



Towards the detection of Promising Processes by Analysing the Relational Data

Belén Ramos Gutiérrez, Luisa Parody, María Teresa Gómez-López

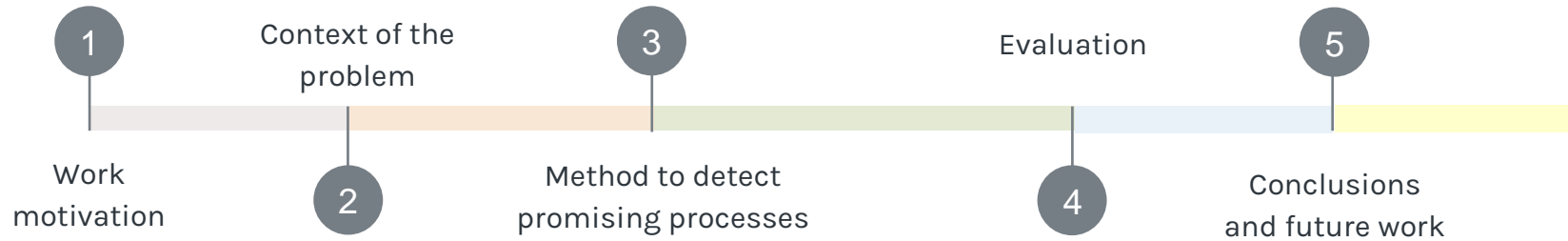
IDEA Research Group

Universidad de Sevilla

www.idea.us.es - brgutierrez@us.es

[@belenramgut](https://twitter.com/belenramgut)

Outline



1. Work motivation



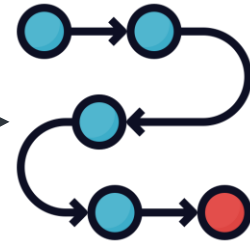
Business Process
Management System
(BPMS)



Process Mining



Inferred behaviour of the
process

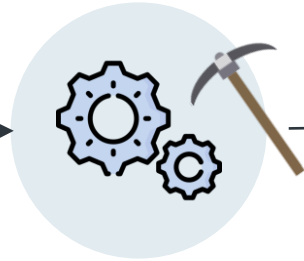


1. Work motivation

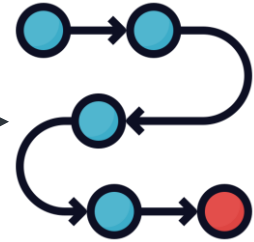
What if there is not a Business Process Management System?



Business Process Management System (BPMS)

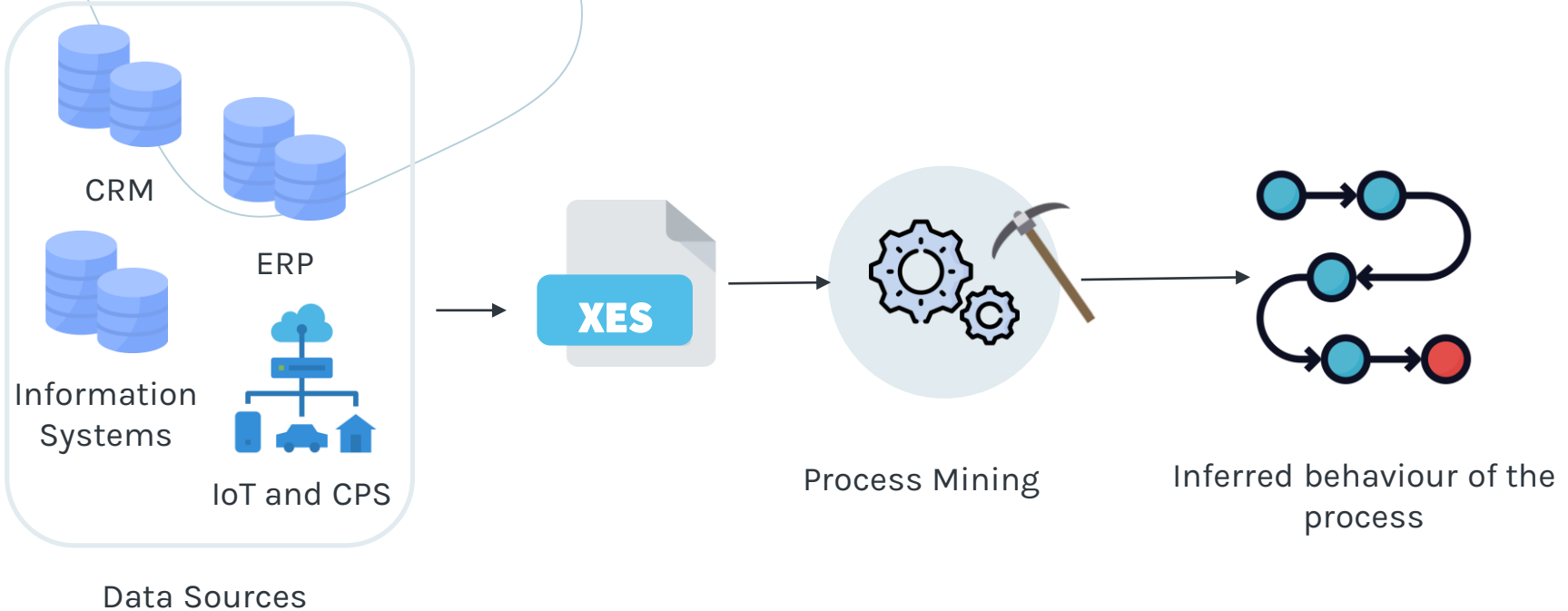


Process Mining

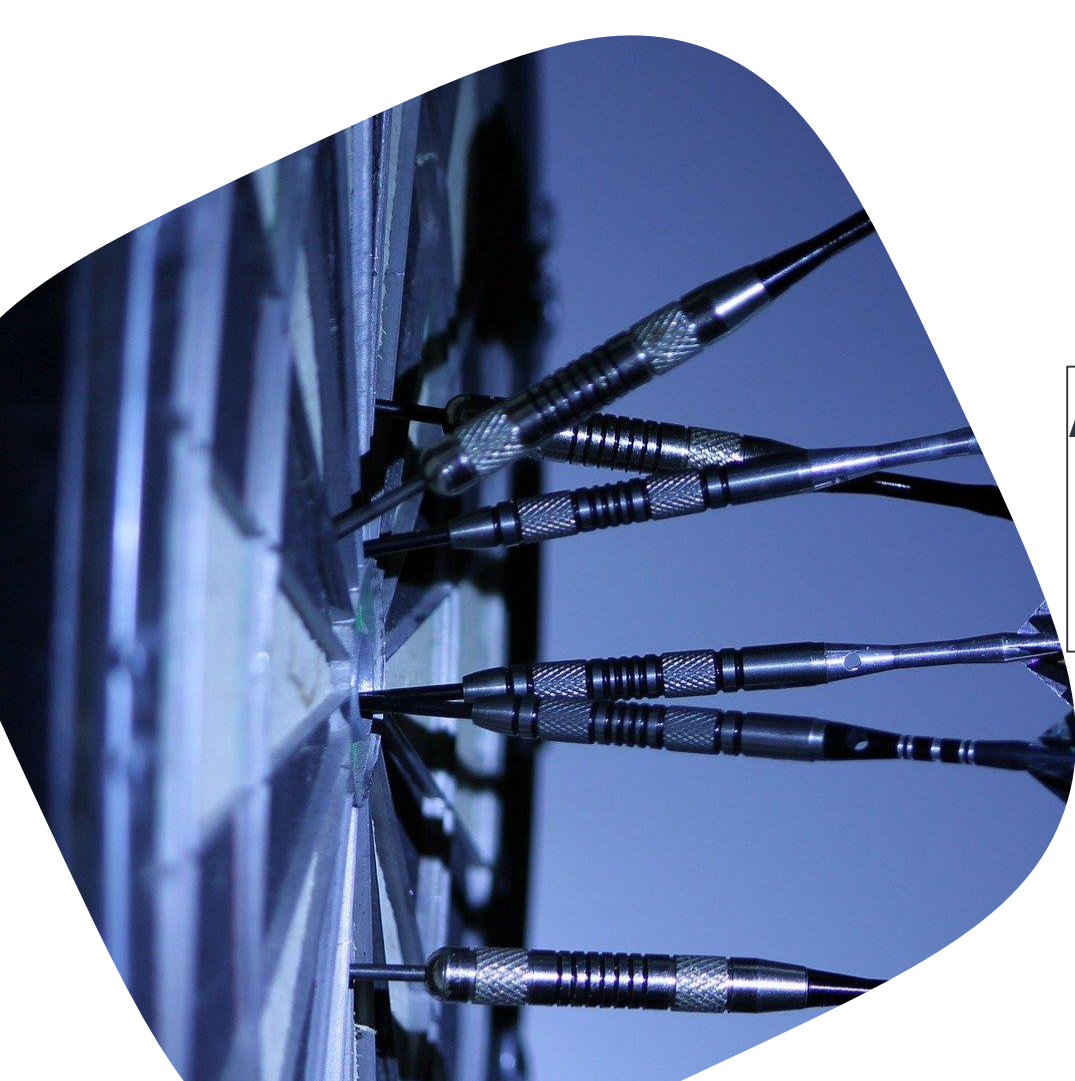


Inferred behaviour of the process

1. Work motivation



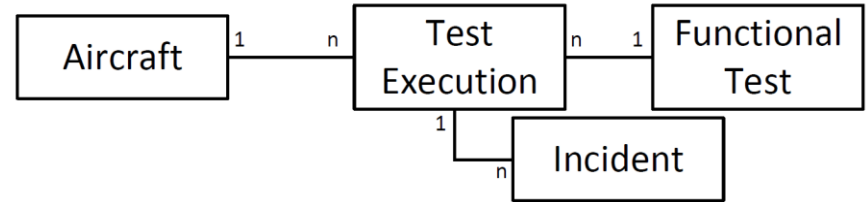




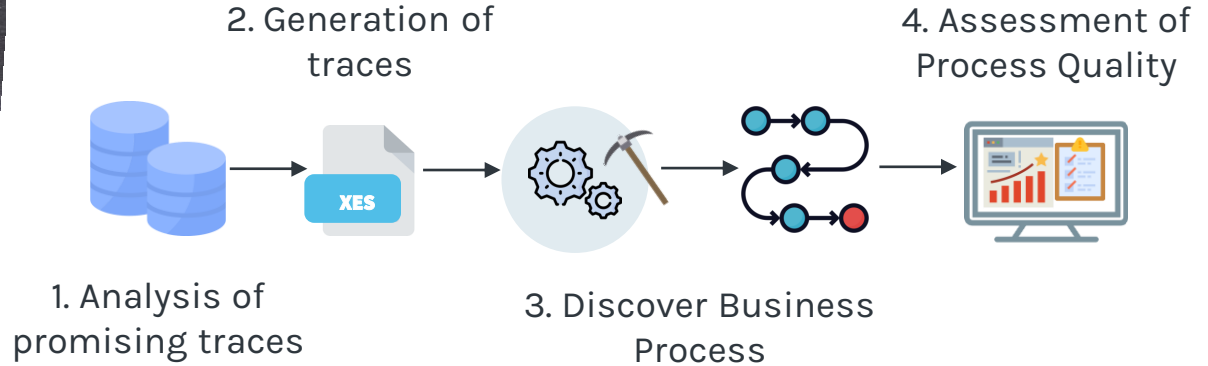
A big challenge

Automate the detection
of promising traces in
relational data sources.

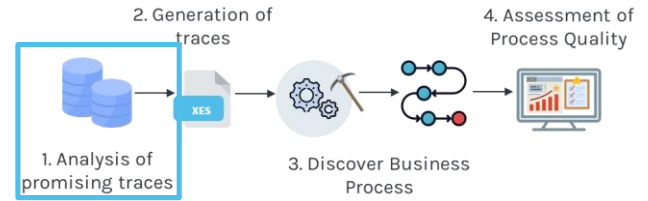
2. Context of the problem



3. Method to detect promising processes



3. Method to detect promising processes



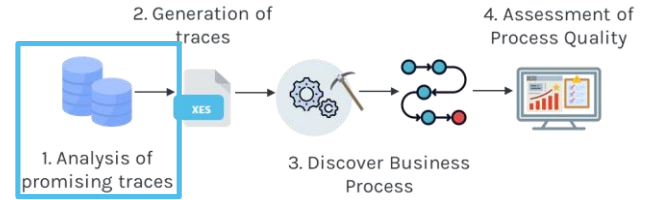
1

IDENTIFY THE POSSIBLE *CASE_ID*



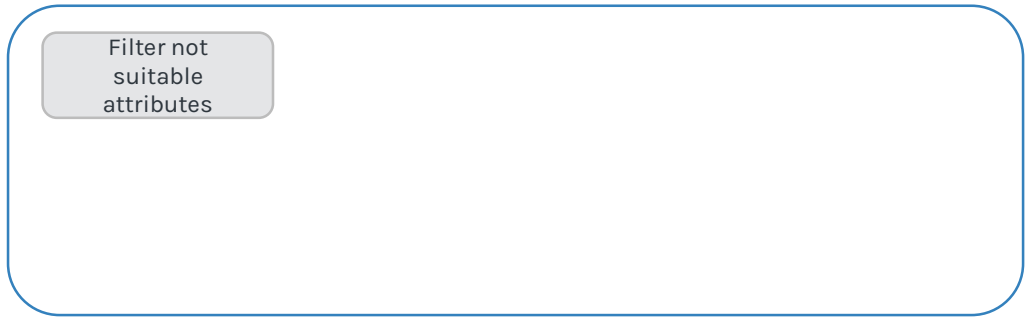
PRIMARY KEYS

3. Method to detect promising processes

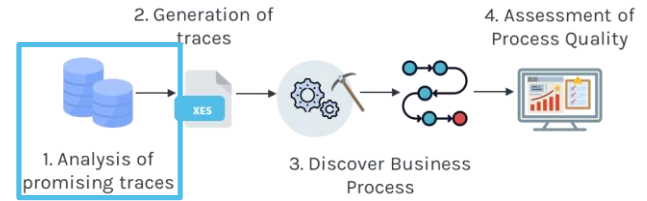


1 IDENTIFY THE POSSIBLE *CASE_ID* → PRIMARY KEYS

2 IDENTIFY THE *EVENTS* →



3. Method to detect promising processes



1

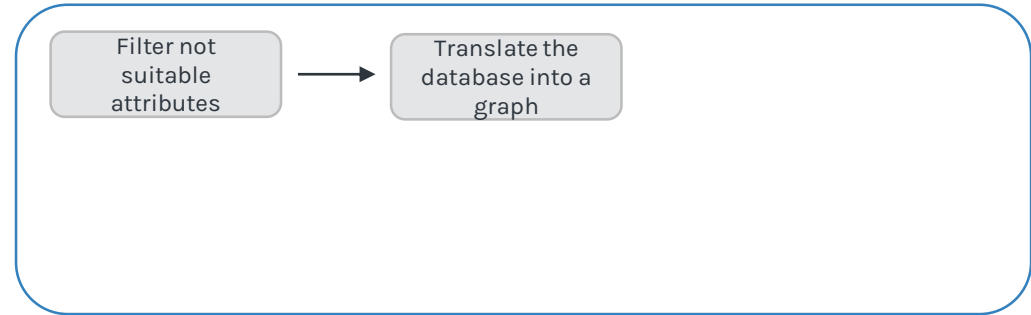
IDENTIFY THE POSSIBLE *CASE_ID*



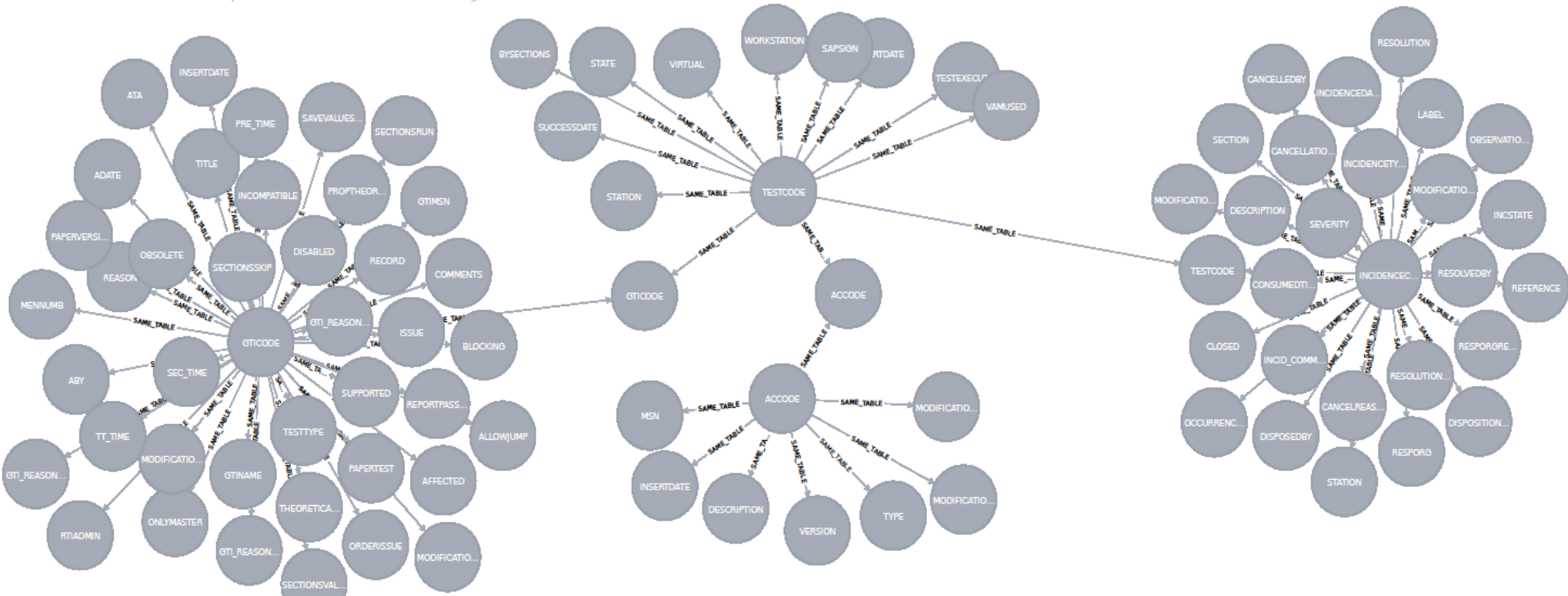
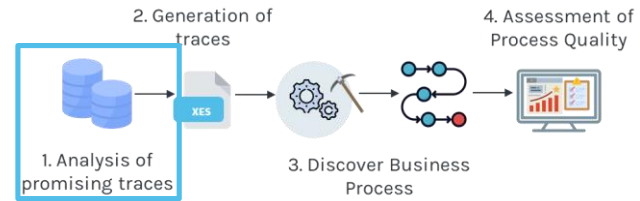
PRIMARY KEYS

2

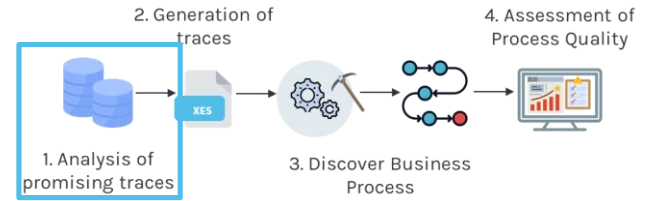
IDENTIFY THE *EVENTS*



3. Method to detect promising processes



3. Method to detect promising processes



1

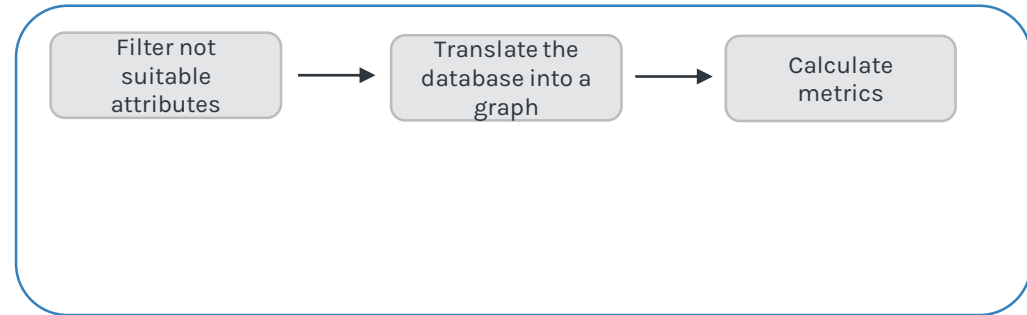
IDENTIFY THE POSSIBLE *CASE_ID*



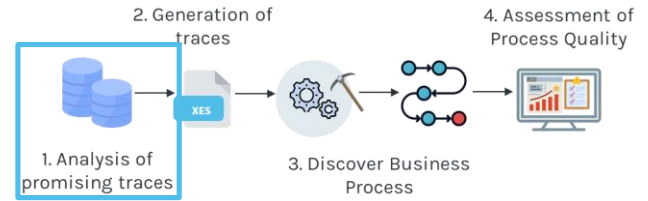
PRIMARY KEYS

2

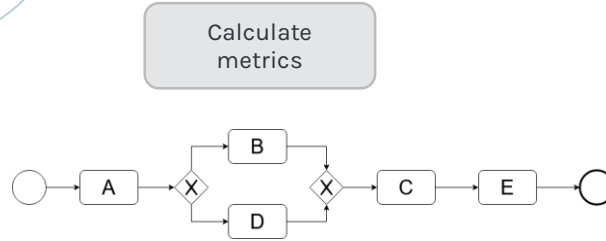
IDENTIFY THE *EVENTS*



3. Method to detect promising processes



- A, B, C
- A, D, C
- A, B, C, E

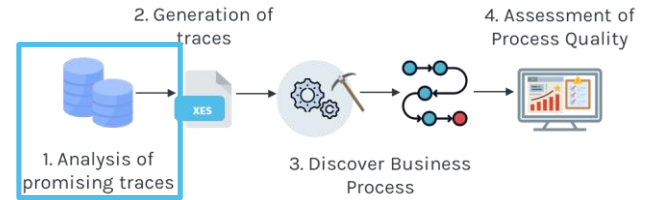


Complexity

Average number of events per trace.

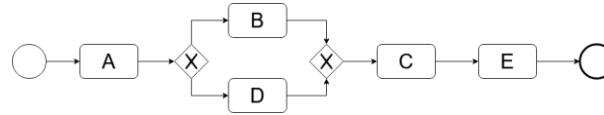
$$\frac{5}{3} = 1.66$$

3. Method to detect promising processes



A	B	C	
A	D	C	
A	B	C	E

Calculate metrics



Complexity

Average number of events per trace.

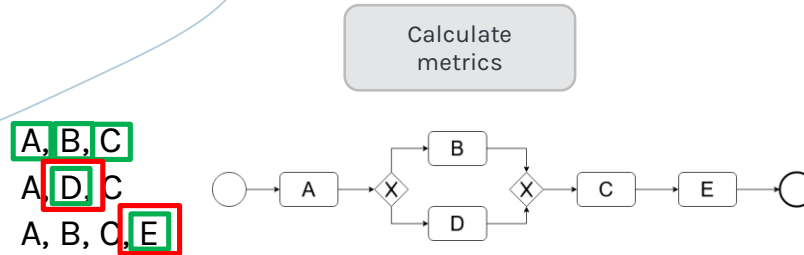
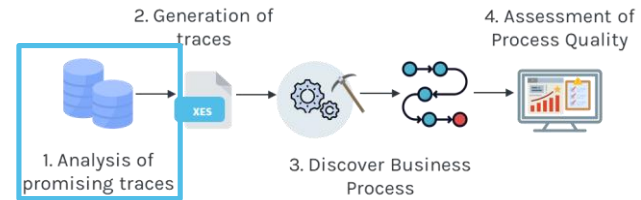
$$\frac{5}{3} = 1.66$$

Diversity

Density of different events that occur in all the traces among all the events that are presented in the log

$$\frac{2}{5} = 0.4$$

3. Method to detect promising processes



Complexity

Average number of events per trace.

$$\frac{5}{3} = 1.66$$

Diversity

Density of different events that occur in all the traces among all the events that are presented in the log

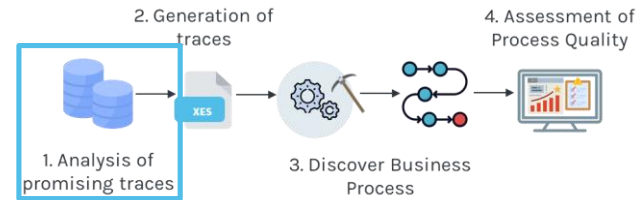
$$\frac{2}{5} = 0.4$$

Noise

Average of events that only occur once in the whole log among all the events inside of it

$$\frac{2}{5} = 0.4$$

3. Method to detect promising processes

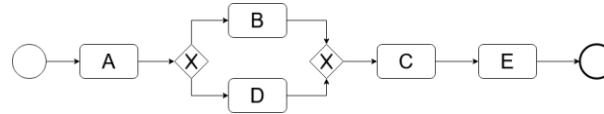


Calculate metrics

A, B, C

A, D, C

A, B, C, E



Complexity

Average number of events per trace.

$$C_n(C_i) = \begin{cases} C_i < C_{q1} & (\frac{1}{C_{q1}}) \cdot C_i \\ C_{q1} < C_i < C_{q3} & 1 \\ C_i > C_{q3} & (\frac{-C_i + C_{q3}}{C_{max} - C_{q3}}) + 1 \end{cases}$$

Diversity

Density of different events that occur in all the traces among all the events that are presented in the log

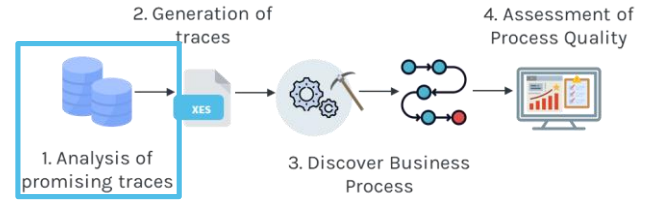
$$D_n(D_i) = \begin{cases} D_i < D_{mean} & (\frac{1}{D_{mean}}) \cdot D_i \\ D_i > D_{mean} & (\frac{-D_i + D_{mean}}{1 - D_{mean}}) + 1 \end{cases}$$

Noise

Average of events that only occur once in the whole log among all the events inside of it

$$N_n(N_i) = -N_i + 1$$

3. Method to detect promising processes



1.1

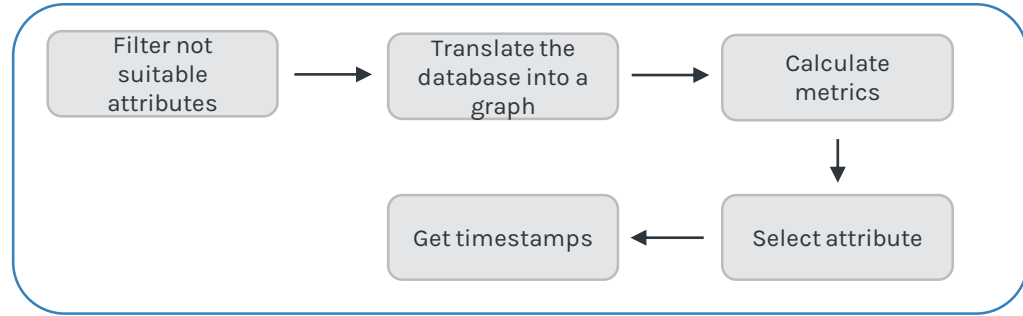
IDENTIFY THE POSSIBLE *CASE_ID*



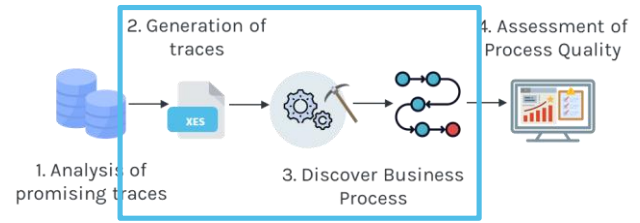
PRIMARY KEYS

1.2

IDENTIFY THE *EVENTS*



3. Method to detect promising processes



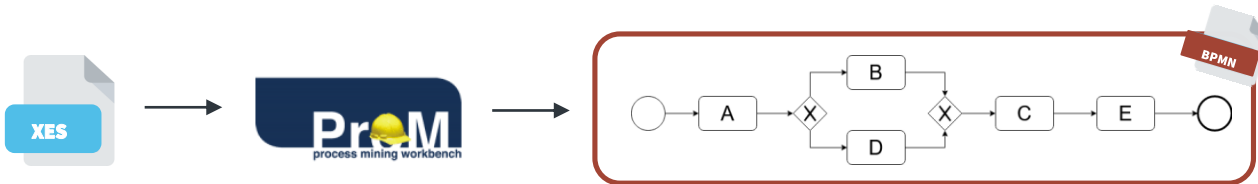
2

EVENT LOG EXTRACTION



3

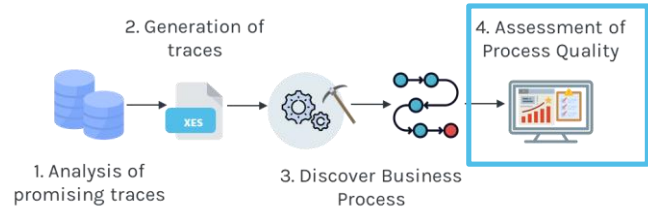
PROCESS DISCOVERY



3. Method to detect promising processes

Will Level metric

The mean of possible tasks that can be selected in each step according to the process model, divided into the number of total tasks of the process.

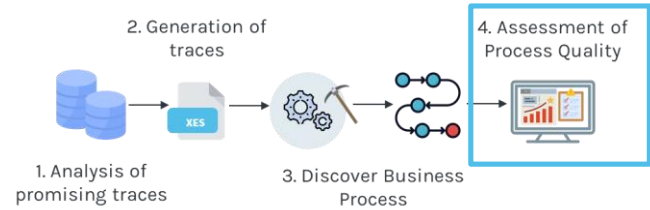


Translate the BPMN model into a graph

3. Method to detect promising processes

Will Level metric

The mean of possible tasks that can be selected in each step according to the process model, divided into the number of total tasks of the process.

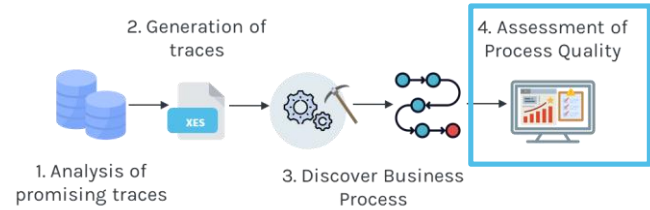


Translate the BPMN model into a graph



Execute Dijkstra algorithm

3. Method to detect promising processes



Will Level metric

The mean of possible tasks that can be selected in each step according to the process model, divided into the number of total tasks of the process.



4. Evaluation

4 TABLES

A very small portion of the database

1 CASE_ID

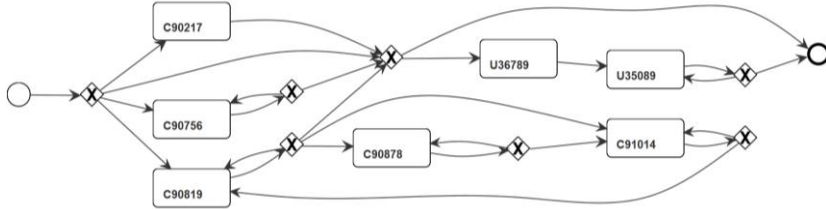
Further reduction of data

58 POSSIBLE EVENT
TYPES

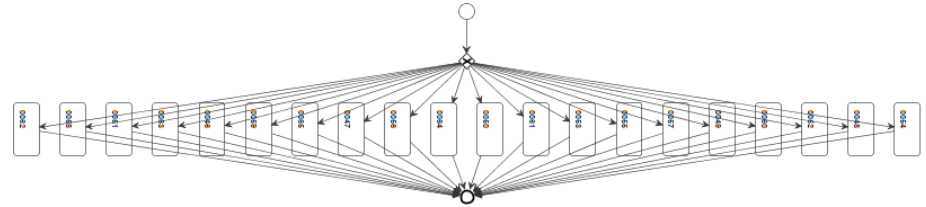


4. Evaluation

Well rated



Bad rated



Complexity

Diversity

Noise

Will Level

Well rated

1

0.893

0.999

0.21

Bad rated

0.005

0.094

0

0.95

PRE DISCOVERY

POST DISCOVERY

5. Conclusions and future work

CONCLUSIONS

- 👍 A method to speed up the process
- 👍 A set of metrics to profile log and process quality
- 👎 Although it is automated it still needs expert knowledge

FUTURE WORK

- New metrics and indicators
- Further analysis on the database structure



ANY QUESTIONS?

You can find me at [@belenramgut](#) & brgutierrez@us.es