



Knowledge Engineering through Process Mining The Practice

By Linda Terlouw (linda.terlouw@icris.nl)

Why process mining?



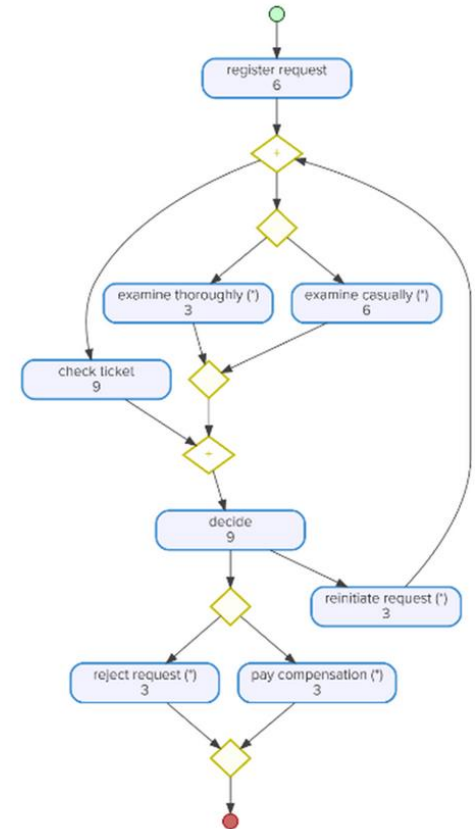
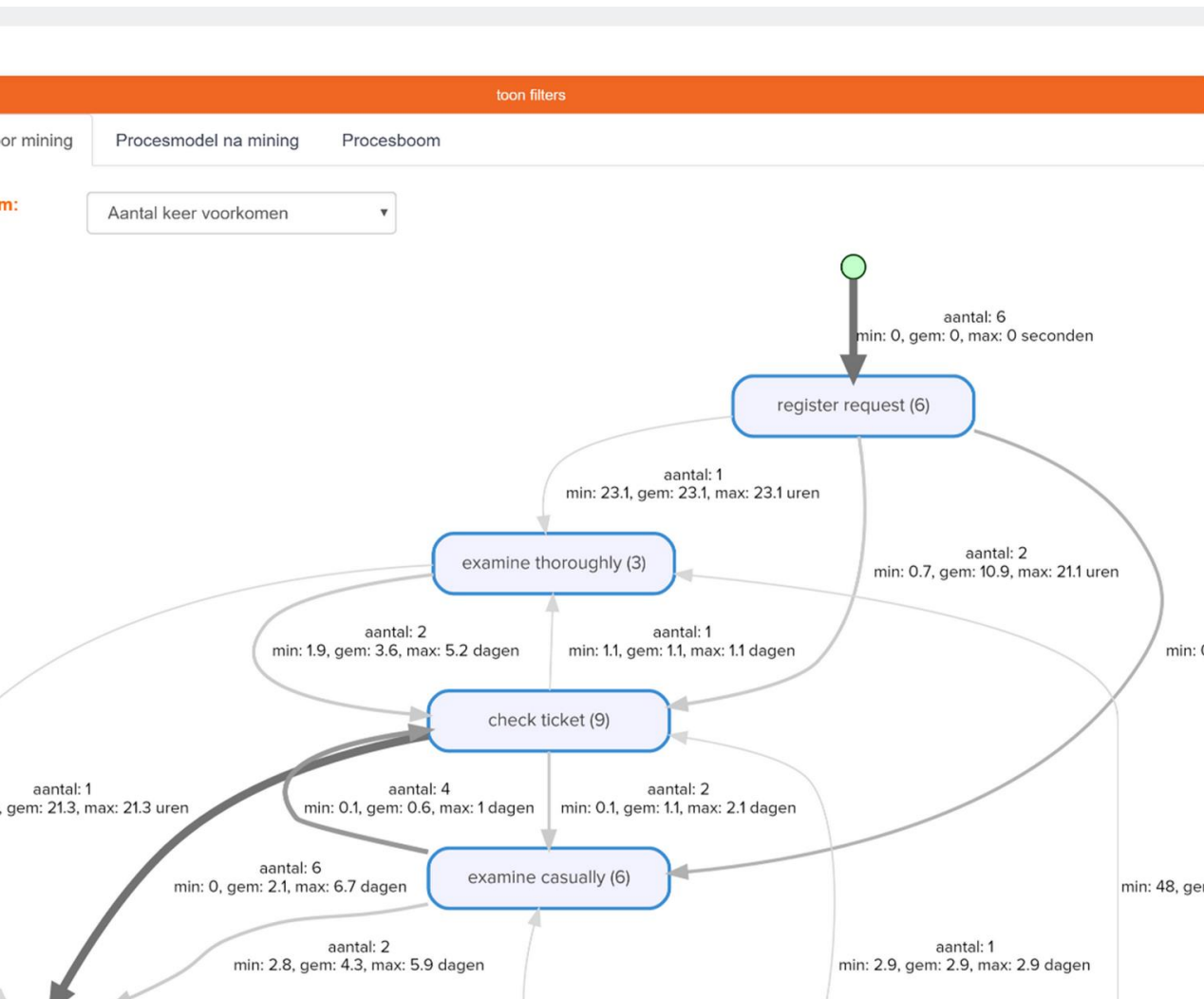
<i>Item</i>	<i>Description</i>
Case id	the identifier of the process instance
Activity	the name of the activity that is performed
Timestamp	the date (and time) the activity is performed (preferably split into a start timestamp and an end timestamp)
Resource	the person/role that performed the activity
Additional attributes of the event	whatever may be of interest, e.g. cost of the activity, responsible department, process category, priority

Contents of an event log

<i>Case id</i>	<i>Activity</i>	<i>Timestamp</i>	<i>Resource</i>
1	Register order	2016-08-01	Tim
1	Order part at supplier	2016-08-03	Ellen
1	Deliver order	2016-08-12	Ali
2	Register order	2016-08-01	Tim
2	Cancel order &	2016-08-02	Tim
3	Register order	2016-08-02	Tim
3	Pick part from warehouse	2016-08-02	Sonya
3	Deliver order	2016-08-03	Ali

An example event log

Mined models

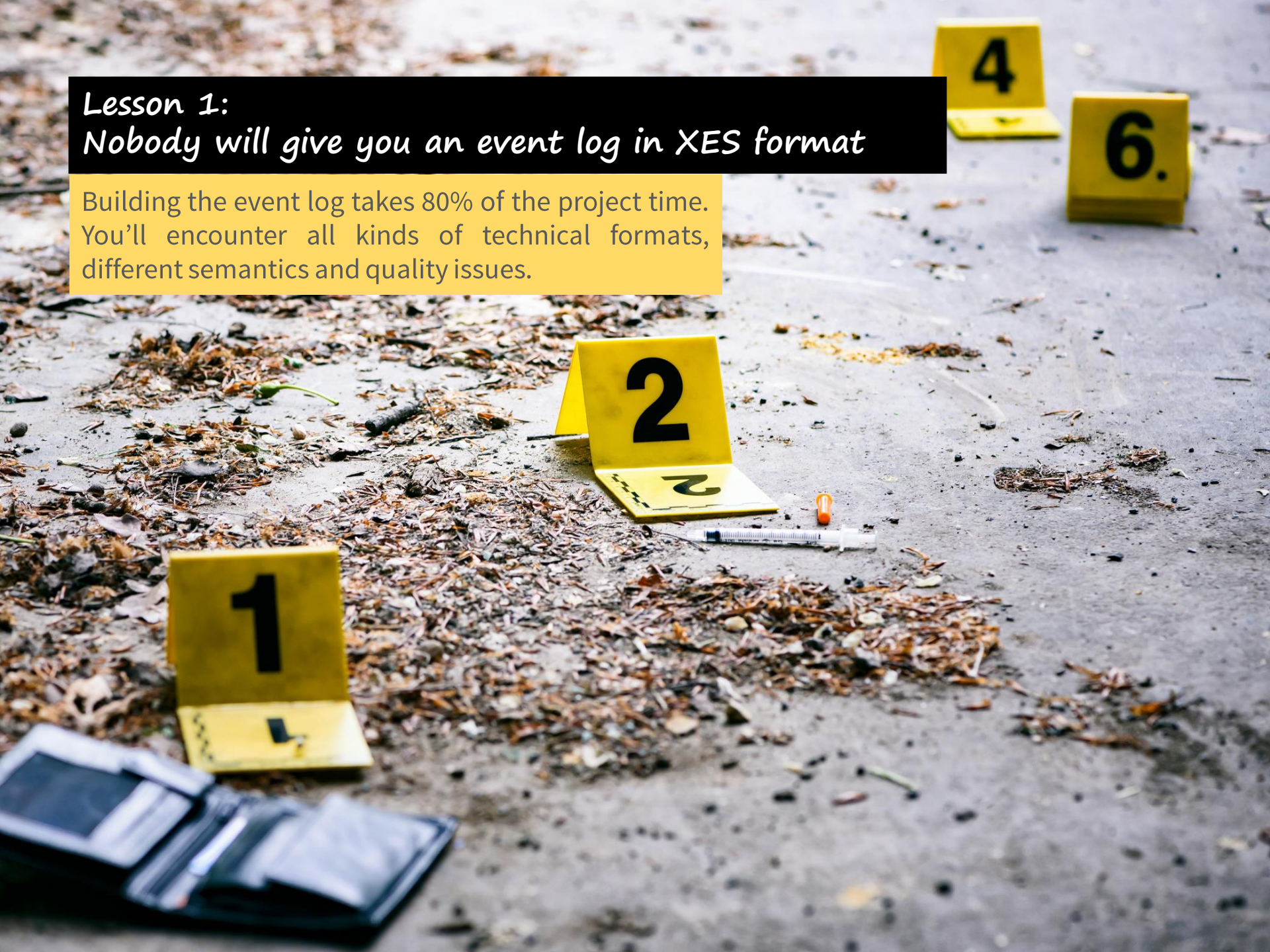





Lessons
learned

*Lesson 1:
Nobody will give you an event log in XES format*

Building the event log takes 80% of the project time. You'll encounter all kinds of technical formats, different semantics and quality issues.





Assemble a *process mining CSI team*. Include domain expert in the team to get an understanding of the data. *Data needs interpretation*. Make sure your data science team members *have scripting skills* to construct the event log from multiple sources.

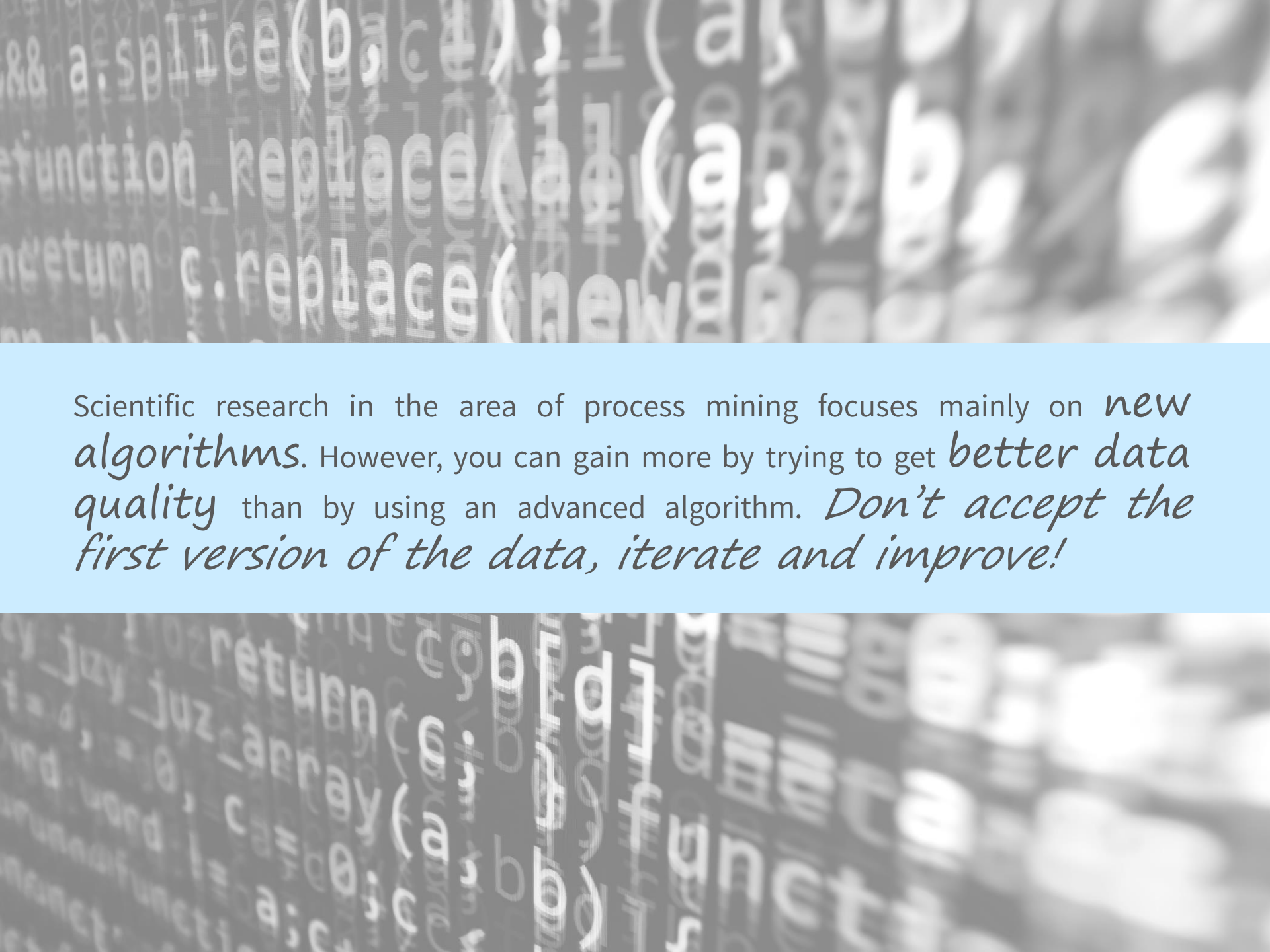
You don't know what you will find!



Lesson 2:

It is the data that matters, not the fancy algorithm

The first data set is never the final one.

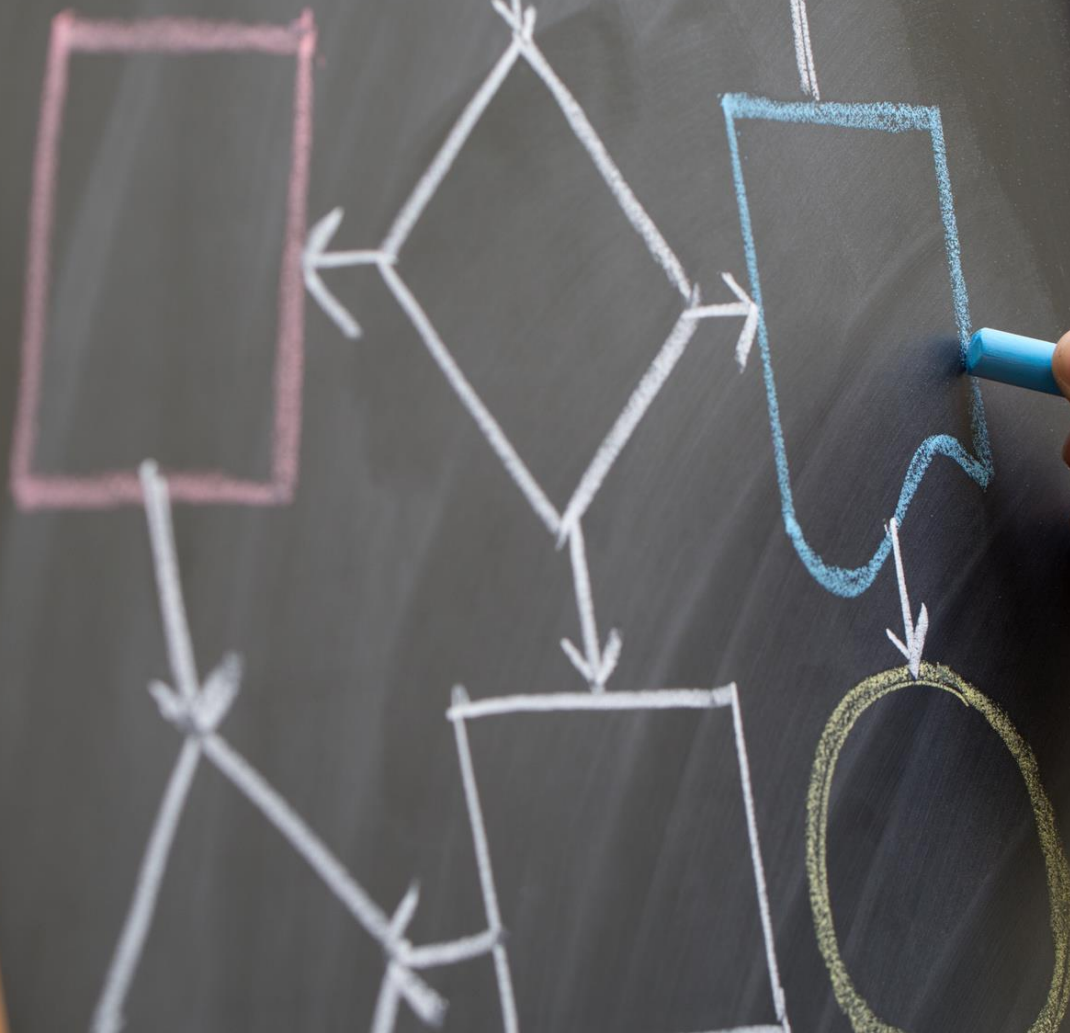
The background of the slide is a dark grey color with a pattern of blurred, light grey text. The text appears to be snippets of code or technical terms, such as 'function', 'return', 'new', 'data', and 'quality', scattered across the surface. The text is out of focus, creating a sense of depth and a technical atmosphere.


Scientific research in the area of process mining focuses mainly on *new algorithms*. However, you can gain more by trying to get *better data quality* than by using an advanced algorithm. *Don't accept the first version of the data, iterate and improve!*

Lesson 3:


Only scientists, BPM people and auditors like flow charts

The rest of the world does not.





The main way of visualizing process mining results is in *process diagrams*, often *petri nets* in scientific tools and less formal *flow charts* in commercial tools. These process diagrams are usually very popular with people designing and auditing processes. People performing the processes usually do not like them. So *provide people with the visualizations that they understand.*




A hand-drawn flowchart on a chalkboard. It features a central rectangular box with an arrow pointing down into it from above. To the right of the box is a circle with an arrow pointing down into it from above. To the left of the box, there are two arrows pointing towards it from the left, one above the other. The drawing is done in white chalk on a dark grey or black chalkboard.

*Lesson 4:
ERP data is just 5% of the process data*

ERP data is relatively easy to extract and reflects the formal process. However, there is always an informal process and many other sources of process data are available.





Don't trust the formal information systems. Ask people how they actually work and you'll discover many *informal information systems* (Excel, Access etc). Also, search for more reliable data source like sensor data. Don't close your eyes for *unstructured data*.

*Lesson 5:
Good processes do not imply good products*

A good process is only a means to an end.





An overengineered process that is perfect on paper can lead to unnecessary *bureaucracy*. This bureaucracy can lead to *frustrated employees*, not feeling in control. They may leave the company because 'people are redesigning their work'. The *process 'quality'* may be improved, but the *product quality* is not. So be careful what you wish for and don't forget to include the people who do the work.



*Lesson 6:
Not everybody will share your enthusiasm*

It can be hard to get data and people might benefit from things not being transparent.



A black and white photograph of a man and a young girl peeking over a yellow horizontal banner. The man on the left has a wide-eyed, surprised expression. The girl on the right is pointing her right index finger upwards and has a happy, open-mouthed smile.

Think about what people have to *gain* or have to *lose*. The turkey may not be willing to give you the Christmas dinner recipe.




Lesson 7:

Process mining is just the first step of a long journey

You're not done when you think you are.

The Beginning...



Knowledge without a way to take actions for improvement is *useless*. If you know things about the process, this does not mean people will *listen to you* and will actually change the way of working. Change management is *hard* and *takes time*. Don't analyze for other people, analyze with other people. Process mining can be a *perfect excuse* not to talk to people!

Recap

- 1: Nobody will give you an event log in XES format
- 2: It is the data that matters, not the fancy algorithm
- 3: Only scientists, BPM people and auditors like flow charts
- 4: ERP data is just 5% of the process data
- 5: Good processes do not imply good products
- 6: Not everybody will share your enthusiasm
- 7: Process mining is just the first step of a long journey

Don't be *afraid* and just start with process mining.
There's nothing to be afraid of.



Don't be *afraid* and just start with process mining.
There's nothing to be afraid of.





Questions?

