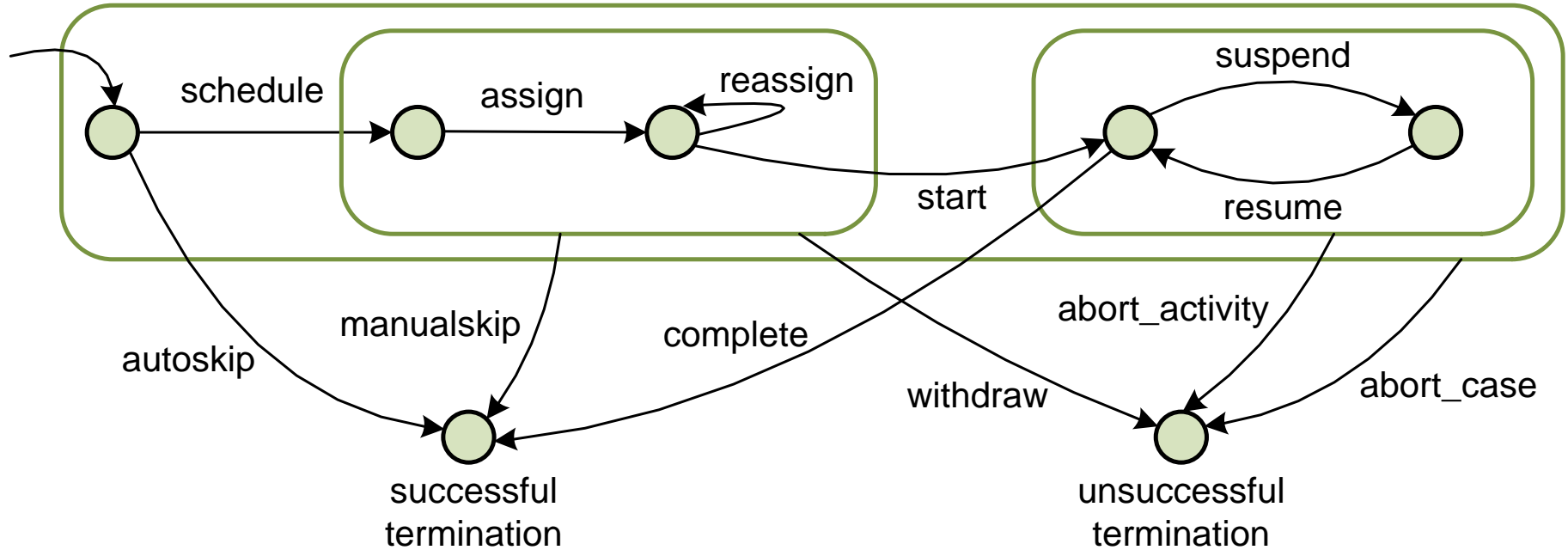
Atomic activities
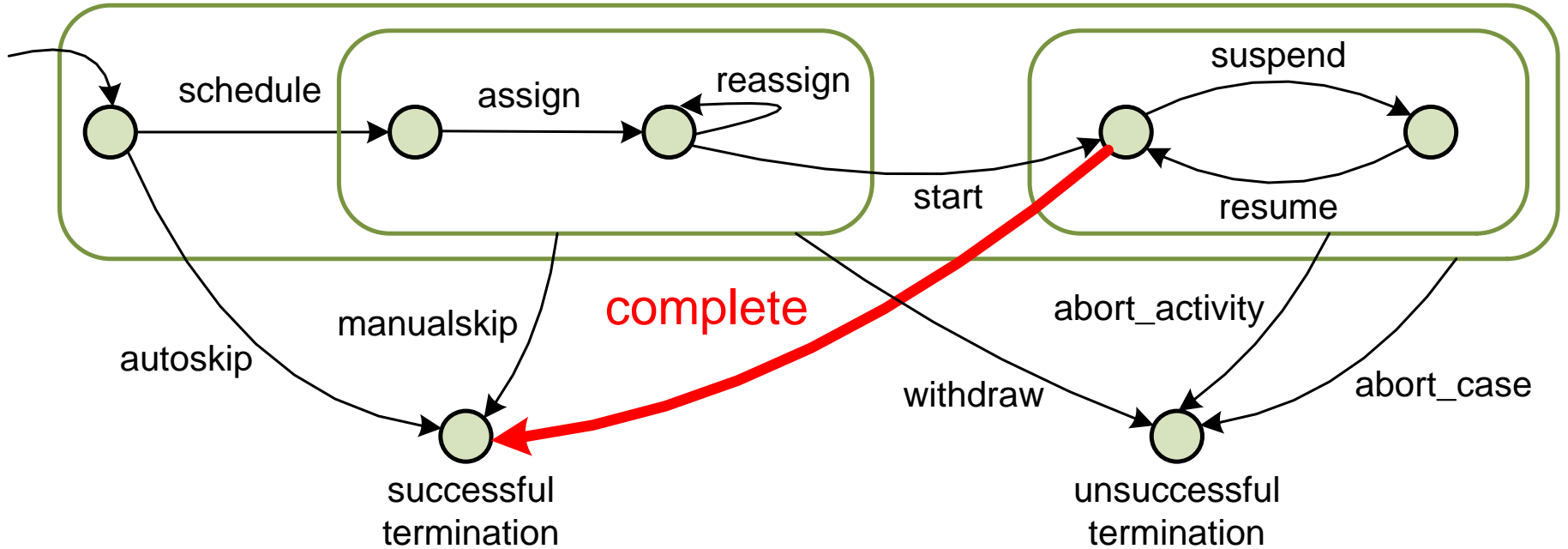
Activities that have a duration

Converting Event Data to XES

- **Easy with tools like ProM, Disco, etc.!**

Import CSV file, then apply "Convert CSV to XES" plug-in.
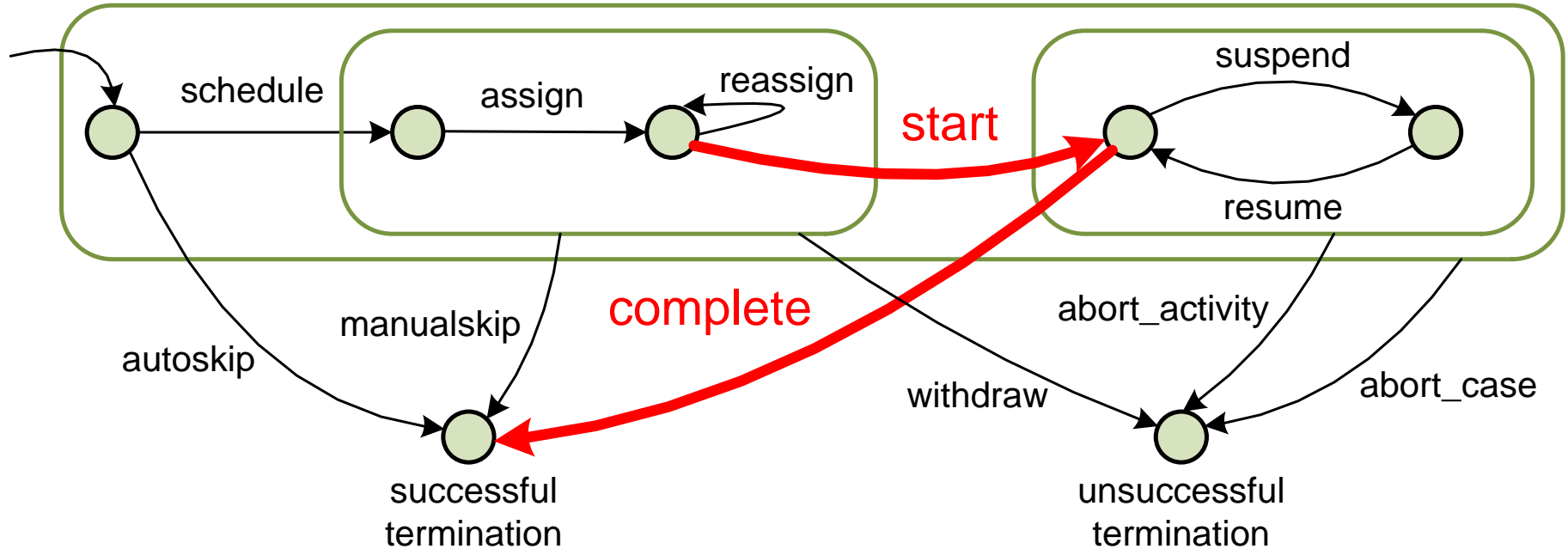
Now you should be able to check:
1. Do my events have indeed
   a) a case id,
   b) activity name, and
   c) timestamp?
2. Can I sketch the expected process model in terms of the activities in the event log?
   (The process will be very different, but you should have some expectations, otherwise it is pointless to talks about processes.)

Please bring fragments of event data and rough sketches of process models to the meeting. This will help to quickly see whether process mining will be feasible and beneficial.

The later slides provide some additional context.

Wil van der Aalst

P D
A S **Chair of Process
and Data Science**

RWTHAACHEN
UNIVERSITY

Do you want to discuss a joint project or seek advice?

Please ensure that you followed the suggestions on the previous slide. If you cannot or do not want to do this, then please do <u>not</u> contact us for free advice or support. This may seem rude, but we are spending a lot of time interacting with people that have lots of "data" and "processes" (everyone has), but do not have a clue what process mining is (despite the abundance of information).

P D
A S
Chair of Process
and Data Science

RWTH AACHEN
UNIVERSITY

A Bit of History

"bridging the gap between process science and data science"

"process management by modeling"

Petri nets
Formal methods
Concurrency theory
BPM, WFM, etc.
Simulation

< 1999

≥ 1999

Process mining
Predictive analytics
Process discovery
Conformance checking

"process management by mining"

Chair of Process and Data Science

research

commercial tools

adoption

ProM process mining workbench

ICPM 2019: First PM conference

30+ process mining companies

first process mining book

Celonis becomes a Unicorn

founding of Celonis

Market Guide for Process Mining by Gartner

first process mining company (Futura PI)

heuristic miner

organization mining

alignment-based conformance checking

"Process Mining: Data Science in Action" book

Alpha miner

decision mining

IEEE Task Force on Process Mining

start of process mining research at TU/e

token-based conformance checking

founding Fluxicon, ProcessGold, …

Coursera process mining MOOC

1997 1998 1999 **2000** 2001 2002 2003 2004 **2005** 2006 2007 2008 2009 **2010** 2011 2012 2013 2014 **2015** 2016 2017 2018 2019 **2020**

Open-Source Process Mining Software

(far from complete)

MiMo (2001-2002)

EMiT (2002-2004)

Thumb (2002-2004)

InWolvE (2000-2004)

ProM 1.1-5.2 (>286 plugins) (2004-2010)

ProM 6.X (>1500 plugins) (2010-now)

PM4Py (2018-now)

RapidProM (2014-now)

Apromore (2015*-now)

bupaR (2016-now)
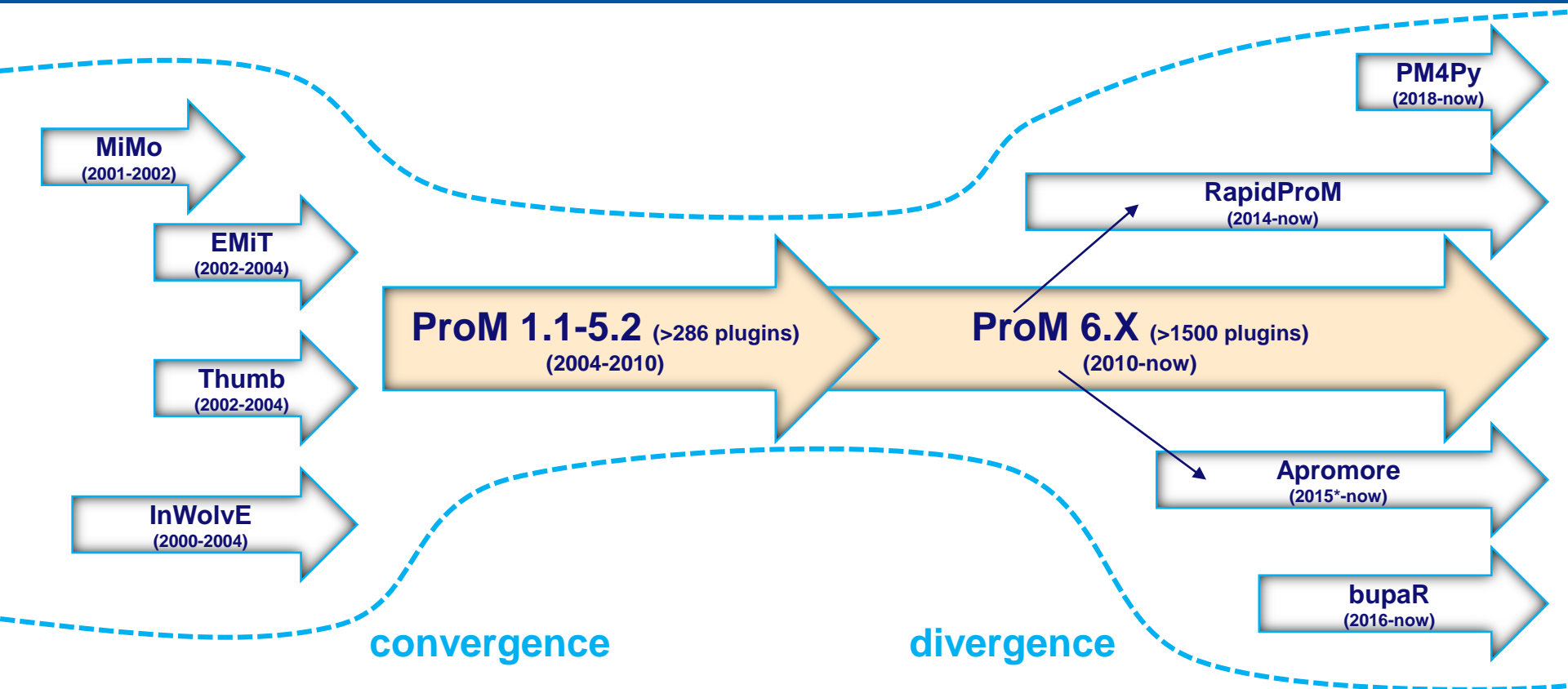
convergence

divergence

Over 35 process mining vendors today

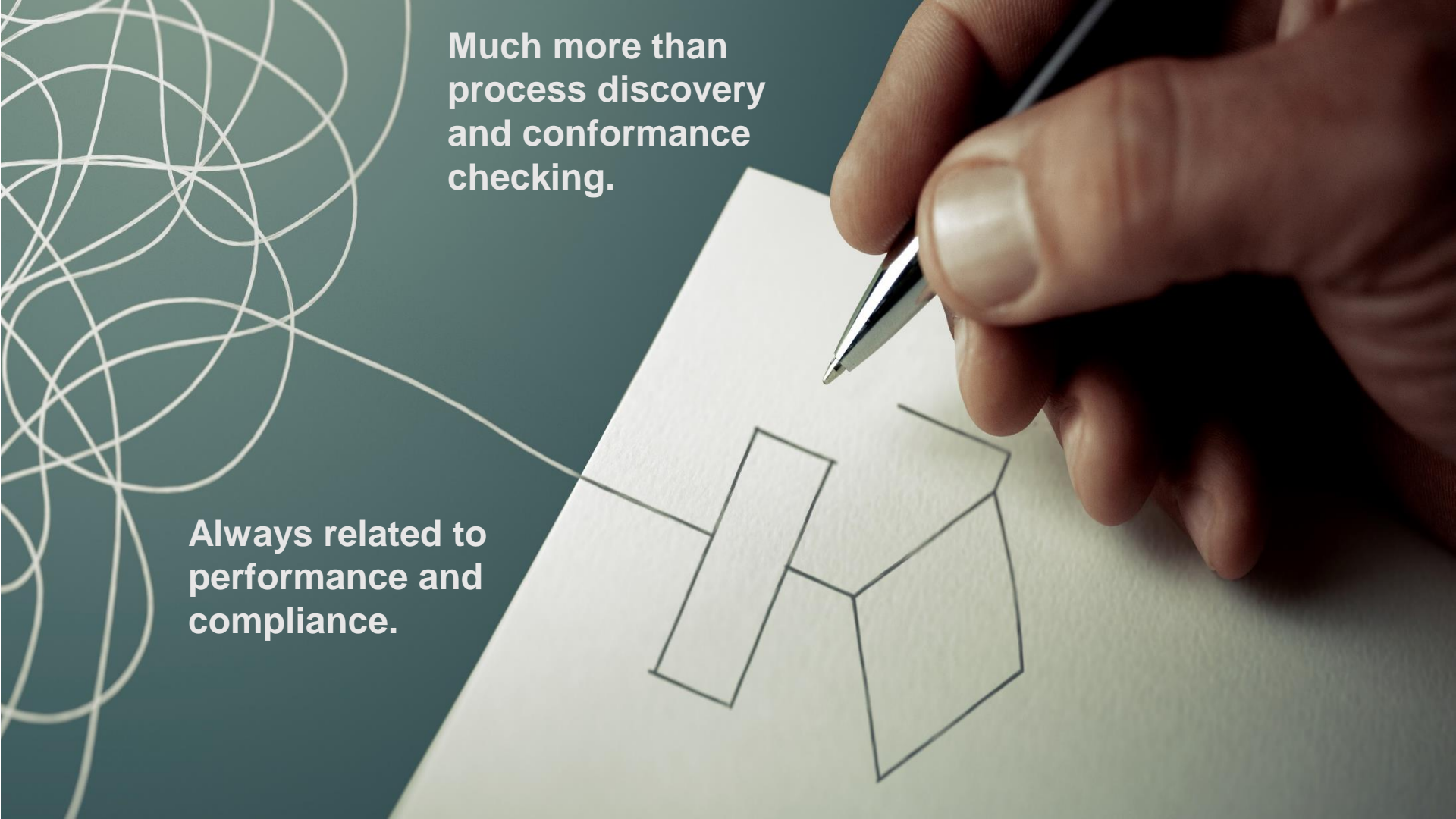# Many of the larger organizations in Europe are using process mining already (and we are just at the beginning!)

Celonis was the first to focus on continuous process mining
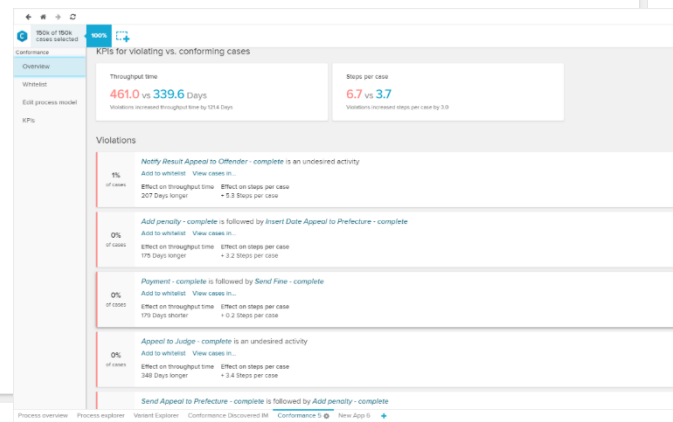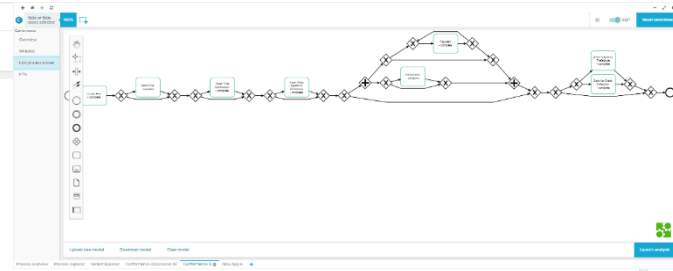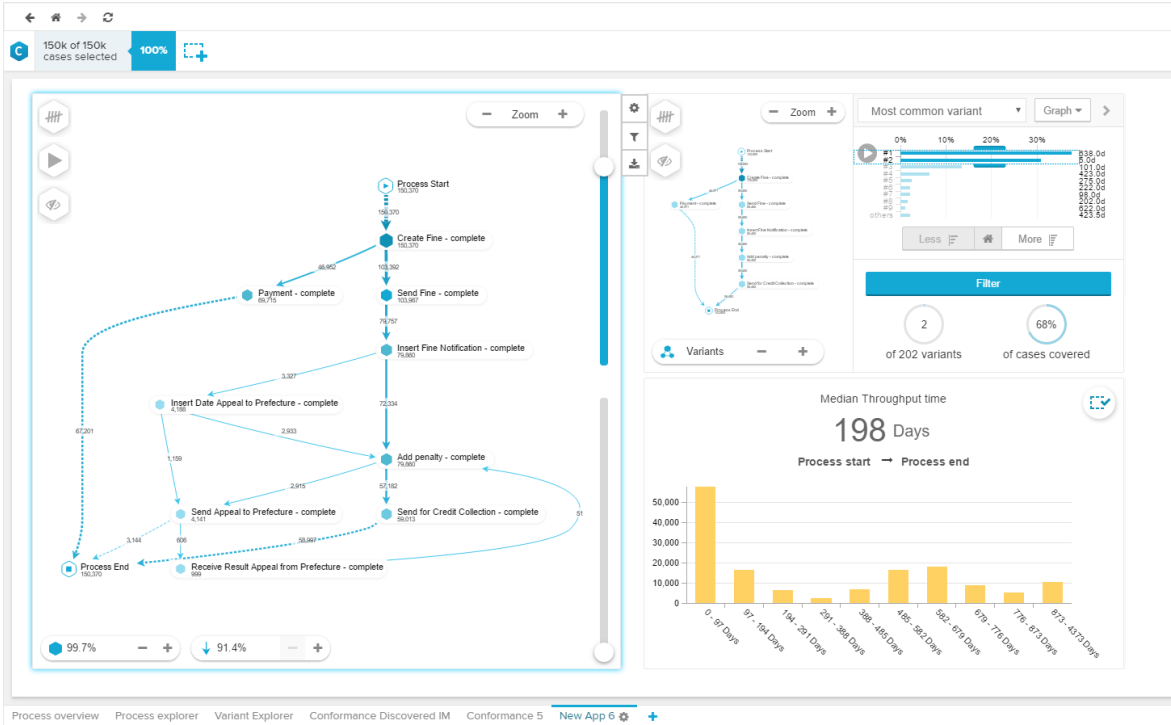
From data scientists to process managers and from insights to actions.

Relation to ML & AI

"Siri and Alexa cannot mine your processes"

PADS Chair of Process and Data Science

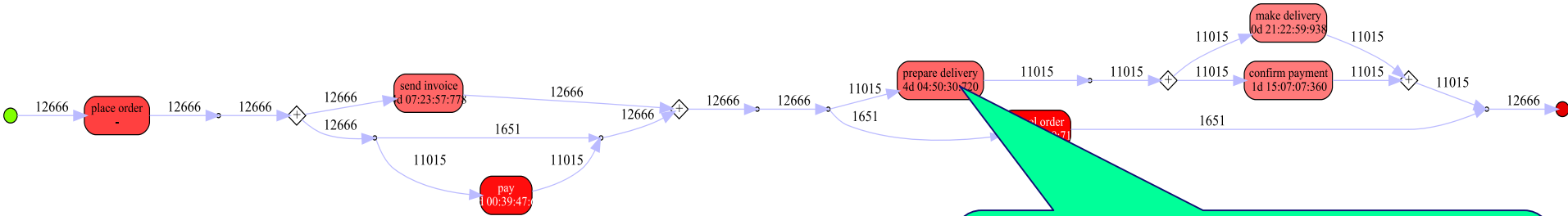# Process mining is very different!

The core process mining techniques and tools do <u>not</u> use techniques from machine learning, artificial intelligence, data mining, etc.
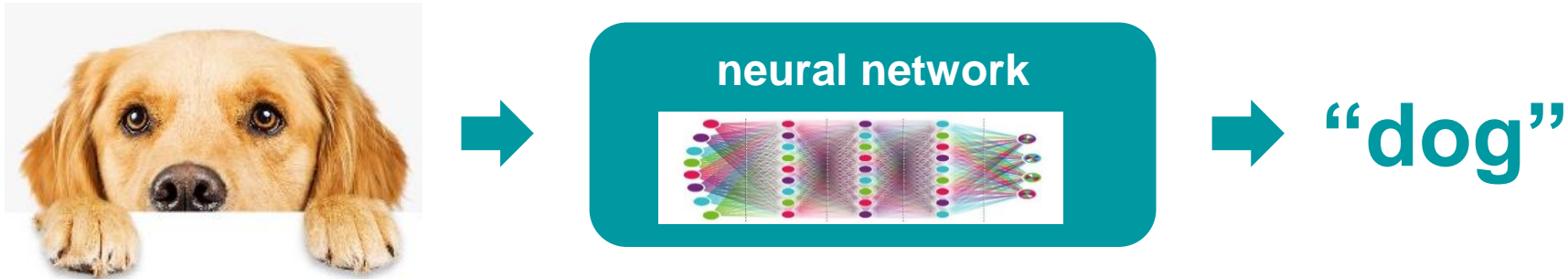
- The model needs to be visible and understandable by stakeholders.
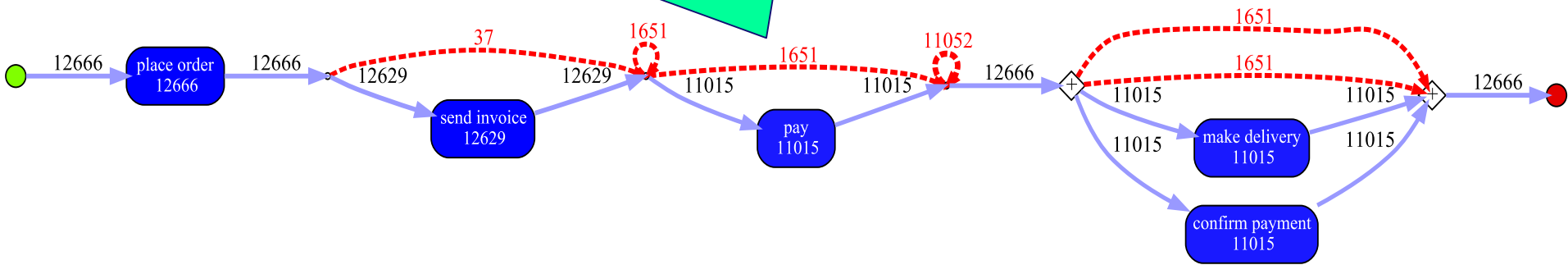- Process owners are not going to label training examples.

However, …
PM can be used to generate ML problems

Why are payments skipped?
What do these cases have in common?
Can we predict such deviations?

Why is the bottleneck here?
What is causing it?
Can we predict such delays?
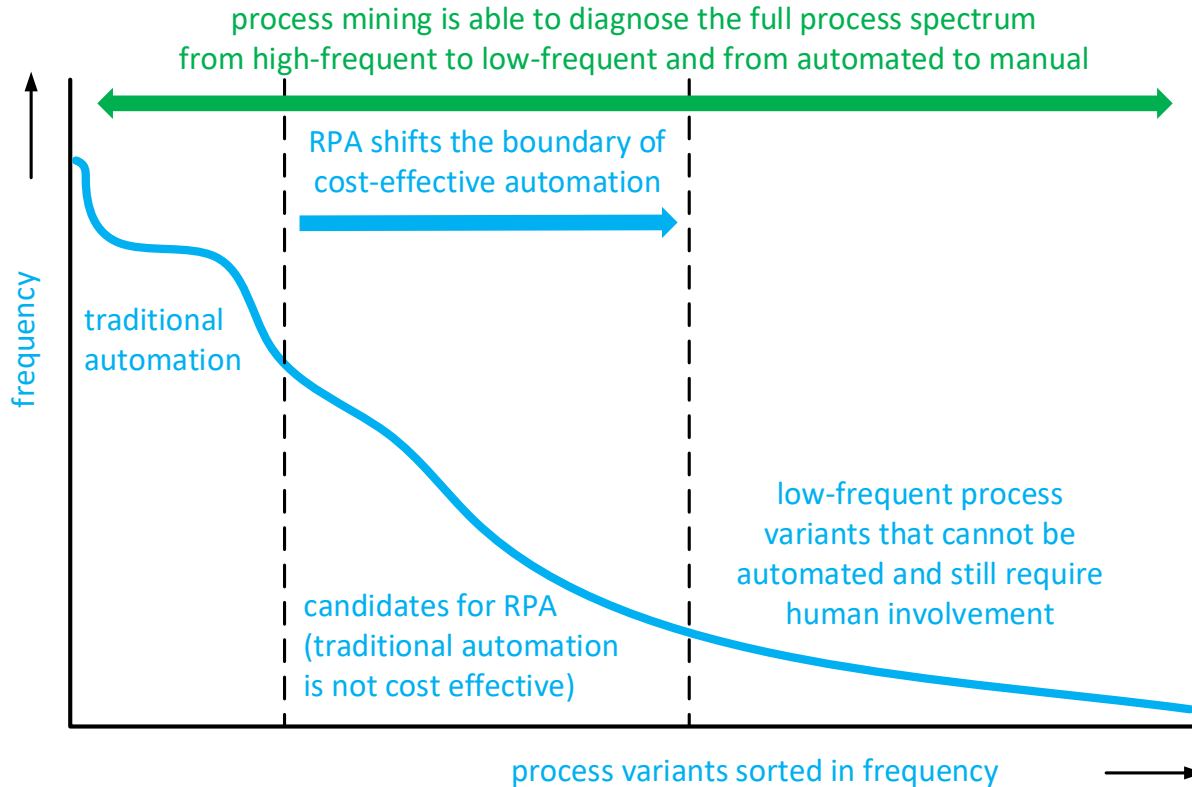
Relation to Robotic Process Automation (RPA)

"enabling the poor man's workflow management solution"

Chair of Process
and Data Science

How to pick your automation battles?

How to pick your automation battles? The RPA connection

process mining is able to diagnose the full process spectrum from high-frequent to low-frequent and from automated to manual

RPA shifts the boundary of cost-effective automation

frequency

traditional automation

low-frequent process variants that cannot be automated and still require human involvement

candidates for RPA (traditional automation is not cost effective)

process variants sorted in frequency

Chair of Process and Data Science

Object-Centric
Event Logs

"Everything should be made as simple as possible - but no simpler!"

PADS Chair of Process and Data Science

Let's make the following assumption

activity	time	orders	items	packages	customers	products	price	weight
pick item	2019-12-26 12:04:46	991224}	{884803}	{}	{Wil van der Aalst}	{iPhone 8}	529.0	0.21
reorder item	2019-12-26 12:37:26	991271}	{885002}	{}	{Mohammadreza Fani Sani}	{Kindle Paperwhite}	129.0	0.495
place order	2019-12-26 12:44:23	991283}	{885038,885039}	{}	{Luis Santos}	{MacBook Air,iPad}	2700.0	1.733
pick item	2019-12-26 14:01:16	991266}	{884983}	{}	{Marco Pegoraro}	{MacBook Air}	2200.0	1.25
create package	2019-12-26 14:01:16	991265}	{884975,884974,884978,884971,884970,884973}	{660798}	{Seran Uysal}	{Fire Stick 4K,iPad Pro,iPad Pro,iPad Pro,Fire Stick,Kindle}	3506.97	2.412
send package	2019-12-26 14:16:11	991265}	{884975,884974,884978,884971,884970,884973}	{660798}	{Seran Uysal}	{Fire Stick 4K,iPad Pro,iPad Pro,iPad Pro,Fire Stick,Kindle}	3506.97	2.412
pick item	2019-12-26 14:16:48	991279}	{884987}	{}	{Julia Graf}	{iPhone 11}	799.0	0.166
confirm order	2019-12-26 14:26:01	991283}	{885038,885039}	{}	{Luis Santos}	{MacBook Air,iPad}	2700.0	1.733
reorder item	2019-12-26 14:32:43	991251}	{884912}	{}	{Tobias Brockhoff}	{Fire Stick}	39.99	0.2
confirm order	2019-12-26 14:32:44	991282}	{885036,885037}	{}	{Lisa Mannel}	{Echo,Echo Dot}	134.98	1.16
pick item	2019-12-26 14:33:28	991278}	{885024}	{}	{Junxiong Gao}	{MacBook Pro}	2500.0	1.17
place order	2019-12-26 14:48:33	991284}	{885040,885041,885042,885043,885044}	{}	{Christine Dobbert}	{iPhone X,Fire Stick,MacBook Air,Echo Show 8,iPhone 11 Pro}	4222.98	2.79
failed delivery	2019-12-26 15:04:53	991240,99116...	{884903,884913,884876,884938,884914,884...}	{}	{Tobias Brockhoff}	{iPad Air,Echo Studio,Echo Studio,Kindle,Kindle,Echo,iPad mini,iPad Pro,iPad P...}	5982.95	3.42
pick item	2019-12-26 15:20:05	991278}	{884502}	{}	{Junxiong Gao}	{iPhone X}	699.0	0.172
confirm order	2019-12-26 15:25:00	991258}	{884938,884939,884940,884941,884942,884943}	{}	{Tobias Brockhoff}	{Echo,Fire Stick,iPad mini,iPad Pro,iPad Pro,iPad Air}	3267.96	2.166
send package	2019-12-26 15:26:49	991247,99125...	{884902,884922,884923,885004,885005,884901}	{660796}	{Mohammadreza Fani Sani}	{MacBook Air,iPad mini,iPad Pro,iPhone 11 Pro,iPad Pro,MacBook Pro}	8496.0	4.054
failed delivery	2019-12-26 15:36:16	991265}	{884975,884974,884978,884971,884970,884973}	{660798}	{Seran Uysal}	{Fire Stick 4K,iPad Pro,iPad Pro,iPad Pro,Fire Stick,Kindle}	3506.97	2.412
confirm order	2019-12-26 15:40:51	991274}	{885008,885009,885010,885011}	{}	{Junxiong Gao}	{Kindle,iPhone X,Fire Stick,iPhone 8}	1352.98	1.035
failed delivery	2019-12-26 15:46:21	991128,99125...	{884424,884913,884799,884908,885010,885011,884903}	{660797}	{Junxiong Gao}	{Echo Show 8,Kindle Paperwhite,Streaming Kindle,iPhone X,iPhone 8,Echo Show...}	2145.97	
payment reminder	2019-12-26 15:54:44	991169}	{884571}	{}	{Gyunam Park}	{iPhone 8,Echo Plus,iPad Air}	1608.99	0.811
pick item	2019-12-26 15:55:38	991201}	{884593}	{}	{Junxiong Gao}	{Echo Show 8}	129.99	
pick item	2019-12-26 15:56:16	991251}	{884912}	{}	{Tobias Brockhoff}	{Fire Stick}	39.99	0.2
reorder item	2019-12-26 16:04:42	991265}	{884977}	{}	{Seran Uysal}	{Fire Stick 4K}	89.99	0.28
payment reminder	2019-12-26 16:11:39	991164}	{884542,884543,884544,884545,884546,884547}	{}	{Junxiong Gao}	{Kindle Paperwhite,iPad Air,iPhone 11,MacBook Air,iPad mini,Echo Dot}	4087.99	3.011
pick item	2019-12-26 16:22:04	991241}	{884882}	{}	{Lisa Mannel}	{iPhone 8}	529.0	0.21
create package	2019-12-26 16:22:04	991263,99126...	{884967,884964,884966}	{660799}	{Luis Santos}	{iPad Air,iPhone 8,iPad}	1500.0	1.133

event = activity + timestamp + objects + attributes

Chair of Process and Data Science

The purpose of the **OCEL** standard is to provide a general standard to interchange object-centric event data with multiple case notions. We set the following goals for the standard:

- **Interoperability:** with the provision of the OCEL standard and JSON/XML serializations of OCEL, we want to support a widespread collection of languages and systems.
- **Generalization:** the standard supports the storage of events, objects, and their attributes. Furthermore, the standard can be extended.
- **Provision of a collection of examples:** example logs, extracted from information systems supporting some widespread business processes, are provided for the OCEL standard.
- **Tool/Library Support:** to support the implementation of OCEL in custom applications, tool/library support shall be provided.

```xml
<events>
  <event>
    <string key="id" value="e1"/>
    <string key="activity" value="place_order"/>
    <date key="timestamp" value="2020-07-09T08:20:01.527+01:00"/>
    <list key="omap">
      <string key="object-id" value="i1"/>
      <string key="object-id" value="o1"/>
      <string key="object-id" value="i2"/>
    </list>
    <list key="vmap">
      <string key="resource" value="Alessandro"/>
      <float key="prepaid-amount" value="200.0"/>
    </list>
  </event>
  <event>
    <string key="id" value="e2"/>
    <string key="activity" value="check_availability"/>
    <date key="timestamp" value="2020-07-09T08:21:01.527+01:00"/>
```

http://ocel-standard.org/

PADS | Chair of Process and Data Science | RWTH AACHEN UNIVERSITY

PROCESS AND DATA SCIENCE GROUP

RWTH AACHEN UNIVERSITY

OCEL Standard

Authors:
Anahita Farhang Ghahfarokhi
Gyunam Park
Alessandro Berti
Wil van der Aalst

January 8th, 2020

Event Identifier

<contains>   1..1

Log   0..*   Event   2   1..1   Activity
<contains>   <relates>

1..* <contains>   3   <relates>

5   4

Timestamp   1..1

1..*

<contains>   0..*   Attribute   <contains>   Attribute Name   1   <relates>   1..1

Object   7   0..*   <relates>   1..1

0..*   <relates>   Attribute Type   1..1

<contains>   <contains>   Attribute Value   1..1

6   <relates>

Object Identifier   1..1   Object Type   1..1   <relates>   1
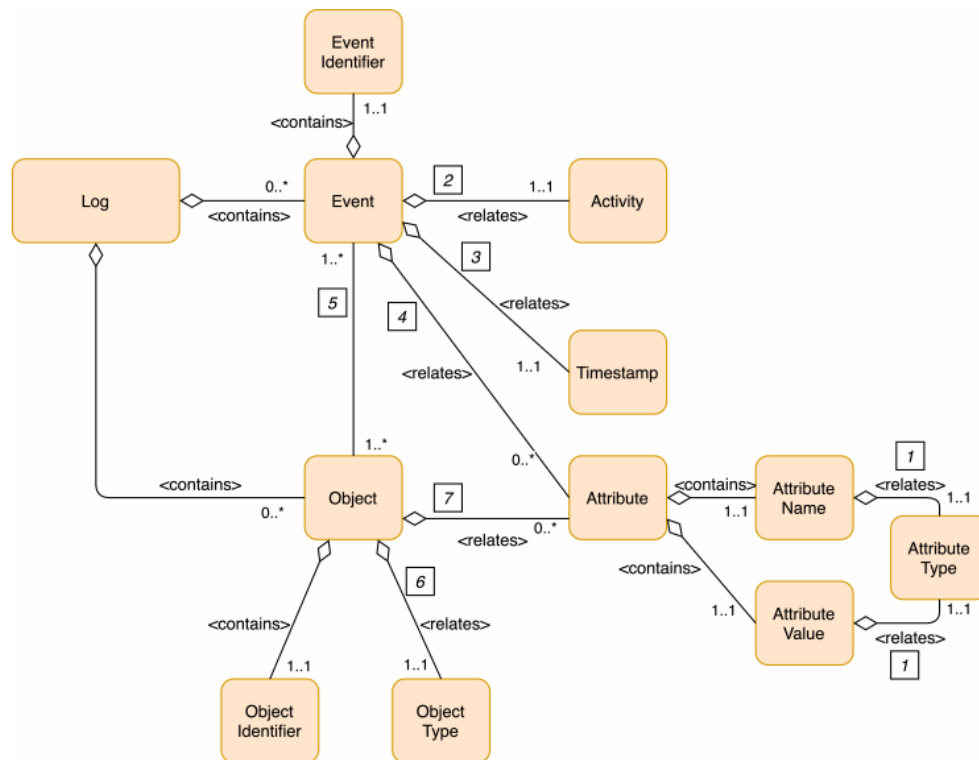
How to deal with this?

Directly extracting one or more conventional event logs (e.g. XES) realizing that there are may be convergence and divergence problems.

Extracting one object-centric event log and creating conventional event logs (e.g. XES) on demand.

Extracting one object-centric event log and using process mining techniques directly working on object-centric event logs.

Object-Centric Process Mining

Free Advice

PDAS Chair of Process and Data Science

Make process mining repeatable and actionable

free advice

- ❑ Event logs provide <u>views on reality</u>.
- ❑ Process mining is <u>not</u> a project, but an ongoing activity.
  - ❑ The return-on-investment is typically low for one time extractions (proof-of-concepts are fine, but …).
  - ❑ Results should be used on a daily basis.
- ❑ Not for one process, but for <u>all</u> processes you would like to improve.
  - ❑ Share efforts and expertise.
  - ❑ Use comparative process mining / benchmarking.

Business process hygiene

Typical excuses: privacy, data quality, workload, etc.

Are you sure you need to have a business case?

www.tf-pm.org

IEEE TASK FORCE ON PROCESS MINING

PADS Chair of Process and Data Science

RWTH AACHEN UNIVERSITY