

# MINERAÇÃO DE PROCESSOS

- o que é?
- para que serve?

**Dra. Sarajane Marques Peres**

Professora Associada na Escola de Artes, Ciências e Humanidades  
Universidade de São Paulo

PPGComp – 2021

Março, 2021

**Ciência de processos**  
melhorar a gestão e execução de processos operacionais



- técnicas de otimização
- métodos formais e teoria da computação
- gestão e pesquisa operacional
- gestão de processos de negócio
- técnicas de descoberta de processos
- técnicas de melhoria de processos
- gestão e automação de processos

**Ciência de dados**  
transformar dados em valor



**Descoberta de modelos de processo**

Fitness

Precisão

Generalização

Simplicidade

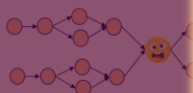
**Log de Eventos**



# Agenda

- Contexto
- Processos + Dados
- Log de eventos
- Descoberta
- Conformidade
- Melhoria
- Suporte operacional
- Informações adicionais

**Melhoria de processos**



Analisar gargalos e valor



Industriais



o processo real



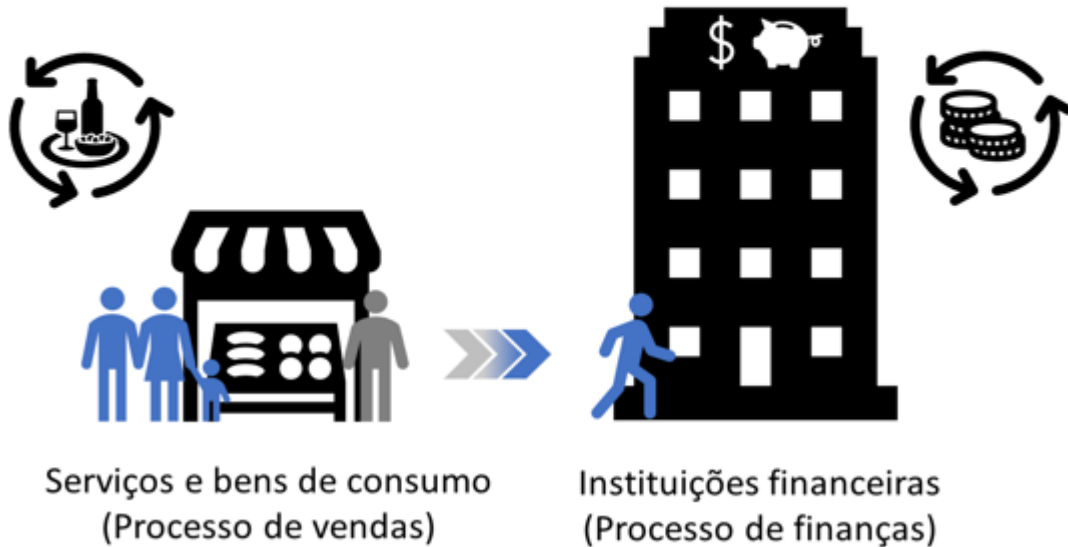
# Contexto



# Contexto

## O que é um processo de negócio?

Conjunto de atividades que uma organização realiza em uma determinada ordem para atingir um objetivo de negócio.



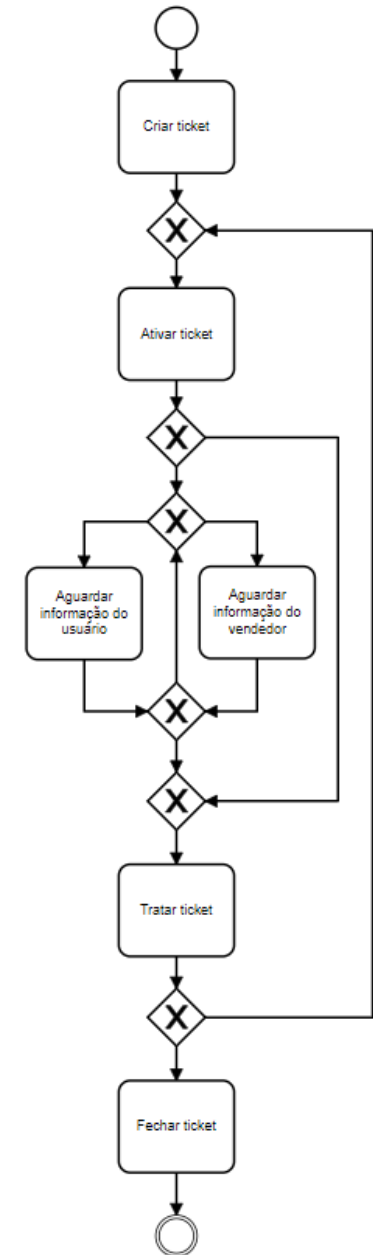
# Contexto

## O que é um processo de negócio?

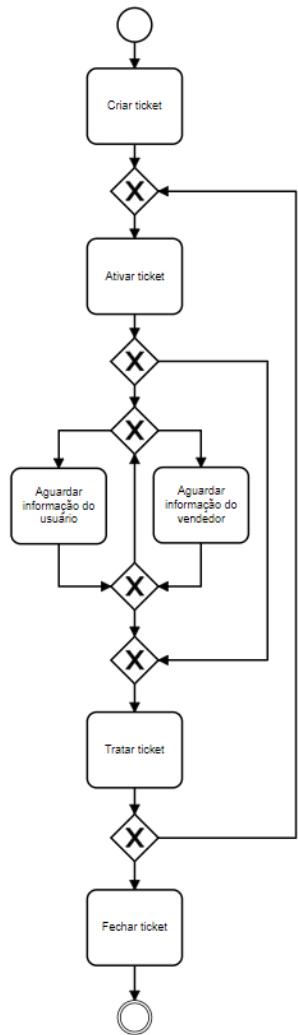
Conjunto de atividades que uma organização realiza em uma determinada ordem para atingir um objetivo de negócio.



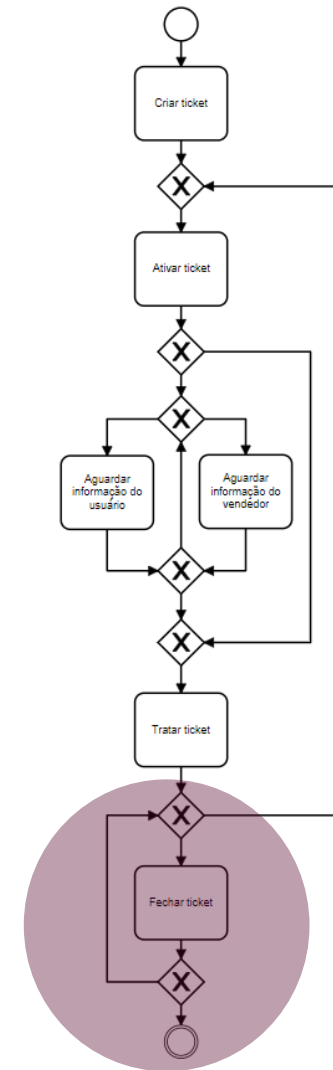
## Modelo de processo



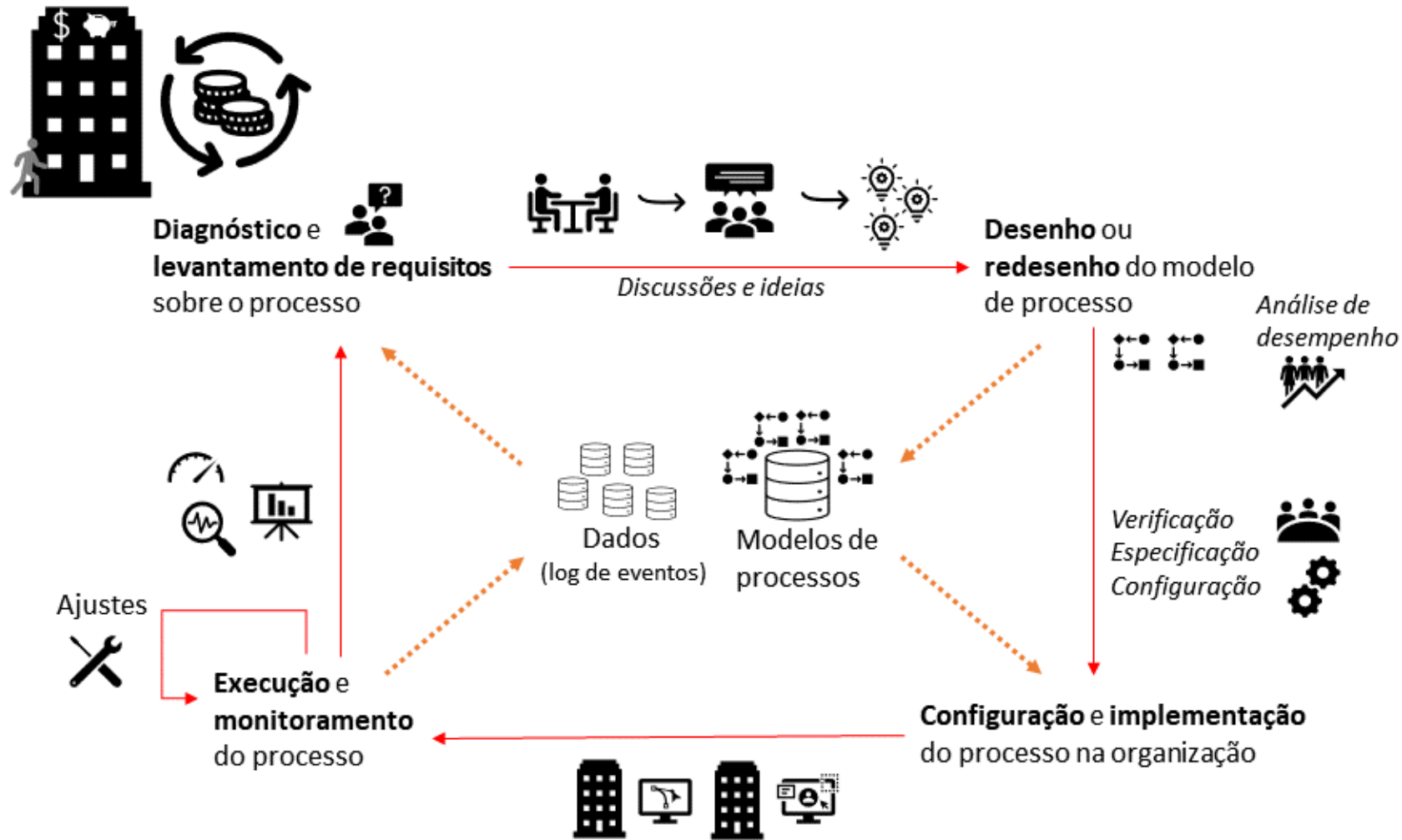
# O modelo de processo normativo *versus* o processo real



<https://medium.com/@bohb.blair/desire-lines-discovering-the-path-people-are-on-despite-the-path-your-brand-planned-83a863827b1c>

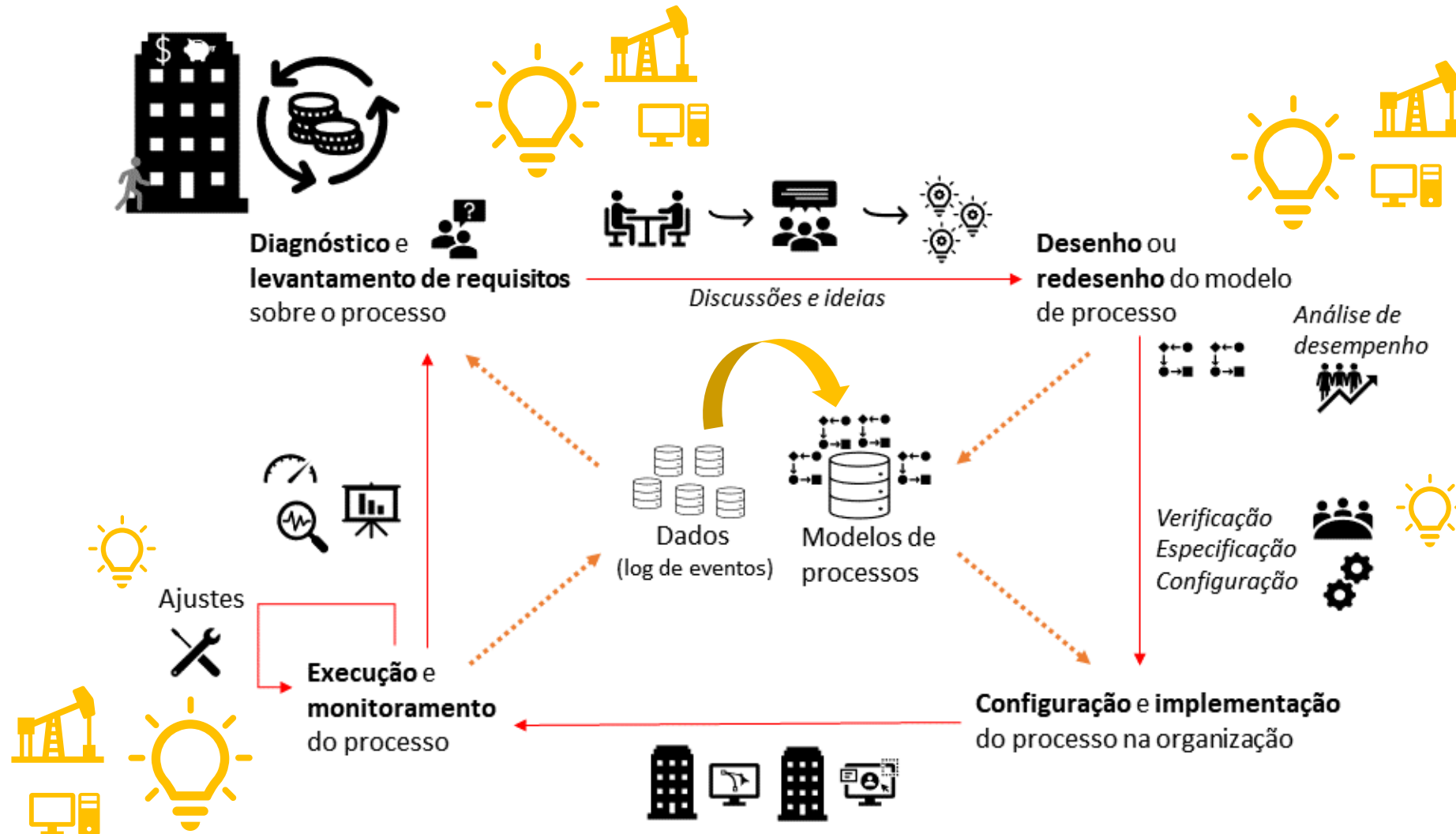


# Contexto



Adaptado de Wil van der Aalst, Process Mining: Data Science in Action, 2nd ed., Springer, 2016, p. 31.

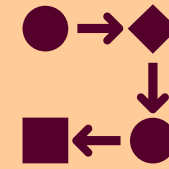
# Contexto



Adaptado de Wil van der Aalst, Process Mining: Data Science in Action, 2nd ed., Springer, 2016, p. 31.



# Processos + Dados



## Ciência de processos

*melhorar a gestão e execução de processos operacionais*



- técnicas de otimização
- métodos formais e teoria da concorrência
- gestão e pesquisa operacional
- gestão de processos de negócio
- técnicas de descoberta de processos
- técnicas de melhoria de processos de negócio
- gestão e automação de processos

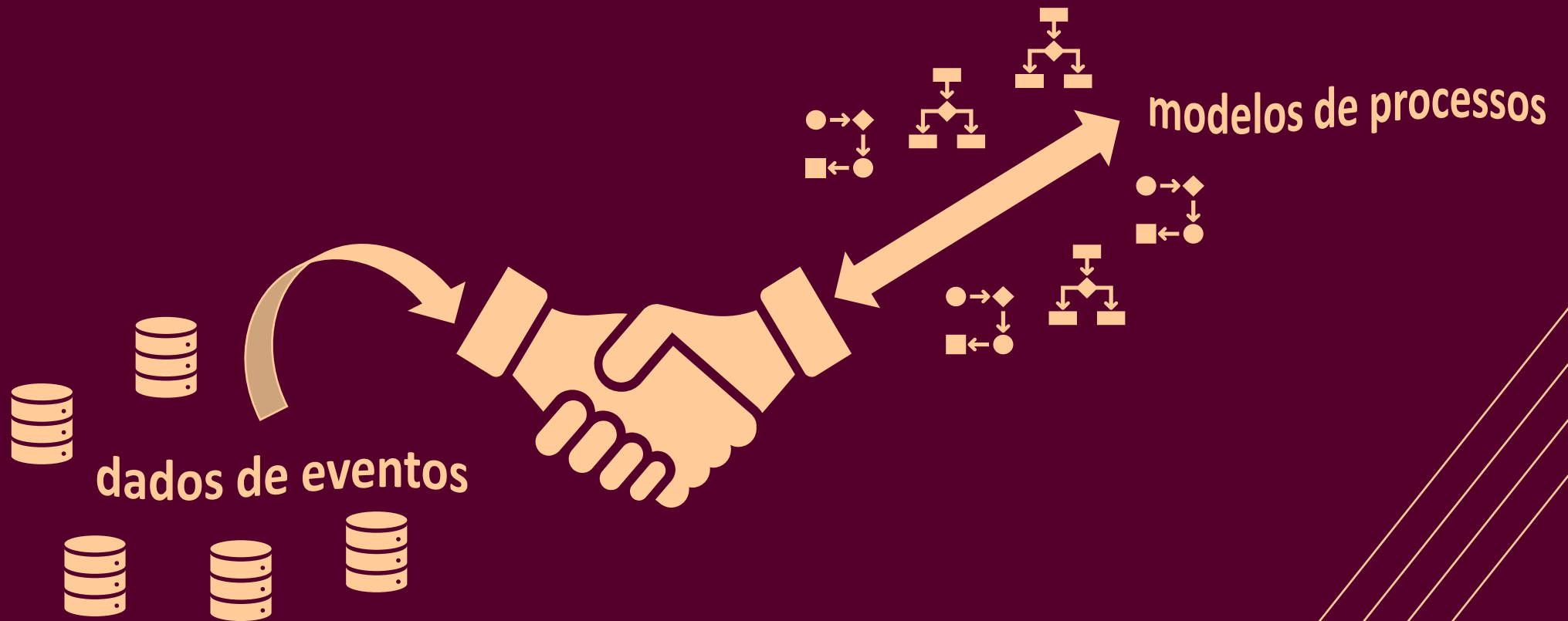
## Ciência de dados

*transformar dados em valor*



- extração, preparação, exploração, transformação de dados
- armazenamento e recuperação
- infraestrutura computacional
- mineração de dados e aprendizado
- apresentação de explicações e previsões
- observação de aspectos éticos, sociais, legais e de negócio

**Ciência de processos + ciência de dados = mineração de processos**

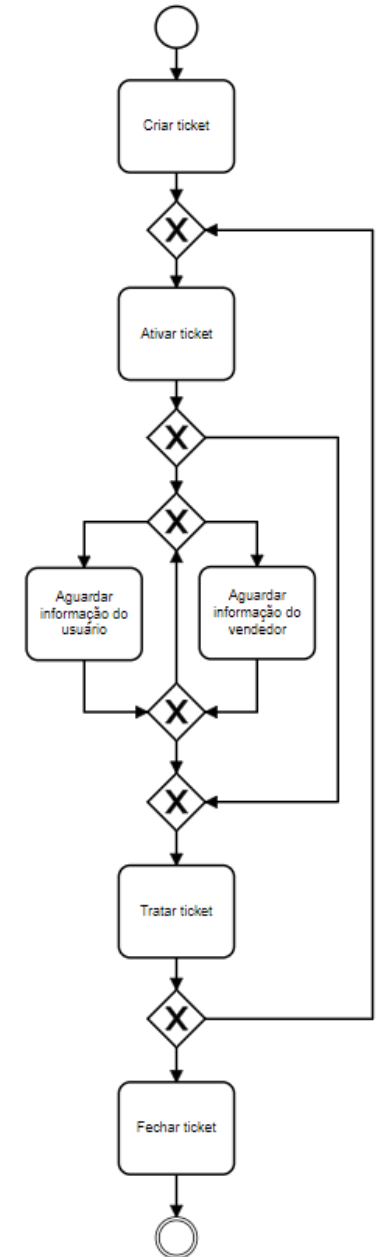


# Log de eventos



# Informação básica e de interesse – log de eventos

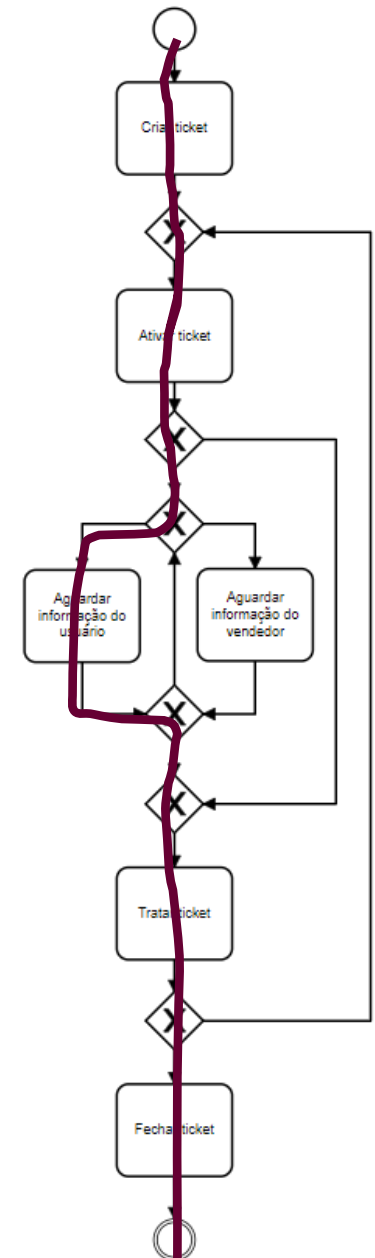
Id. Caso	Id. Evento	Propriedades				
		Timestamp	Atividade	Recurso	Custo	...
1	1	20/02/2021-10:30	Criar ticket	Joana	10	
	2	20/02/2021-10:35	Ativar ticket	Paulo	50	
	3	20/02/2021-11:10	Aguardar informação do usuário	Cris	100	
	4	26/02/2021-15:00	Tratar ticket	Cris	50	
	5	26/02/2021-18:00	Fechar ticket	Cris	10	
2	6	23/02/2021-16:00	Criar ticket	Joana	10	
	7	23/02/2021-16:10	Ativar ticket	Paulo	50	
	8	23/02/2021-16:12	Tratar ticket	Paulo	50	
	9	23/02/2021-16:48	Fechar ticket	Paulo	10	
3	10	20/02/2021-10:35	Criar ticket	Ana	10	
	11	20/02/2021-10:40	Ativar ticket	Paulo	50	
	12	20/02/2021-15:50	Aguardar informação do usuário	Cris	100	
	13	25/02/2021-11:40	Tratar ticket	Cris	50	
	14	25/02/2021-12:40	Ativar ticket	Cris	70	
	15	26/02/2021-08:00	Tratar ticket	Cris	70	
	16	26/02/2021-14:15	Fechar ticket	Cris	20	
...	...	...	...	...	...	



## Informação básica e de interesse – log de eventos

Id. Caso	Id. Evento	Propriedades				
		Timestamp	Atividade	Recurso	Custo	...
1		1 20/02/2021-10:30	Criar ticket	Joana	10	
		2 20/02/2021-10:35	Ativar ticket	Paulo	50	
		3 20/02/2021-11:10	Aguardar informação do usuário	Cris	100	
		4 26/02/2021-15:00	Tratar ticket	Cris	50	
		5 26/02/2021-18:00	Fechar ticket	Cris	10	
2		6 23/02/2021-16:00	Criar ticket	Joana	10	
		7 23/02/2021-16:10	Ativar ticket	Paulo	50	
		8 23/02/2021-16:12	Tratar ticket	Paulo	50	
		9 23/02/2021-16:48	Fechar ticket	Paulo	10	
3		10 20/02/2021-10:35	Criar ticket	Ana	10	
		11 20/02/2021-10:40	Ativar ticket	Paulo	50	
		12 20/02/2021-15:50	Aguardar informação do usuário	Cris	100	
		13 25/02/2021-11:40	Tratar ticket	Cris	50	
		14 25/02/2021-12:40	Ativar ticket	Cris	70	
		15 26/02/2021-08:00	Tratar ticket	Cris	70	
		16 26/02/2021-14:15	Fechar ticket	Cris	20	
...	...	...	...	...	...	

Caso -- Evento -- Trace



Extra !!!

## Informação básica e de interesse – log de eventos



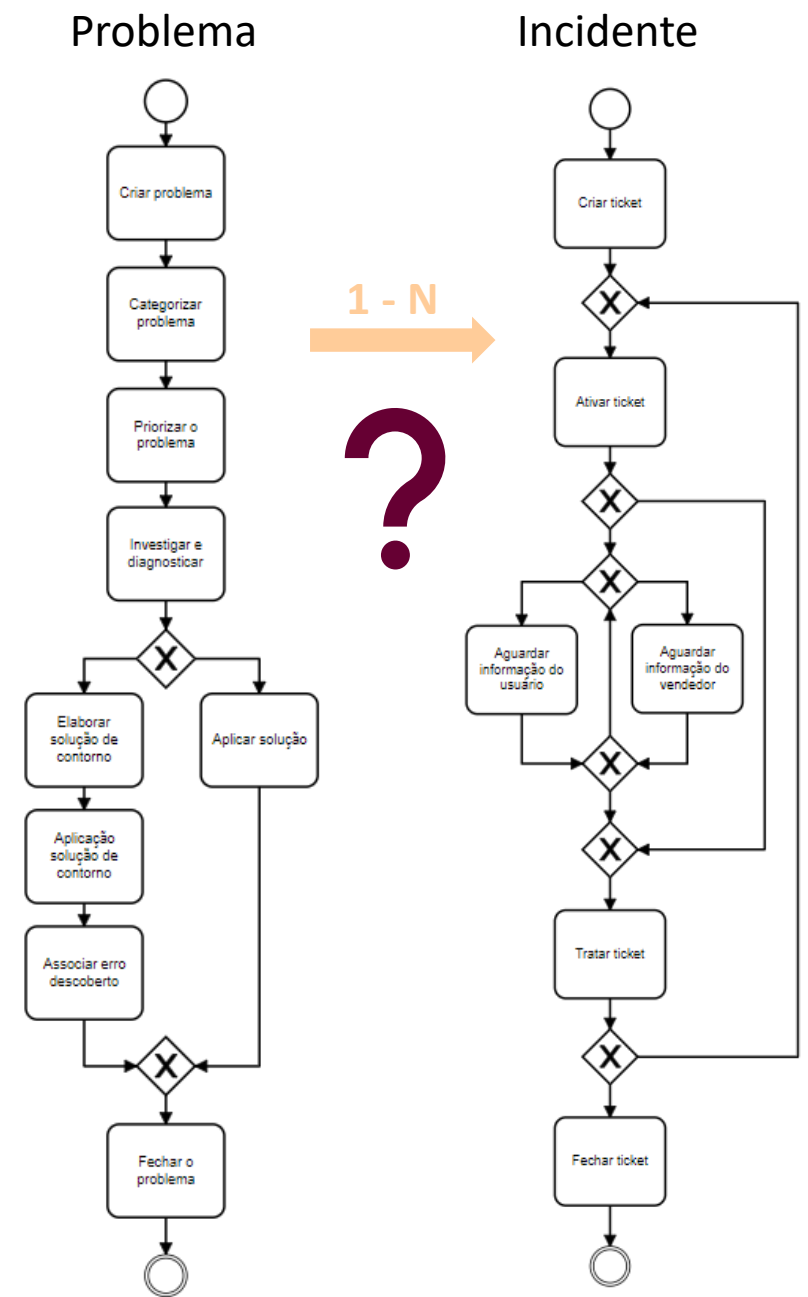
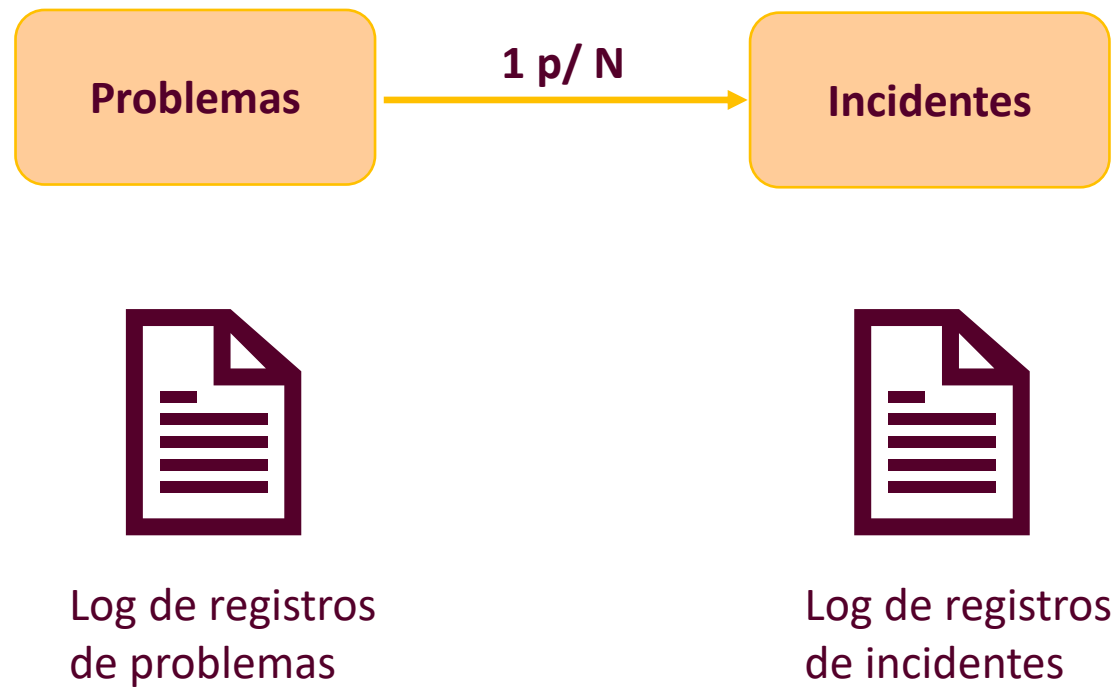
Log de registros  
de problemas



Log de registros  
de incidentes

Extra !!!

# Informação básica e de interesse – log de eventos





Extra !!!

## Informação básica e de interesse – log de eventos centrados em objetos

Estado da arte

Atividade	Timestamp	Problema	Incidente	...
...	...	...	...	...
Criar incidente	20/02/2021-08:01	$\phi$	{10}	...
Ativar incidente	20/02/2021-10:15	$\phi$	{10}	...
Tratar incidente	20/02/2021-10:26	$\phi$	{10}	...
Criar incidente	20/02/2021-10:27	$\phi$	{11}	...
Ativar incidente	20/02/2021-11:00	$\phi$	{11}	...
Criar problema	21/02/2021-08:55	{101}	{10,11}	...
Categorizar problema	21/02/2021-08:56	{101}	$\phi$	...
Priorizar problema	21/02/2021-08:57	{101}	$\phi$	...
Investigar e diagnosticar problema	21/02/2021-08:58	{101}	$\phi$	...
Aplicar solução ao problema	21/02/2021-14:01	{101}	$\phi$	...
Criar incidente	21/02/2021-15:10	$\phi$	{20}	...
Ativar incidente	21/02/2021-15:11	$\phi$	{20}	...
Criar problema	21/02/2021-15:12	{202}	{20}	...
Categorizar problema	21/02/2021-15:13	{202}	$\phi$	...
Fechar incidente	21/02/2021-15:14		{10}	...
Fechar problema	21/02/2021-15:15	{101}	$\phi$	...
Priorizar problema	21/02/2021-15:16	{202}	$\phi$	...
Tratar incidente	22/02/2021-09:35	$\phi$	{20}	...
Priorizar problema	22/02/2021-8:01	{202}	$\phi$	...
Investigar e diagnosticar problema	22/02/2021-8:02	{202}	$\phi$	...
Aplicar solução ao problema	22/02/2021-15:05	{202}	$\phi$	...
Tratar incidente	22/02/2021-15:30	$\phi$	{11}	...
Fechar incidente	22/02/2021-15:31	$\phi$	{11}	...
...	...	...	...	...

## Formatos – os logs e as ferramentas

- 2003-2010: **MXML** (Mining XML)
  - Várias extensões *ad hoc* foram desenvolvidas naturalmente para MXML.
  - Exemplo: extensão para “anotação de semântica”.
- 2010: **XES** (eXtensible Event Stream)
  - <http://www.xes-standard.org/>
  - Padrão oficial IEEE: <https://standards.ieee.org/standard/1849-2016.html>



# Formatos – os logs e as ferramentas

## Formatos



- .mxml, .xml
- .csv , .xlsx, ...
- .json
- conectores  
(SAP, ServiceNow ...)

## Acadêmicas



## Industriais



Indústria nacional

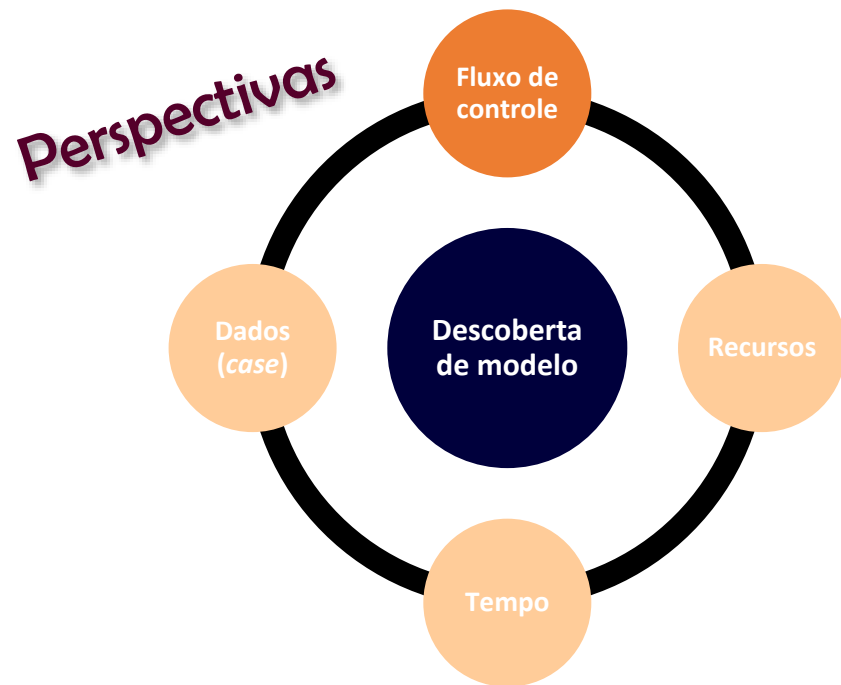
# Tipos de mineração de processos:

## Descoberta de modelos de processo



# Descoberta de modelos de processo

- Construção de um modelo de processo a partir de um log de eventos.
  - Captura o comportamento que existe no log de eventos (**play-in**)



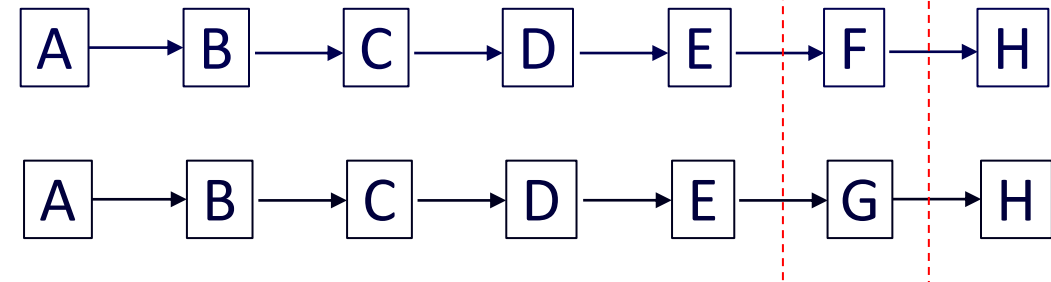
## Log de eventos simples

< A, B, C, D, E, F, H >

< A, B, C, D, E, G, H >

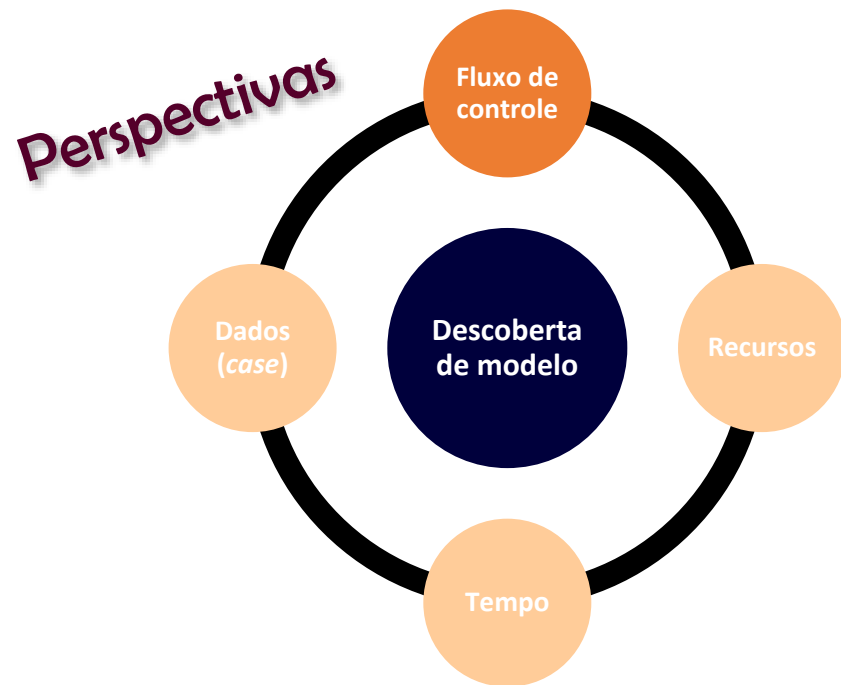
< A, C, B, D, E, F, H >

< A, C, B, D, E, G, H >



# Descoberta de modelos de processo

- Construção de um modelo de processo a partir de um log de eventos.
  - Captura o comportamento que existe no log de eventos (play-in)



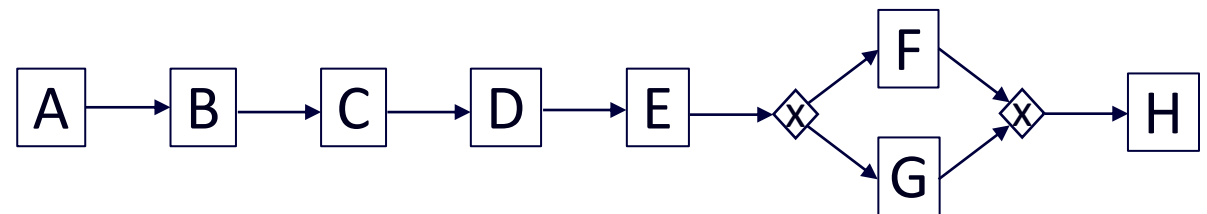
## Log de eventos simples

< A, B, C, D, E, F, H >

< A, B, C, D, E, G, H >

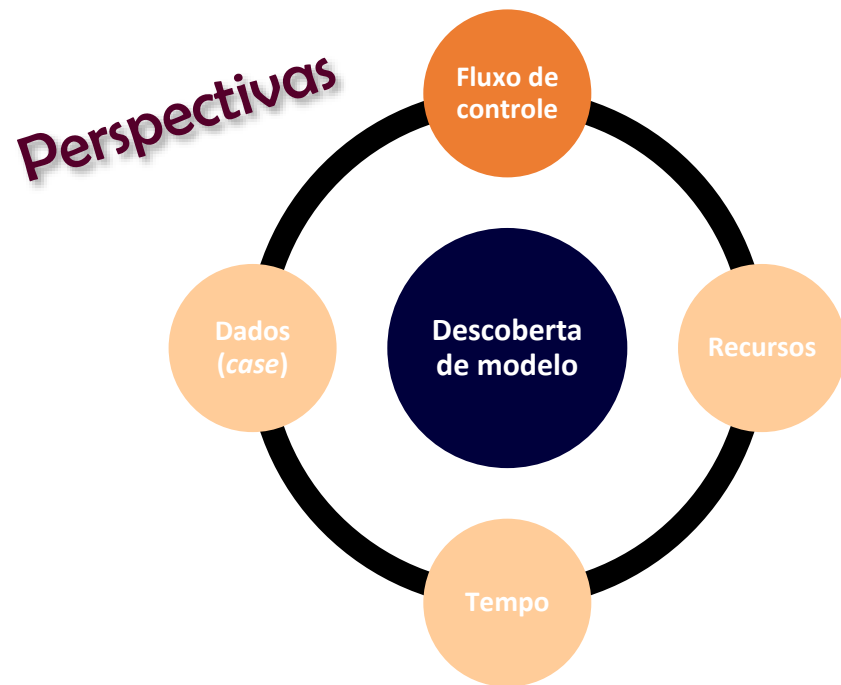
< A, C, B, D, E, F, H >

< A, C, B, D, E, G, H >



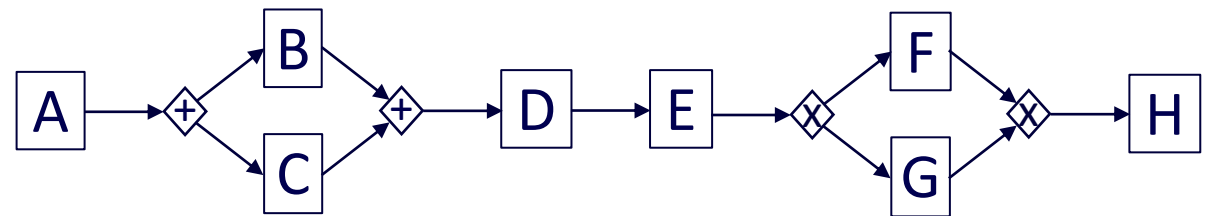
# Descoberta de modelos de processo

- Construção de um modelo de processo a partir de um log de eventos.
  - Captura o comportamento que existe no log de eventos (**play-in**)



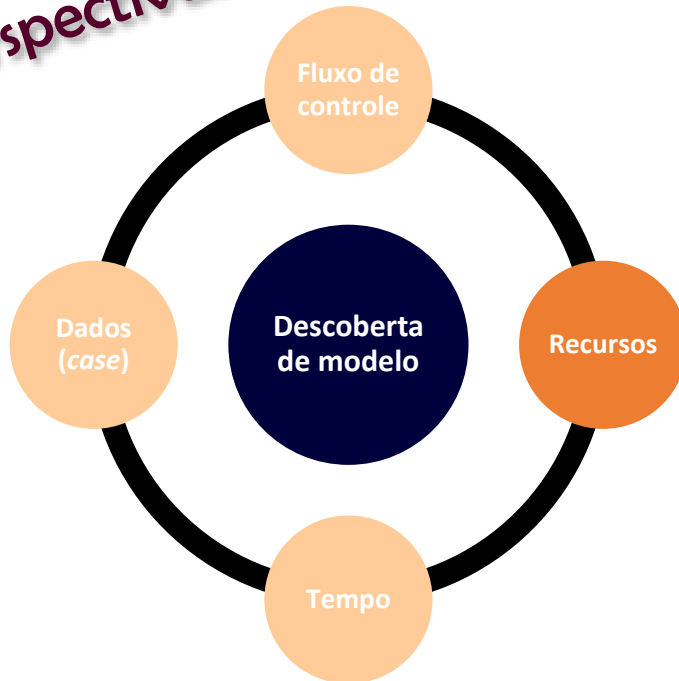
## Log de eventos simples

< A, B, C, D, E, F, H >  
< A, B, C, D, E, G, H >  
< A, C, B, D, E, F, H >  
< A, C, B, D, E, G, H >

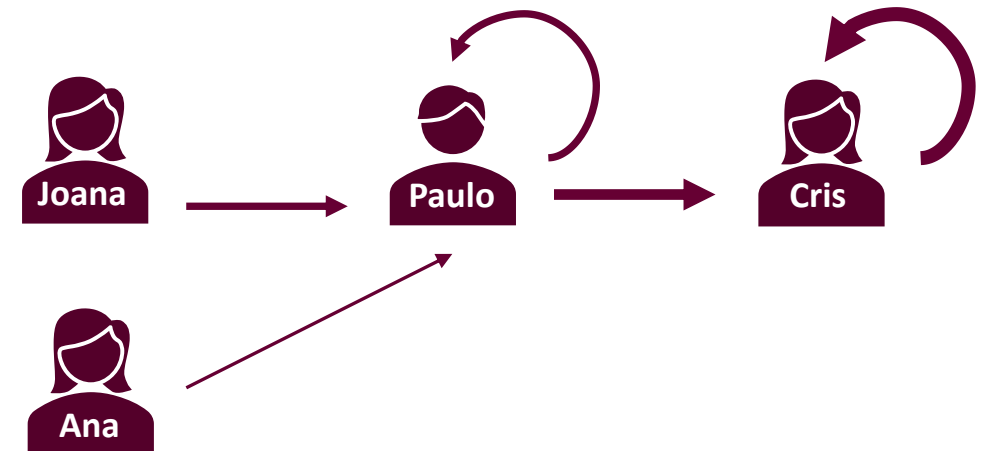


# Descoberta de modelos de processo

Perspectivas



Id. Caso	Id. Evento	Propriedades				
		Timestamp	Atividade	Recurso	Custo	...
1	1	20/02/2021-10:30	Criar ticket	Joana	10	
	2	20/02/2021-10:35	Ativar ticket	Paulo	50	
	3	20/02/2021-11:10	Aguardar informação do usuário	Cris	100	
	4	26/02/2021-15:00	Tratar ticket	Cris	50	
	5	26/02/2021-18:00	Fechar ticket	Cris	10	
2	6	23/02/2021-16:00	Criar ticket	Joana	10	
	7	23/02/2021-16:10	Ativar ticket	Paulo	50	
	8	23/02/2021-16:12	Tratar ticket	Paulo	50	
	9	23/02/2021-16:48	Fechar ticket	Paulo	10	
3	10	20/02/2021-10:35	Criar ticket	Ana	10	
	11	20/02/2021-10:40	Ativar ticket	Paulo	50	
	12	20/02/2021-15:50	Aguardar informação do usuário	Cris	100	
	13	25/02/2021-11:40	Tratar ticket	Cris	50	
	14	25/02/2021-12:40	Ativar ticket	Cris	70	
	15	26/02/2021-08:00	Tratar ticket	Cris	70	
	16	26/02/2021-14:15	Fechar ticket	Cris	20	
...	...	...	...	...	...	

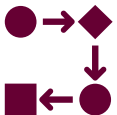




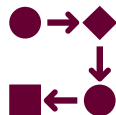
Extra !!!

## Descoberta de modelos de processo

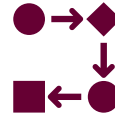
**Fitness**



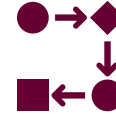
**Precisão**



**Generalização**



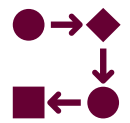
**Simplicidade**



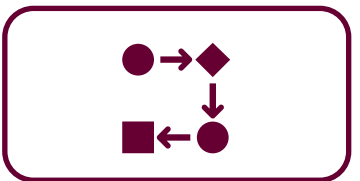
Extra !!!

# Descoberta de modelos de processo

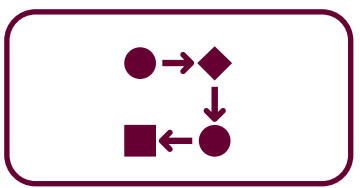
**Fitness**



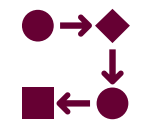
**Precisão**



**Generalização**



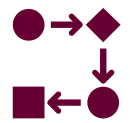
**Simplicidade**



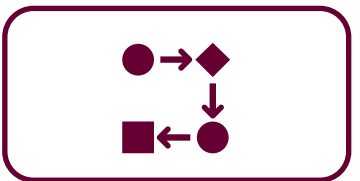
Extra !!!

# Descoberta de modelos de processo

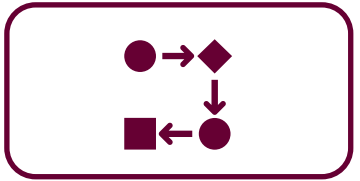
**Fitness**



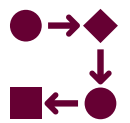
**Precisão**



**Generalização**

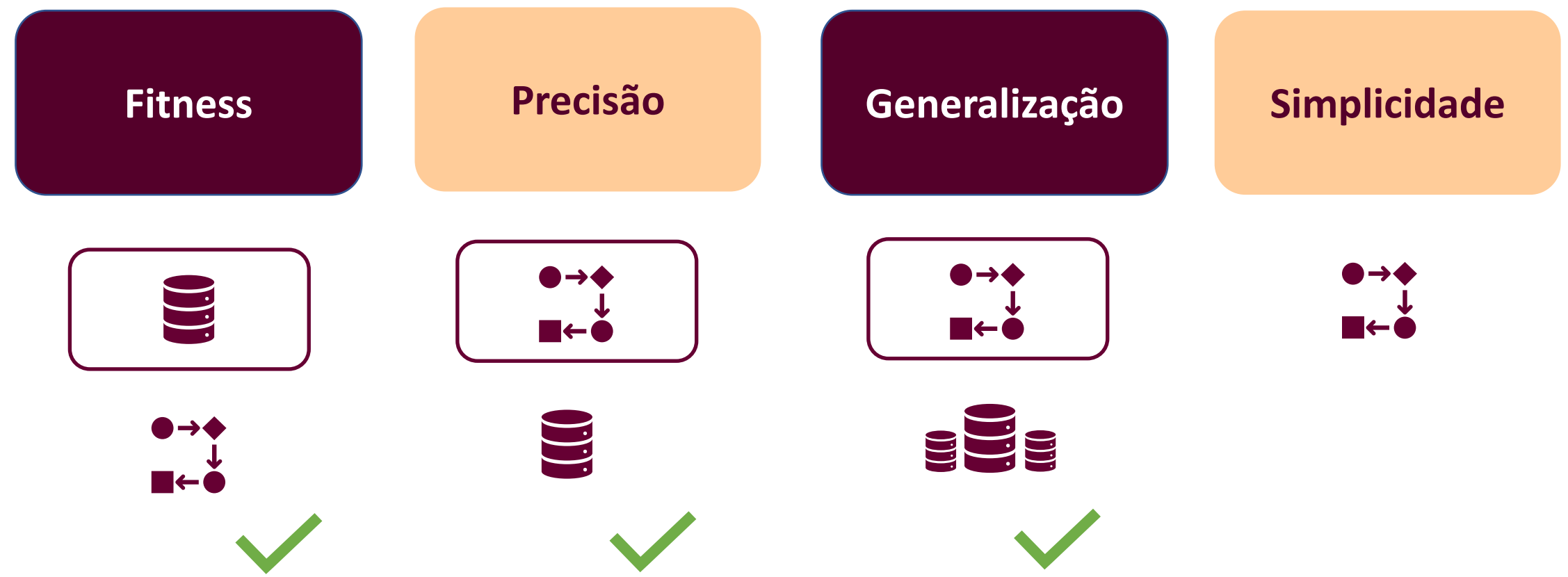


**Simplicidade**



Extra !!!

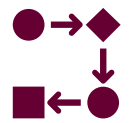
# Descoberta de modelos de processo



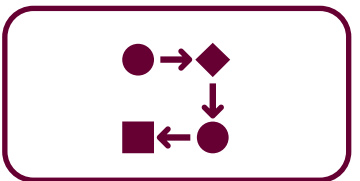
Extra !!!

# Descoberta de modelos de processo

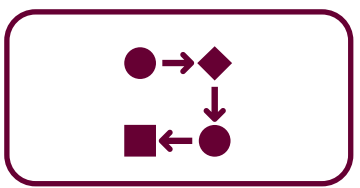
**Fitness**



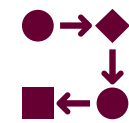
**Precisão**



**Generalização**



**Simplicidade**



# Descoberta de modelos de processo

## Algoritmos

Alpha-algorithm

Heuristic Miner

Fuzzy Miner

Genetic Miner

Inductive Miner

Split Miner

## Representação

Redes de Petri

Causal nets

Direct flow graph

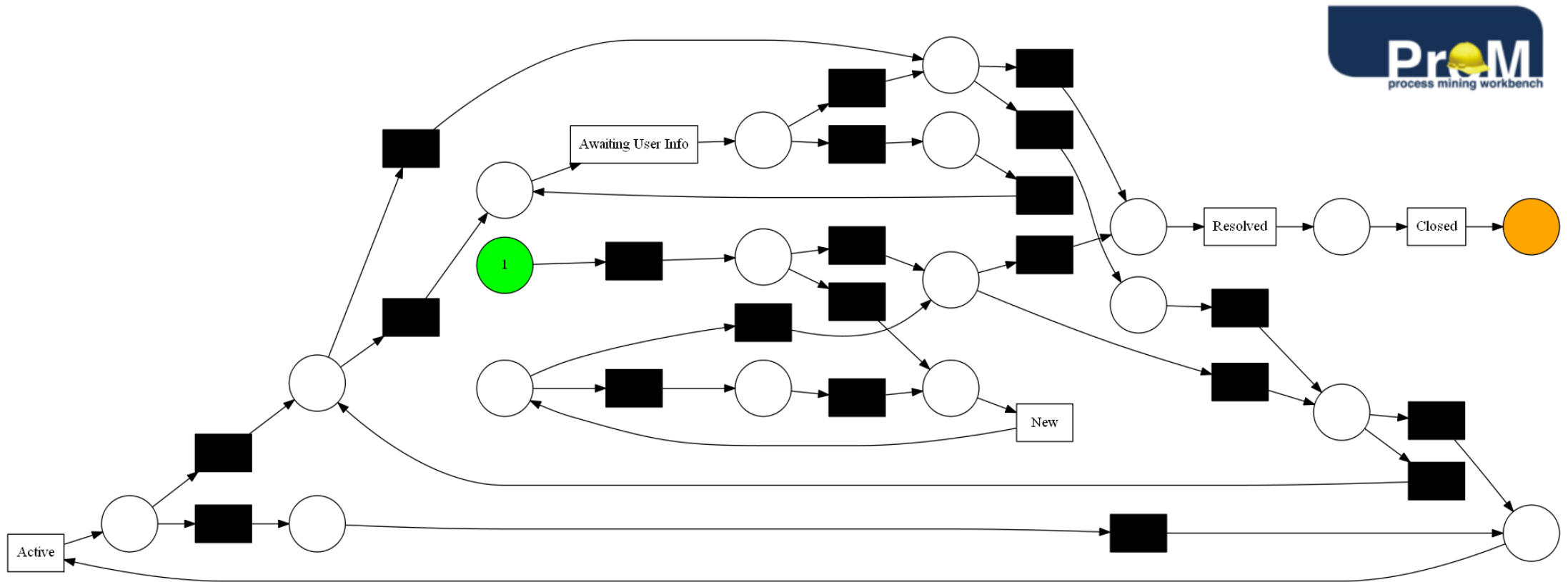
Process tree

BPMN

# Descobertas nas ferramentas

Em geral, as ferramentas acadêmicas apresentam os modelos de processo em diferentes notações.

- A notação usada está associada ao algoritmo de descoberta usado.

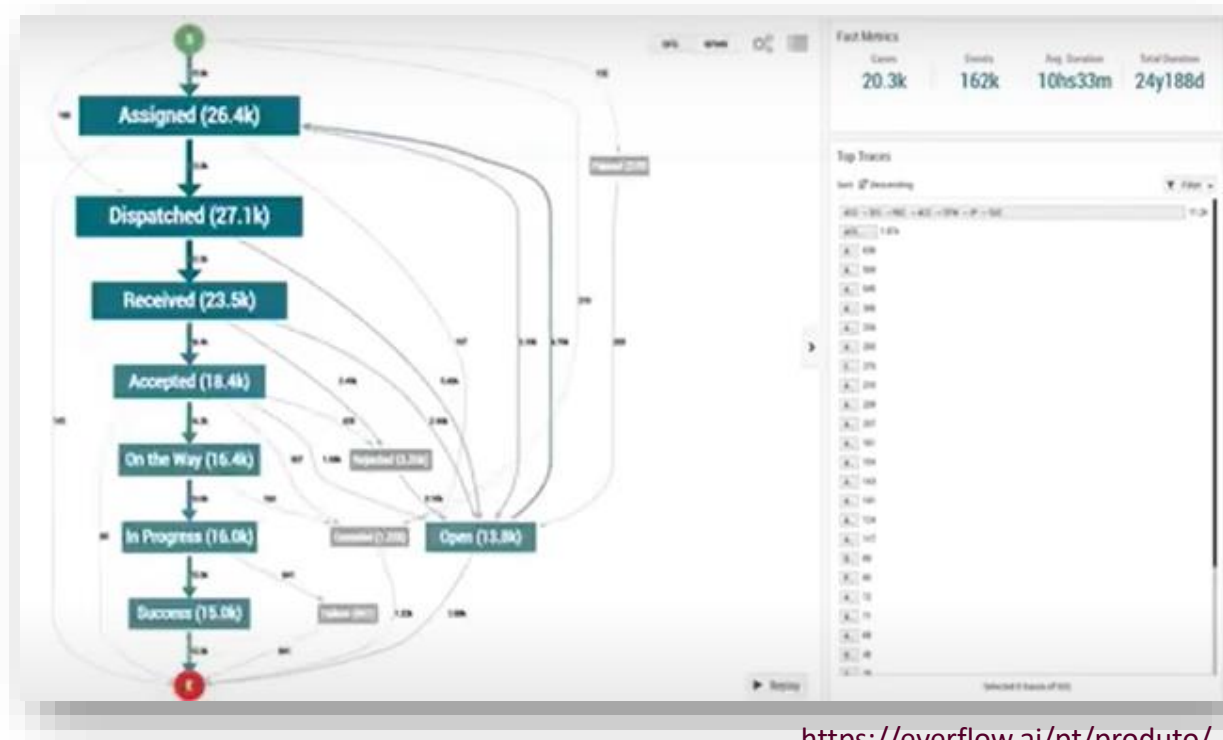


# Descobertas nas ferramentas

Em geral, as ferramentas industriais usam a representação com sistemas de transições, que possui expressividade limitada.

- Isso é um problema????

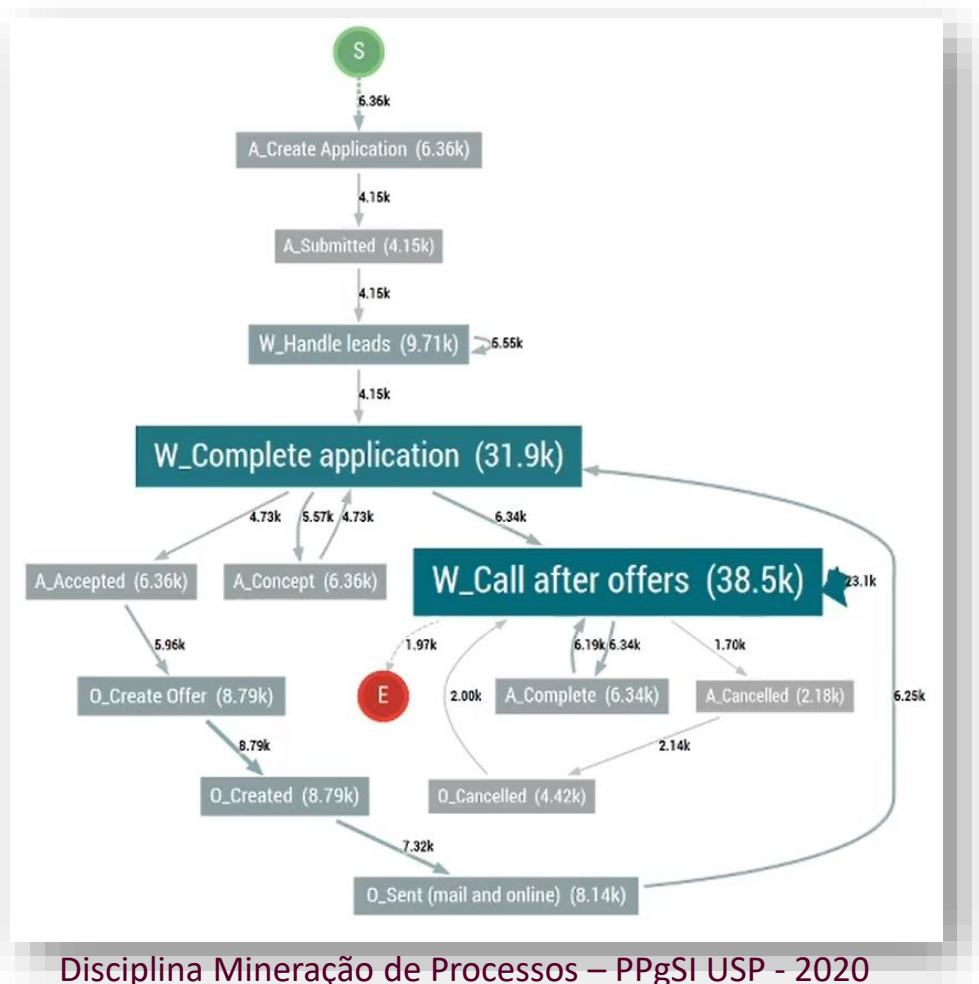
Mapeamentos para modelos BPMN também estão disponíveis.



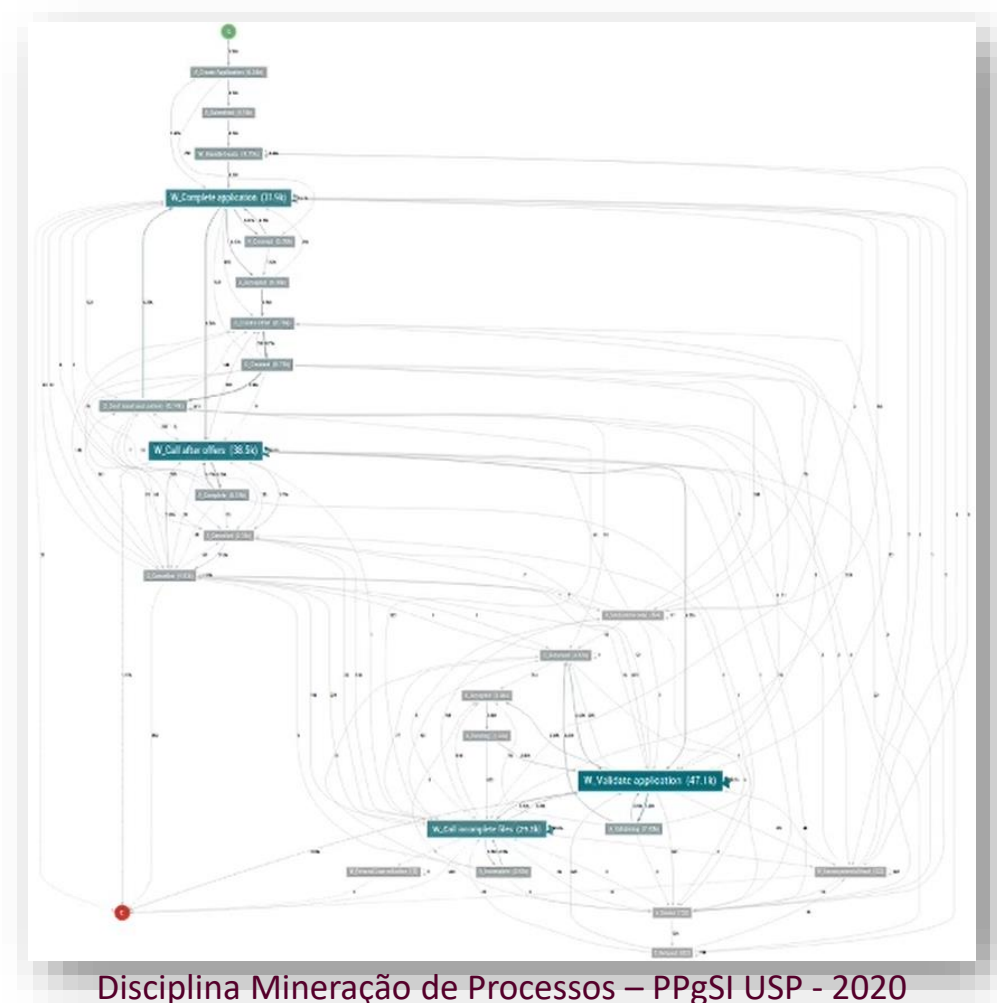
<https://everflow.ai/pt/produto/>



# Descobertas nas ferramentas

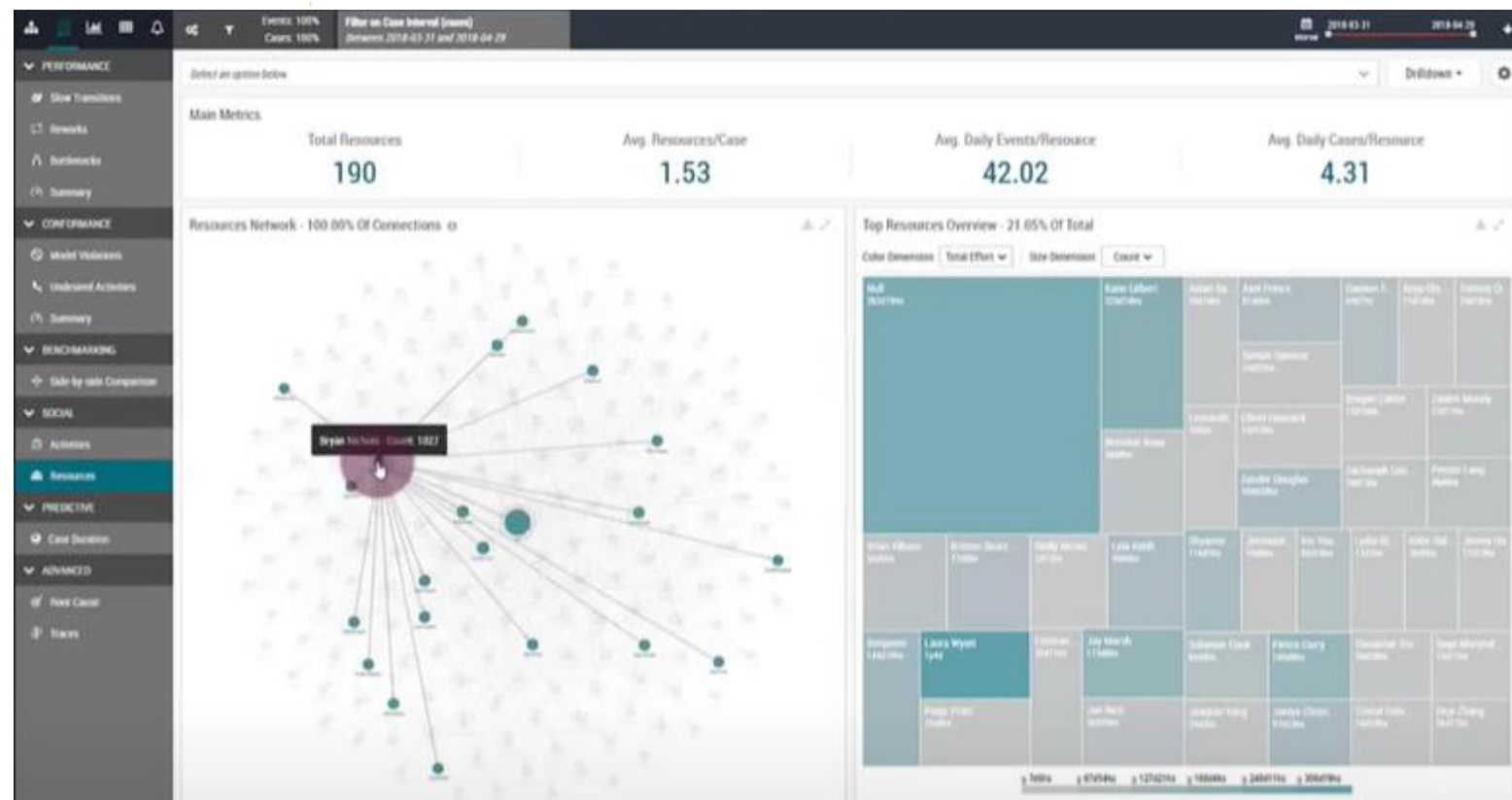
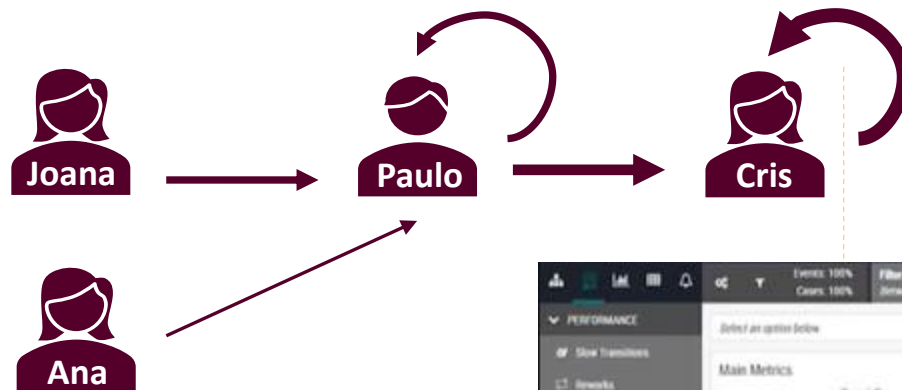


**Modelo lasanha**



**Modelo espaguete**

# Descobertas nas ferramentas



<https://everflow.ai/pt/produto/>

Extra !!!

## Descoberta de modelos de processo

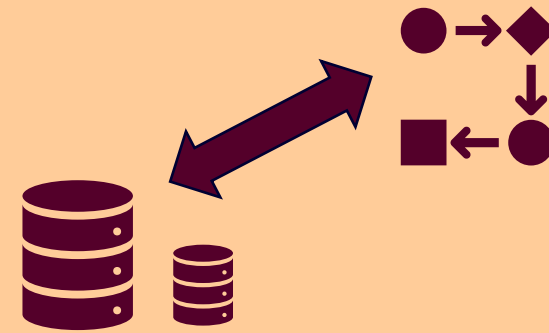
O problema de incompletude do log:

- Se há loops  
não teremos todos os *traces* possíveis;
- Se há concorrência (AND)  
não teremos todas as possíveis ordenações;
- Se há múltiplas escolhas (XOR)  
não veremos todas as combinações.

**Não existe almoço grátis!**

# Tipos de mineração de processos:

## Checagem de conformidade



# Checagem de conformidade

## Contexto:

- temos o modelo de processo (descoberto - descrito ou construído à mão - normativo)
- temos o log de eventos.

- **Oportunidade:**

- relacionar eventos do log de eventos com atividades no modelo de processo
- compará-los.

- **Meta:**

- encontrar coisas em comum e discrepâncias entre o comportamento modelado (no modelo) e o comportamento observado (no log).

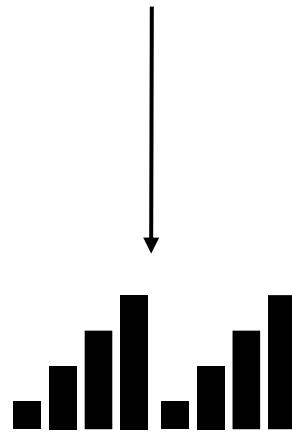
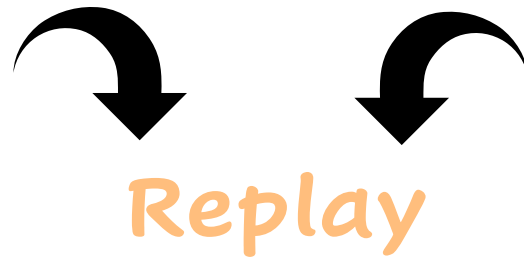
## Oportunidades

Auditoria

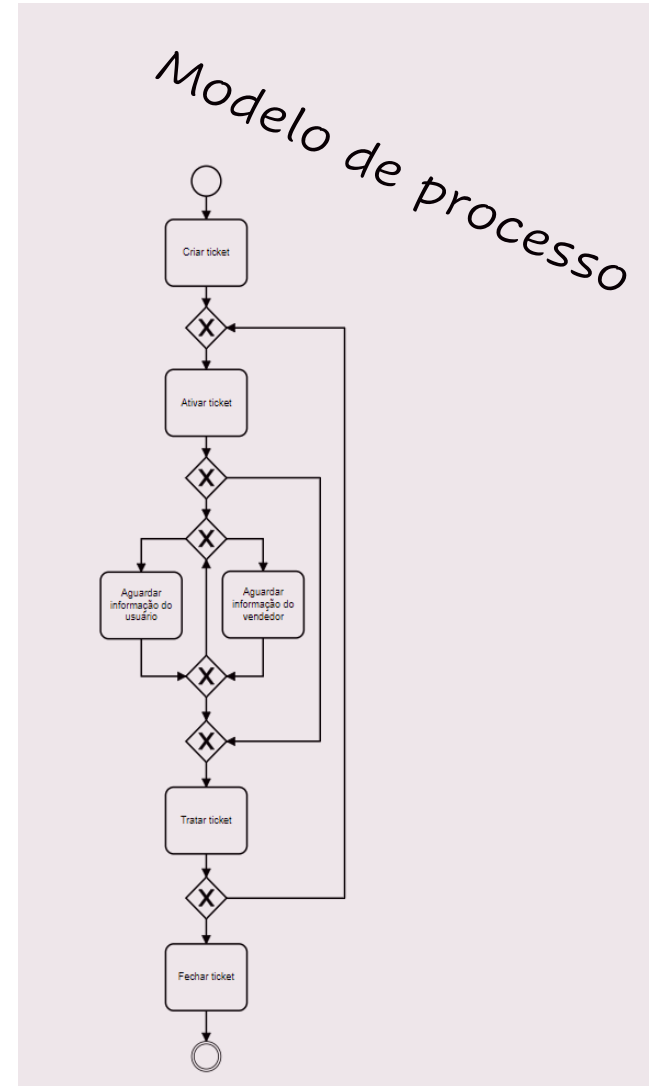
Alinhamento de negócios

Avaliação de algoritmos

# Checagem de conformidade



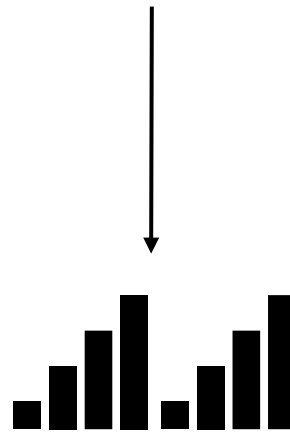
Medidas de conformidade global



# Checagem de conformidade

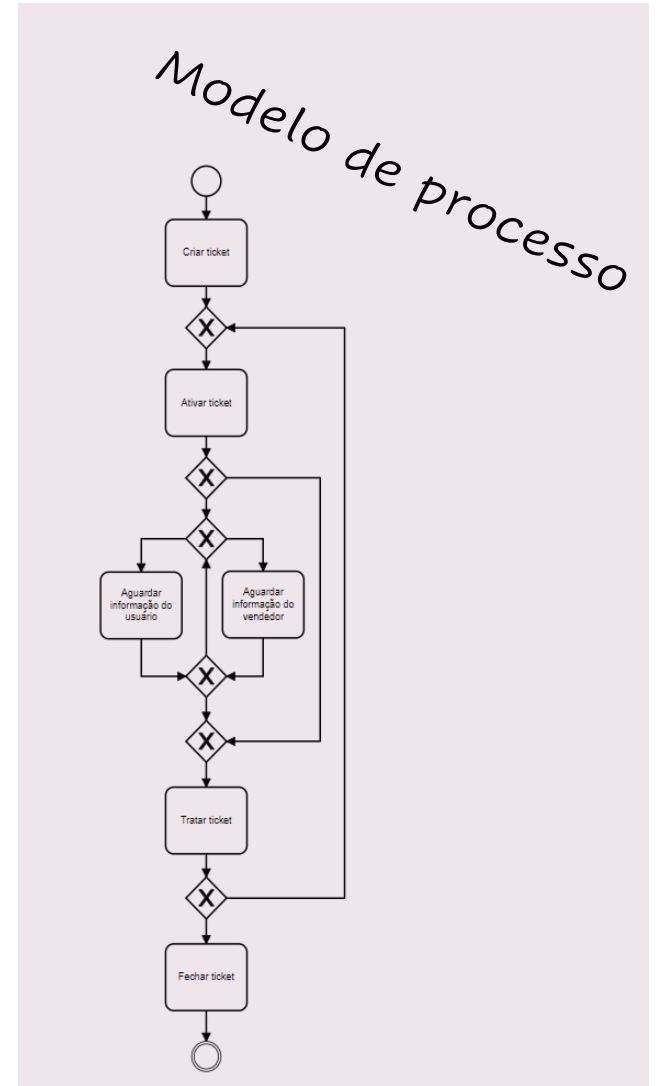


Replay



Medidas de conformidade global

85% dos casos no log podem  
ser reproduzidos pelo modelo



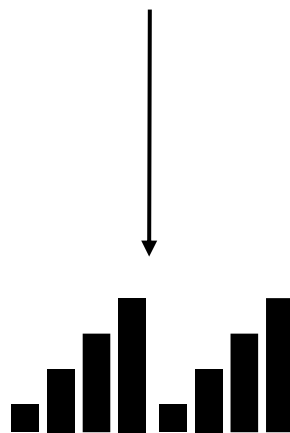
# Checagem de conformidade



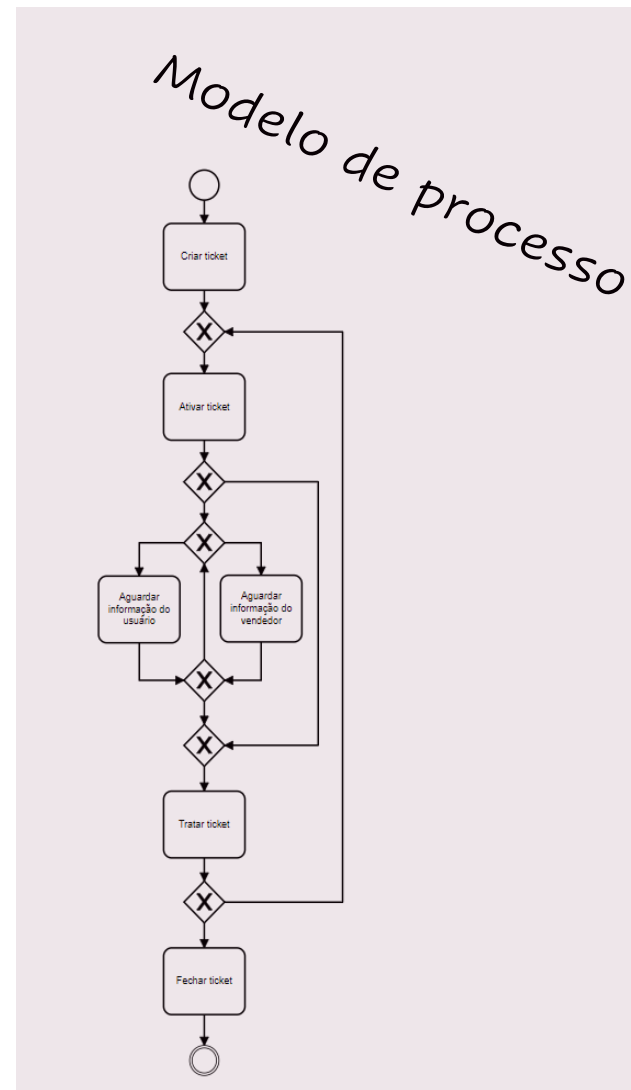
85% dos casos no log podem ser reproduzidos pelo modelo.



Replay



Medidas de conformidade global



A atividade X foi executada 15 vezes, embora isso não fosse permitido de acordo com o modelo.



# Checagem de conformidade

Imagine!!!!

O *trace*: <a1, a2, a3, ..., a100>

Modelo A

Consegue reproduzir 99 das  
100 atividades

Modelo B

Consegue reproduzir 10 das  
100 atividades

Em uma métrica *naïve*, o *trace* seria considerado *non-fitting* em ambos os modelos.

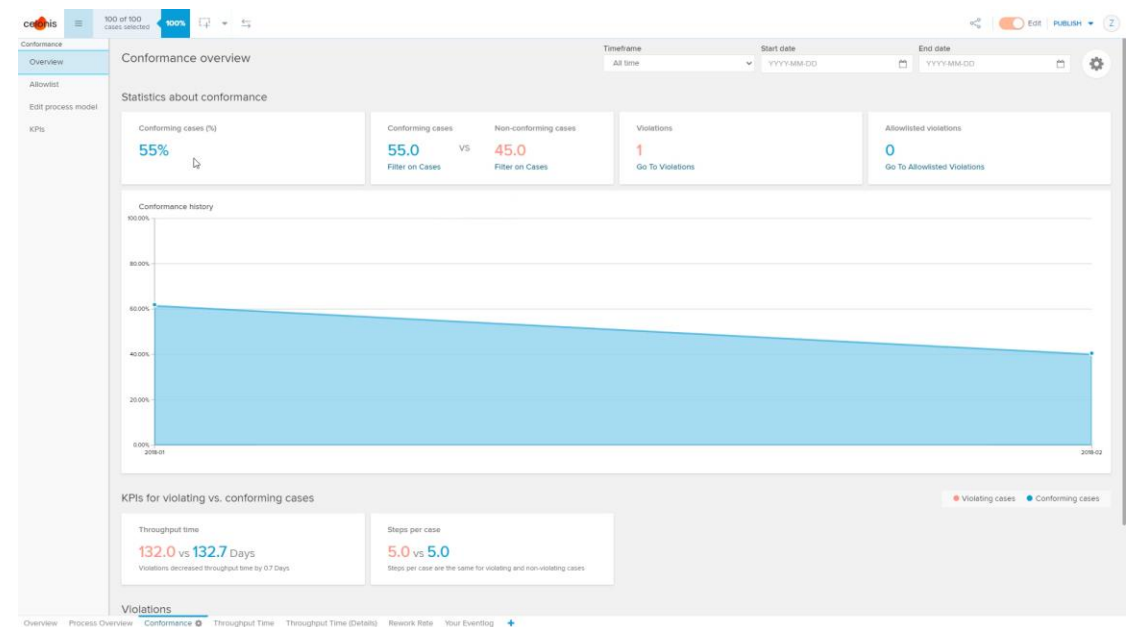
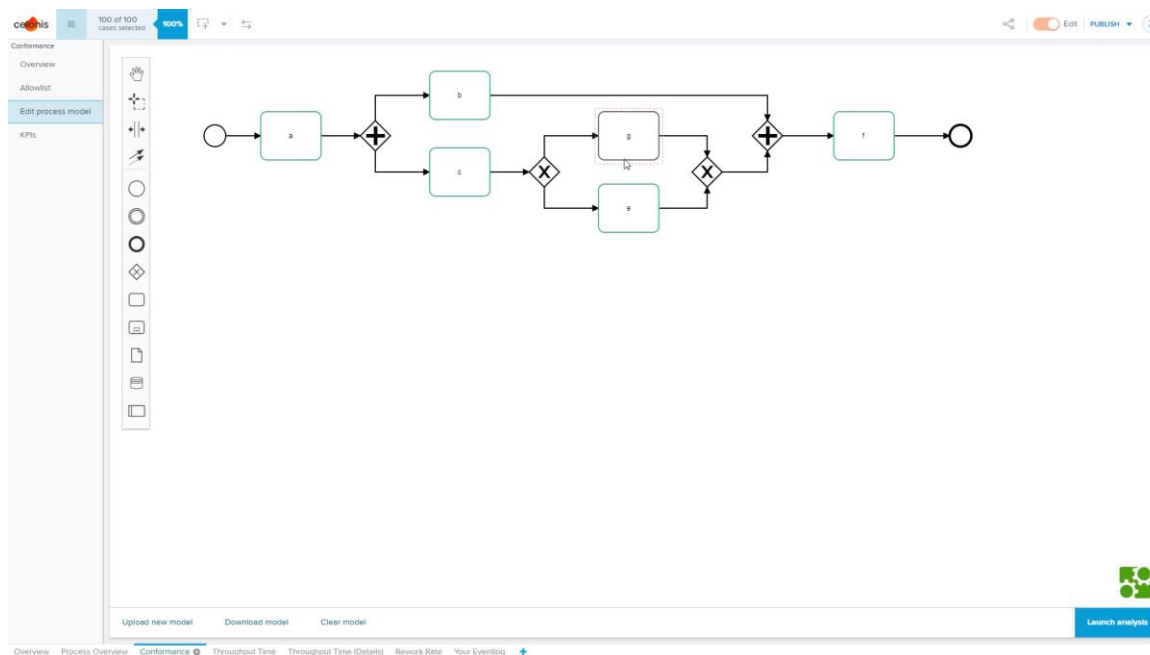
Precisamos de uma noção de *fitness* definida **no nível dos eventos** em vez de ser definida sobre *traces* completos.

Token replay

Alinhamento

Comparação de assinaturas

# Checagem de conformidade nas ferramentas

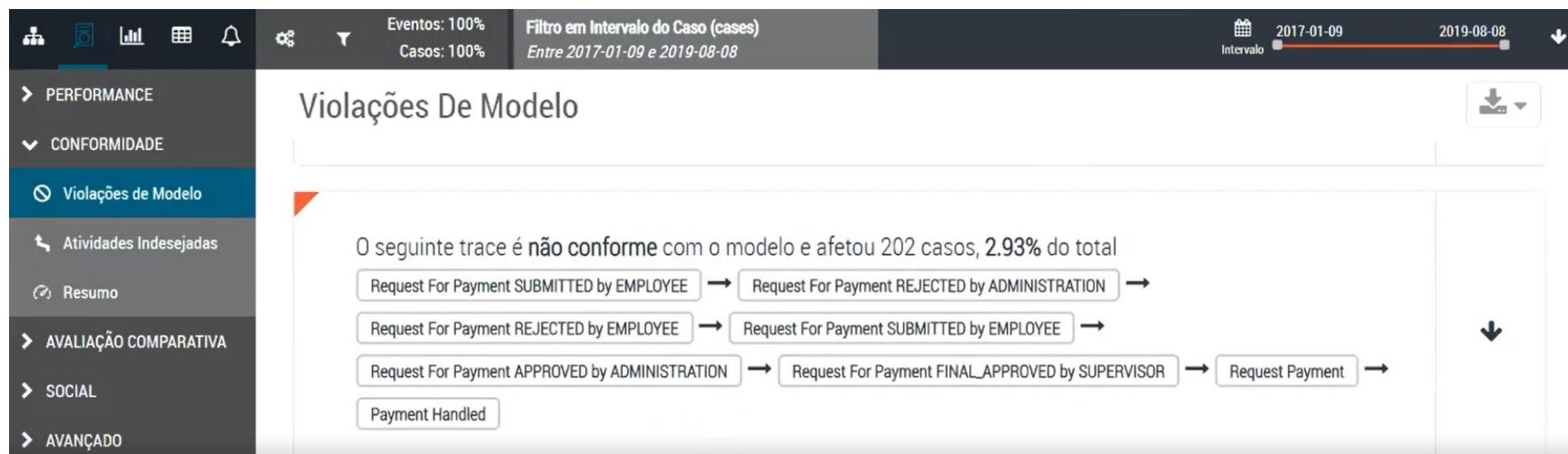


<http://processmining.each.webhostusp.sti.usp.br/index.php/teaching-research-resources/>



<https://www.celonis.com/snap-signup/>

# Checagem de conformidade nas ferramentas



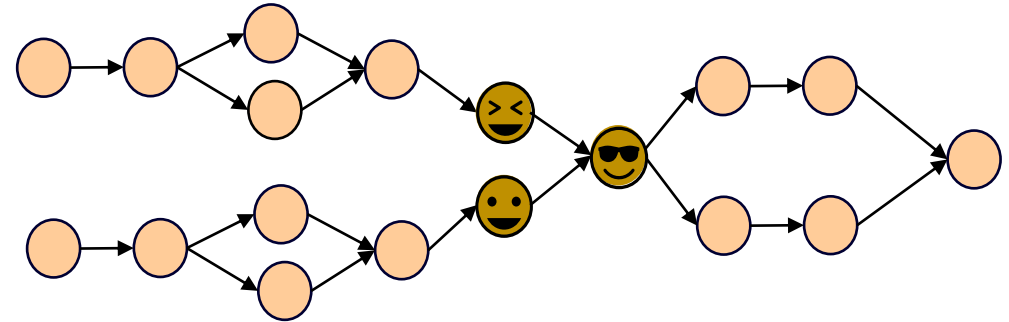
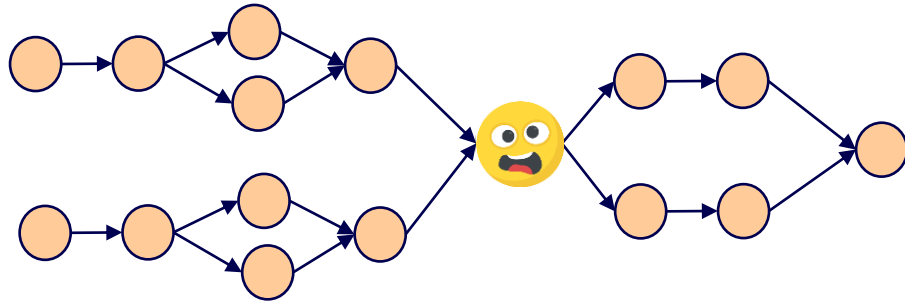
Disciplina Mineração de Processos – PPgSI USP - 2020

# Tipos de mineração de processos:

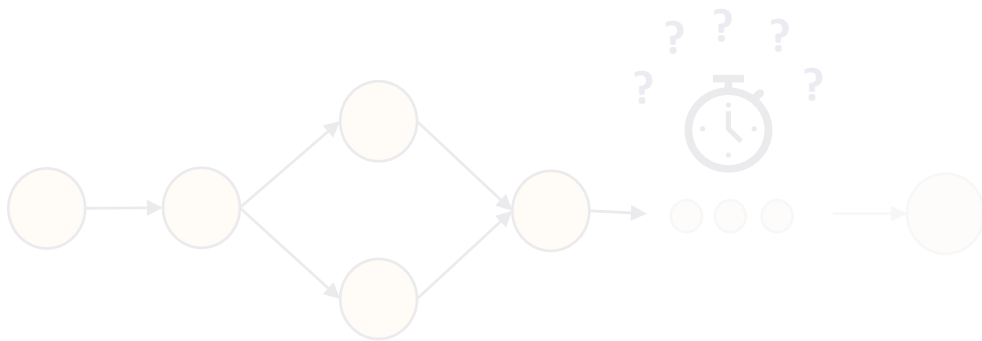
## Melhoria de processos



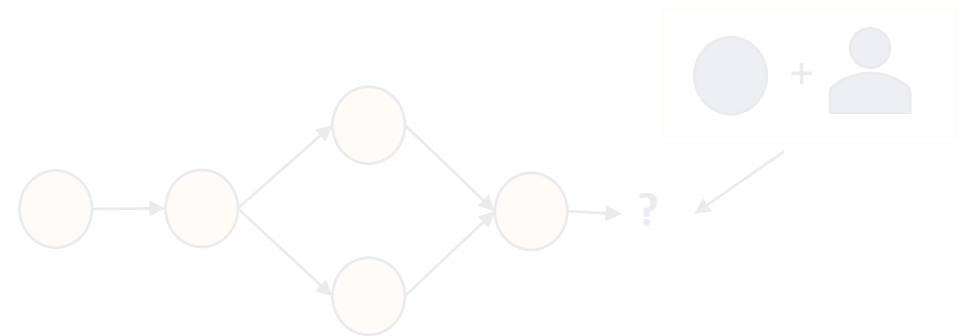
# Melhoria de processos



Analisar gargalos e variantes do processo e sugerir melhorias para o processo

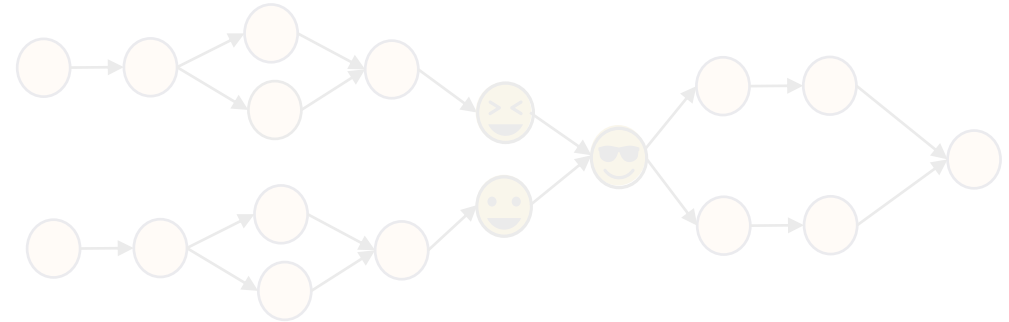
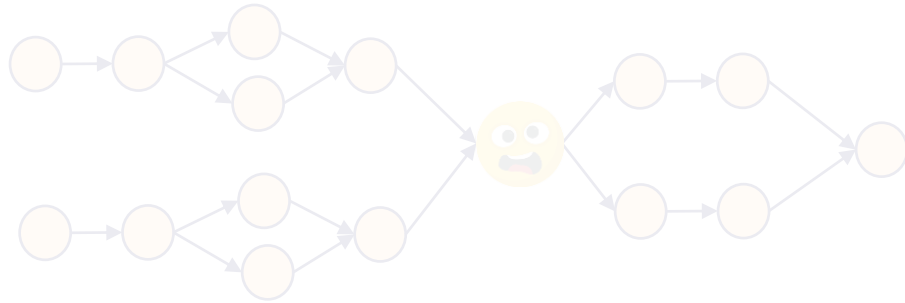


Predição de tempo de resolução de um caso

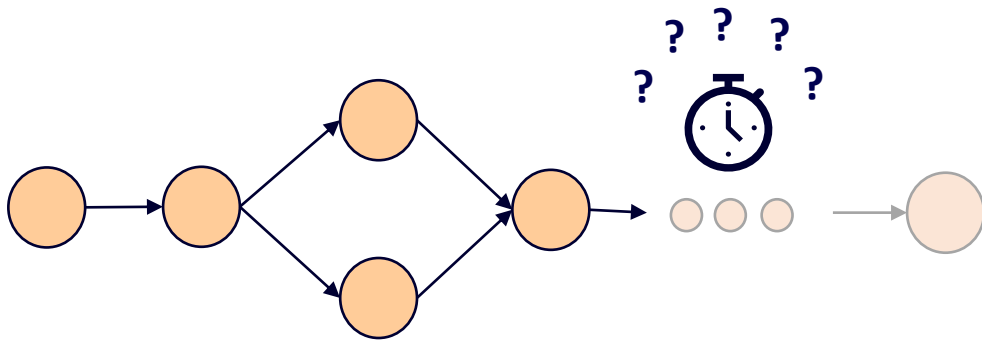


Predição da próxima atividade e recomendação do recurso

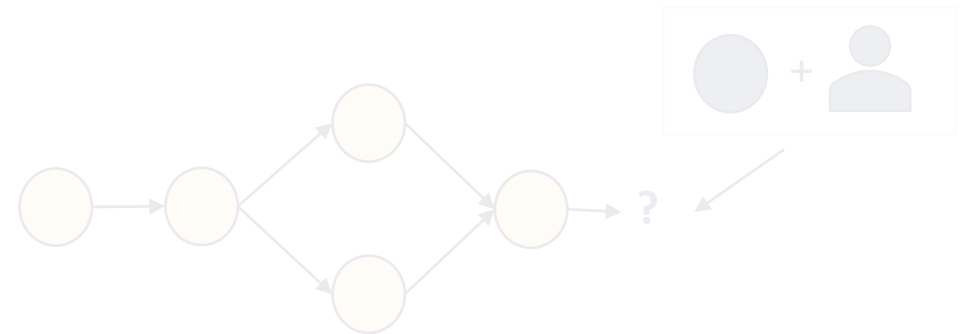
# Melhoria de processos



Analisar gargalos e variantes do processo e sugerir melhorias para o processo

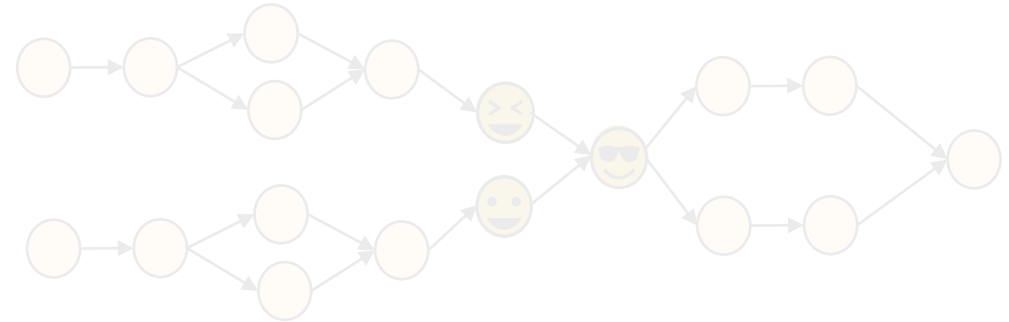
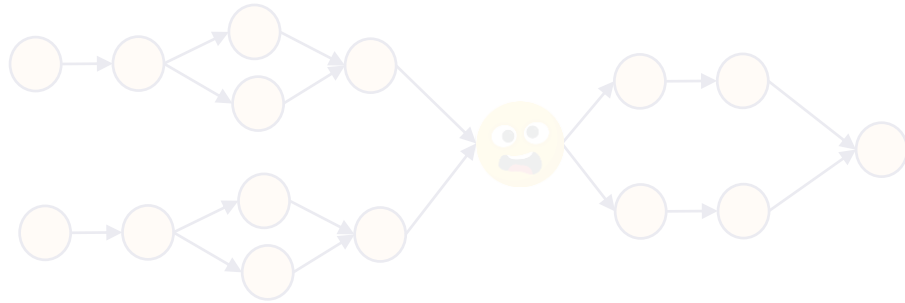


Predição de tempo de resolução de um caso

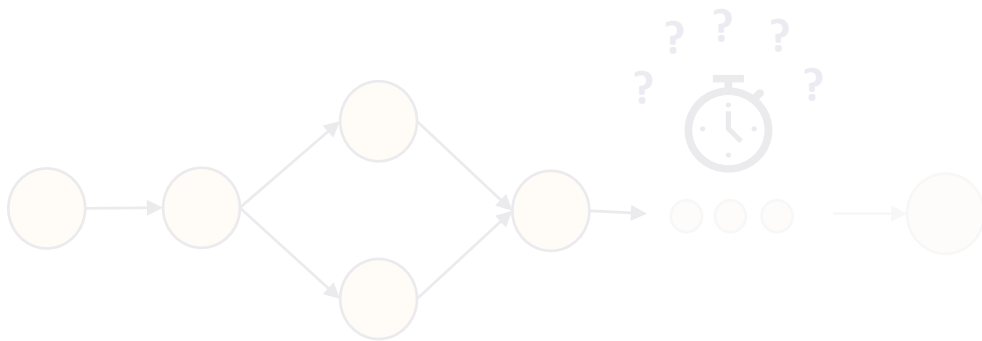


Predição da próxima atividade e recomendação do recurso

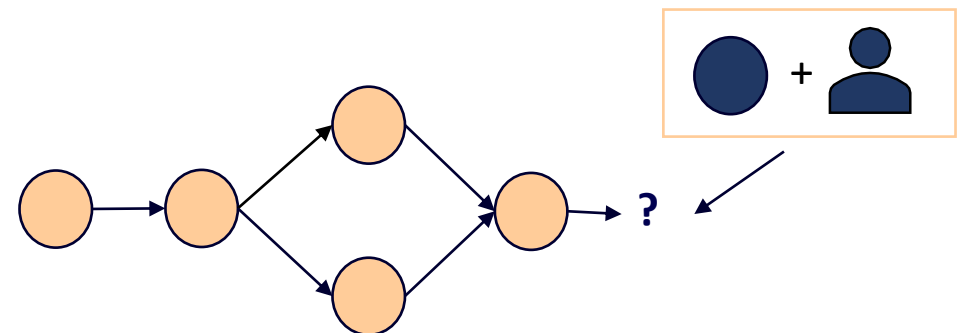
# Melhoria de processos



Analisar gargalos e variantes do processo e sugerir melhorias para o processo

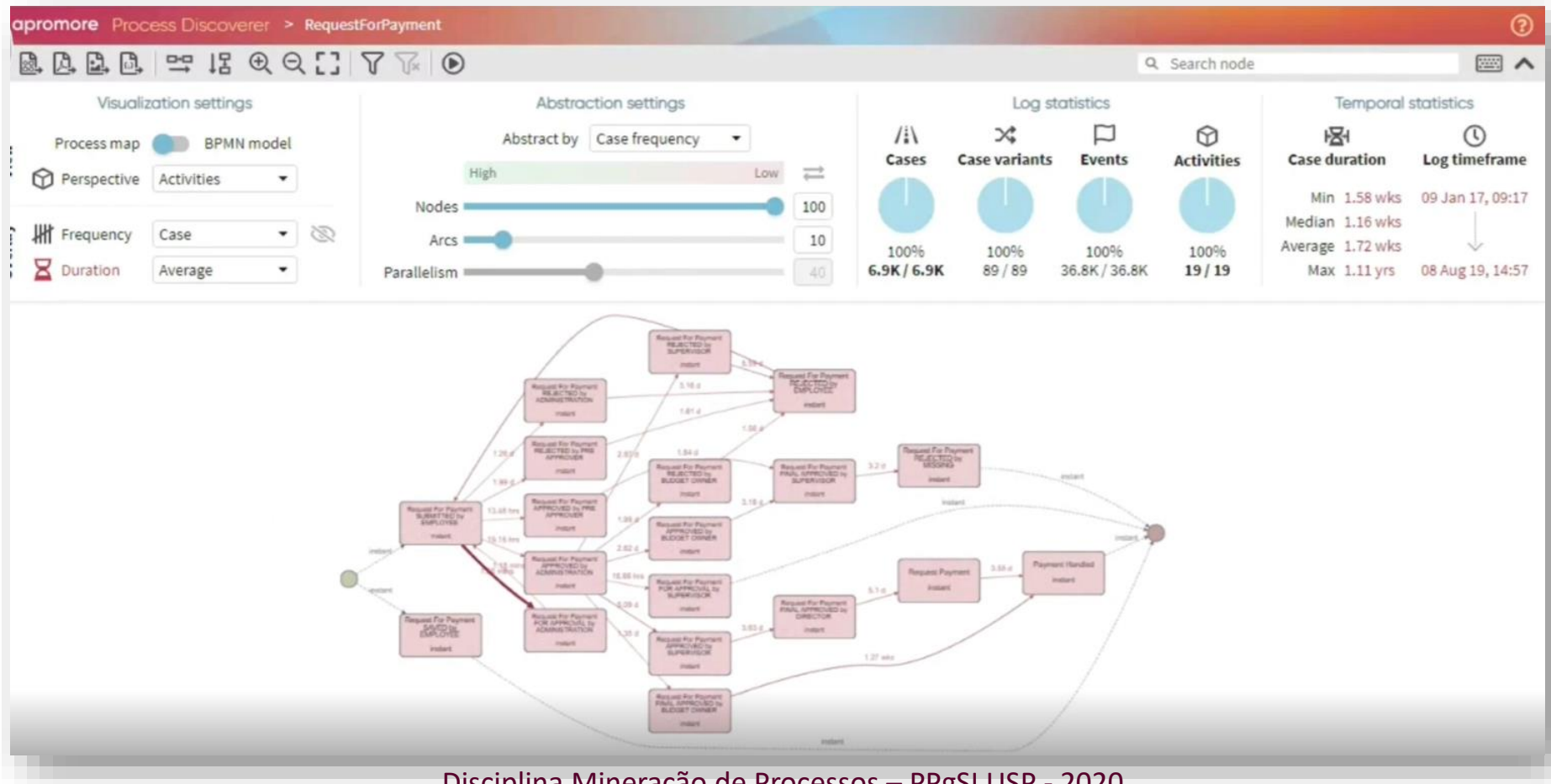


Predição de tempo de resolução de um caso



Predição da próxima atividade e recomendação do recurso

# Melhoria de processos



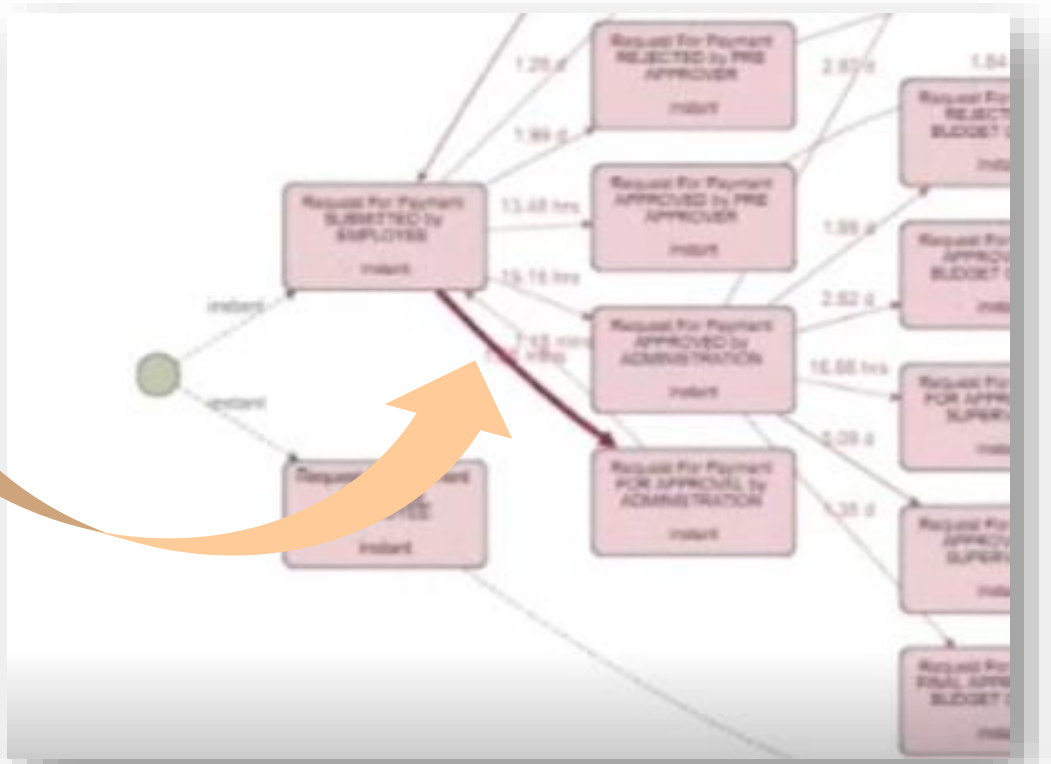
Disciplina Mineração de Processos – PPgSI USP - 2020



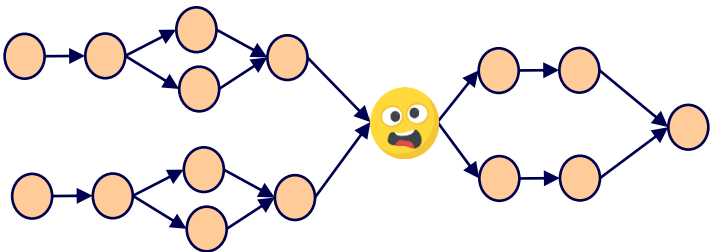
# Melhoria de processos



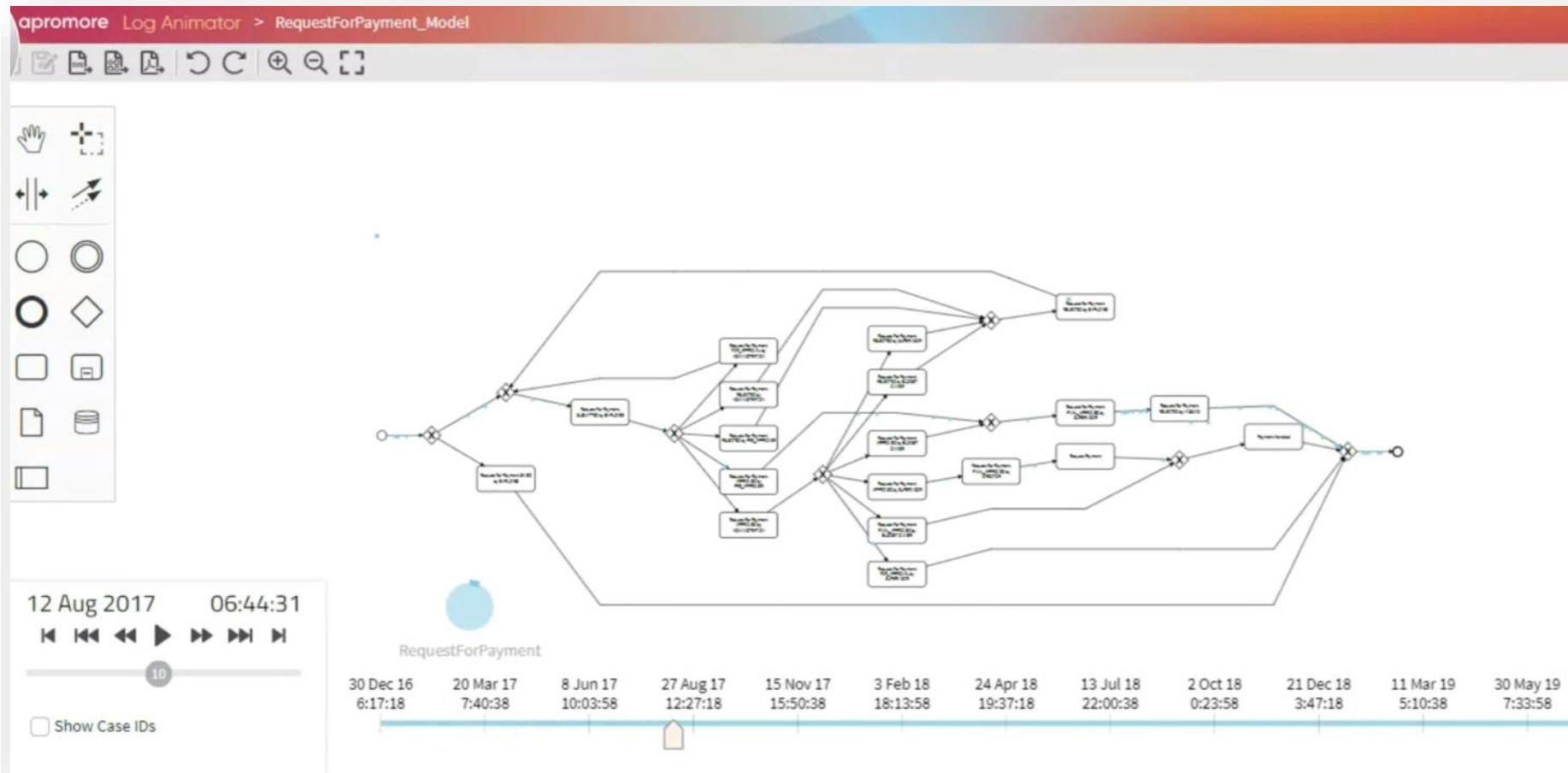
?



Disciplina Mineração de Processos – PPgSI USP - 2020

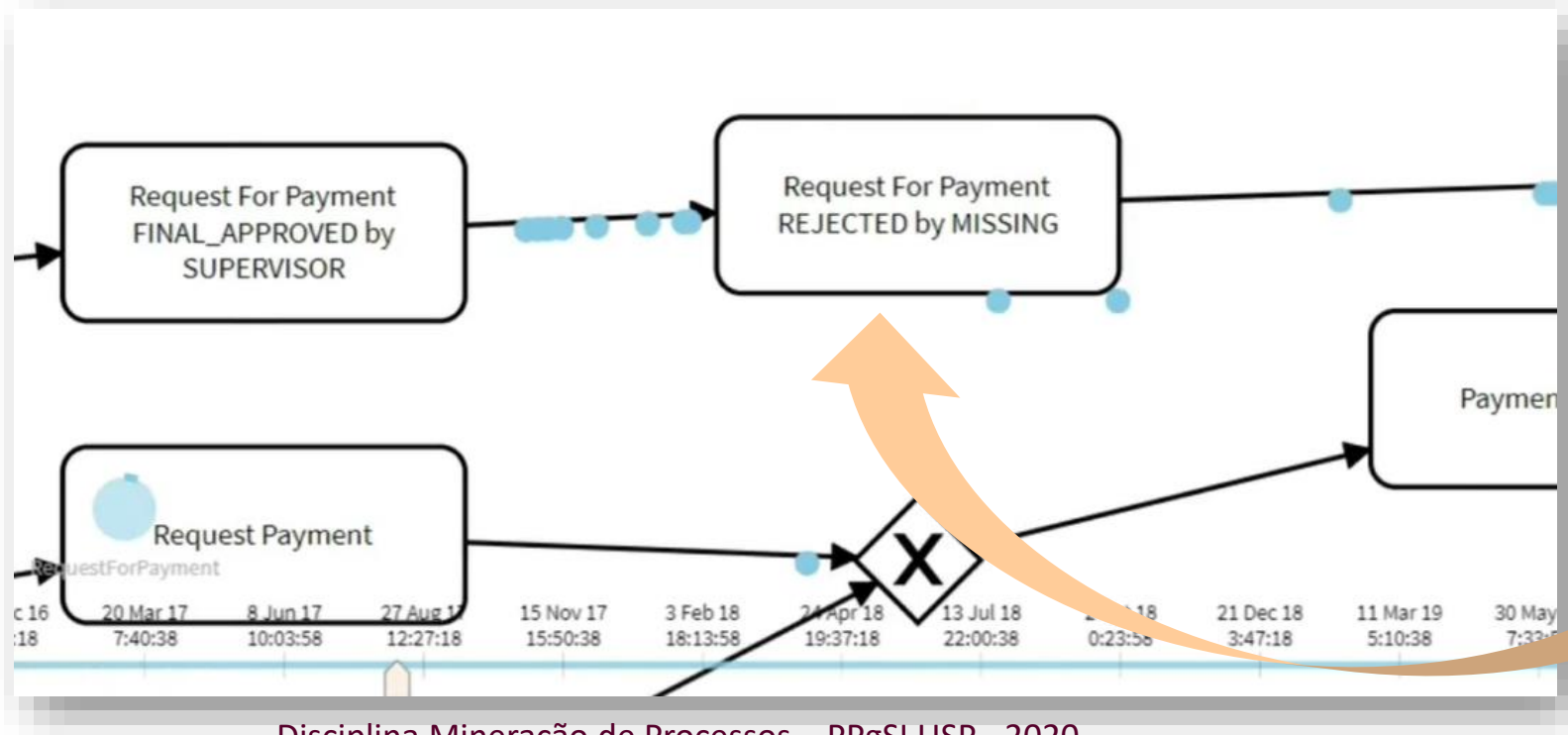


# Melhoria de processos



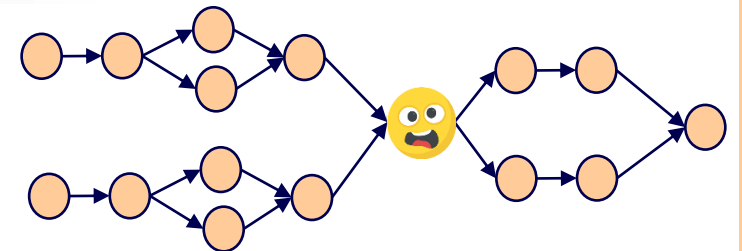
Disciplina Mineração de Processos – PPgSI USP - 2020

# Melhoria de processos



?

Disciplina Mineração de Processos – PPgSI USP - 2020



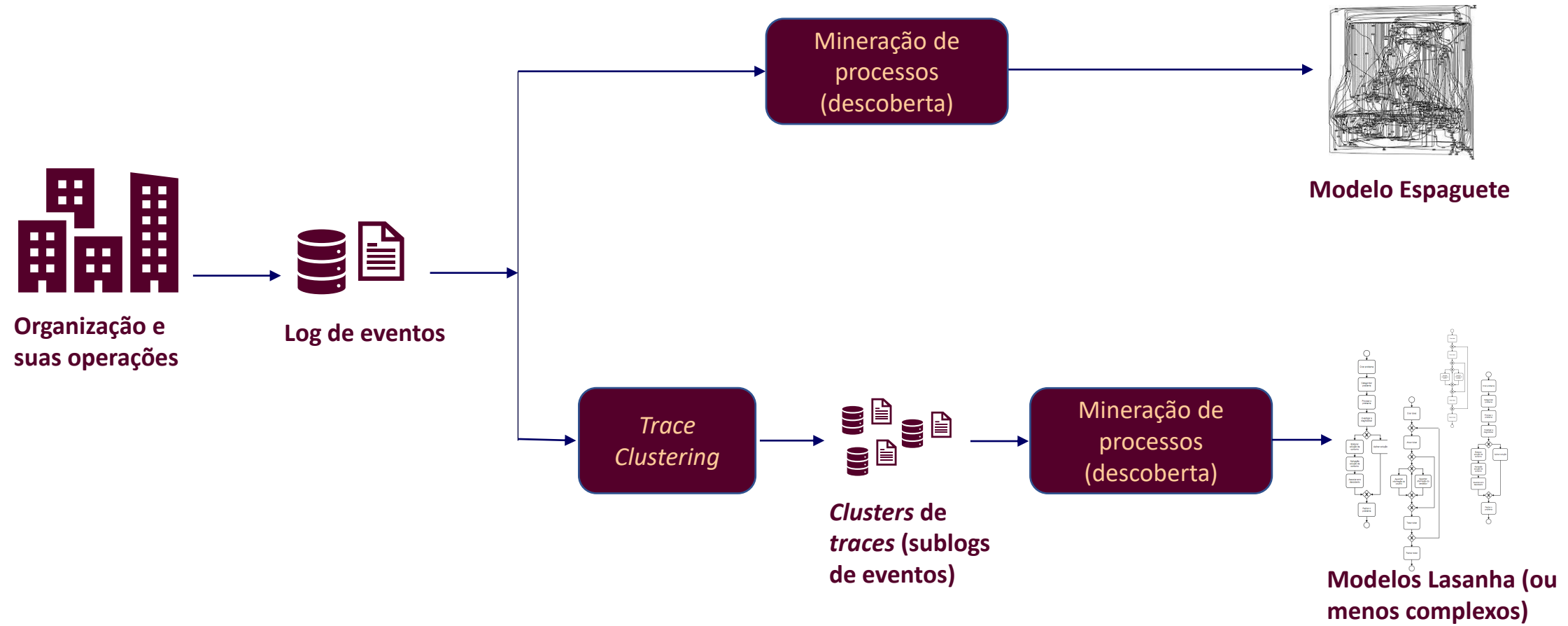
# Suporte operacional



Extra !!!

# Suporte operacional

- *Trace clustering*



Extra !!!

# Suporte operacional

- Detecção de anomalias

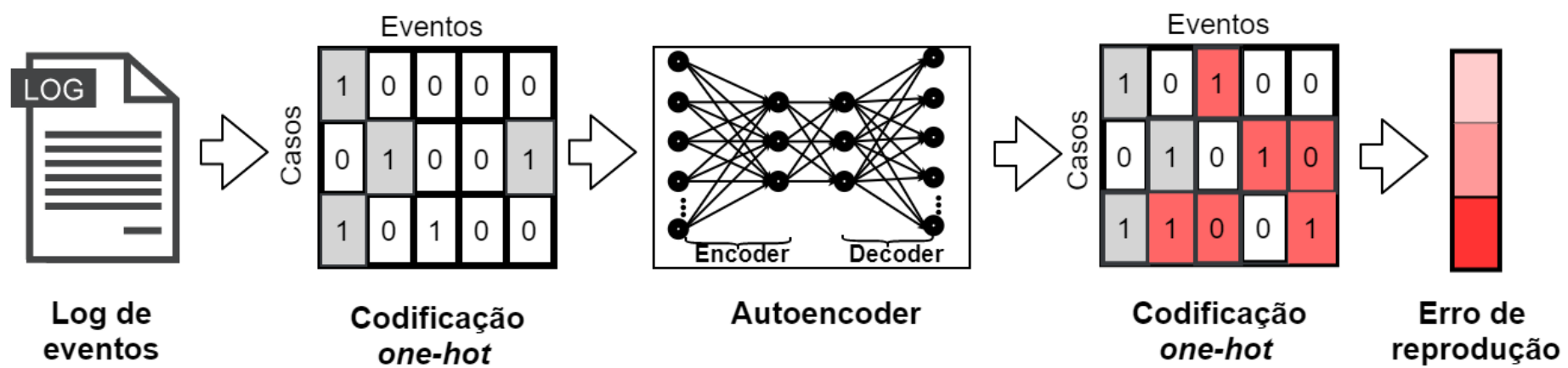


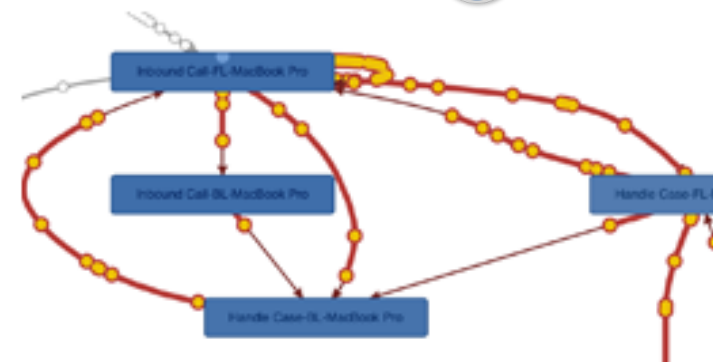
Figura de Esther M. Rojas Krugger - Documento de qualificação de mestrado: Detecção de anomalias em logs de eventos de processos de negócio: um estudo comparativo entre abordagens baseadas em redes neurais e baseadas em contagem. Universidade de São Paulo, 2020.

Abordagem proposta em: Nolle, T.; Luettgen, S.; Seeliger, A.; Mühlhäuser, M. Analyzing business process anomalies using autoencoders. Mach. Learn., Kluwer Academic Publishers, v. 107,n. 11, p. 1875–1893, nov. 2018.

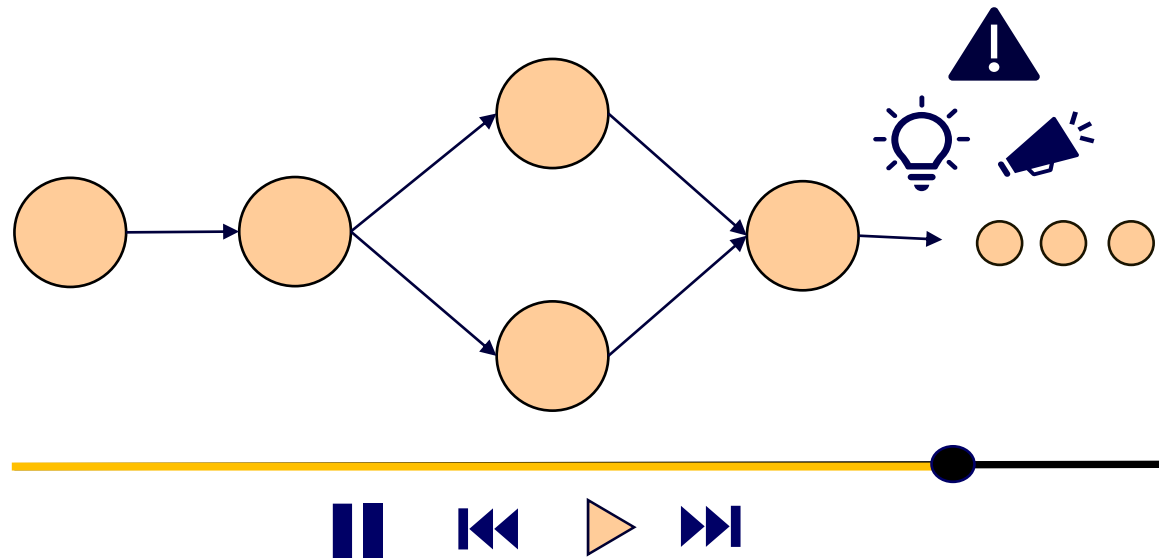
# Suporte operacional

- Análise descritiva

- Quantidade de casos e quantidade de variantes
- Média de tempo por variante
- Taxa de chegada de casos
- Alinhamento de atividades em relação ao tempo
- Períodos de tempo com poucas atividades
- Quantidade de casos tratados por período de tempo
- Simulação
- ...



# Suporte operacional



Predição de tempo

Detecção de desvio

Concept drift

Análise online



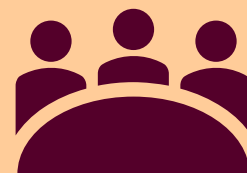
# Informações adicionais

...

# Trabalhando com mineração de processos: a necessidade do profissional para o mercado e para a pesquisa

Cientista de processos

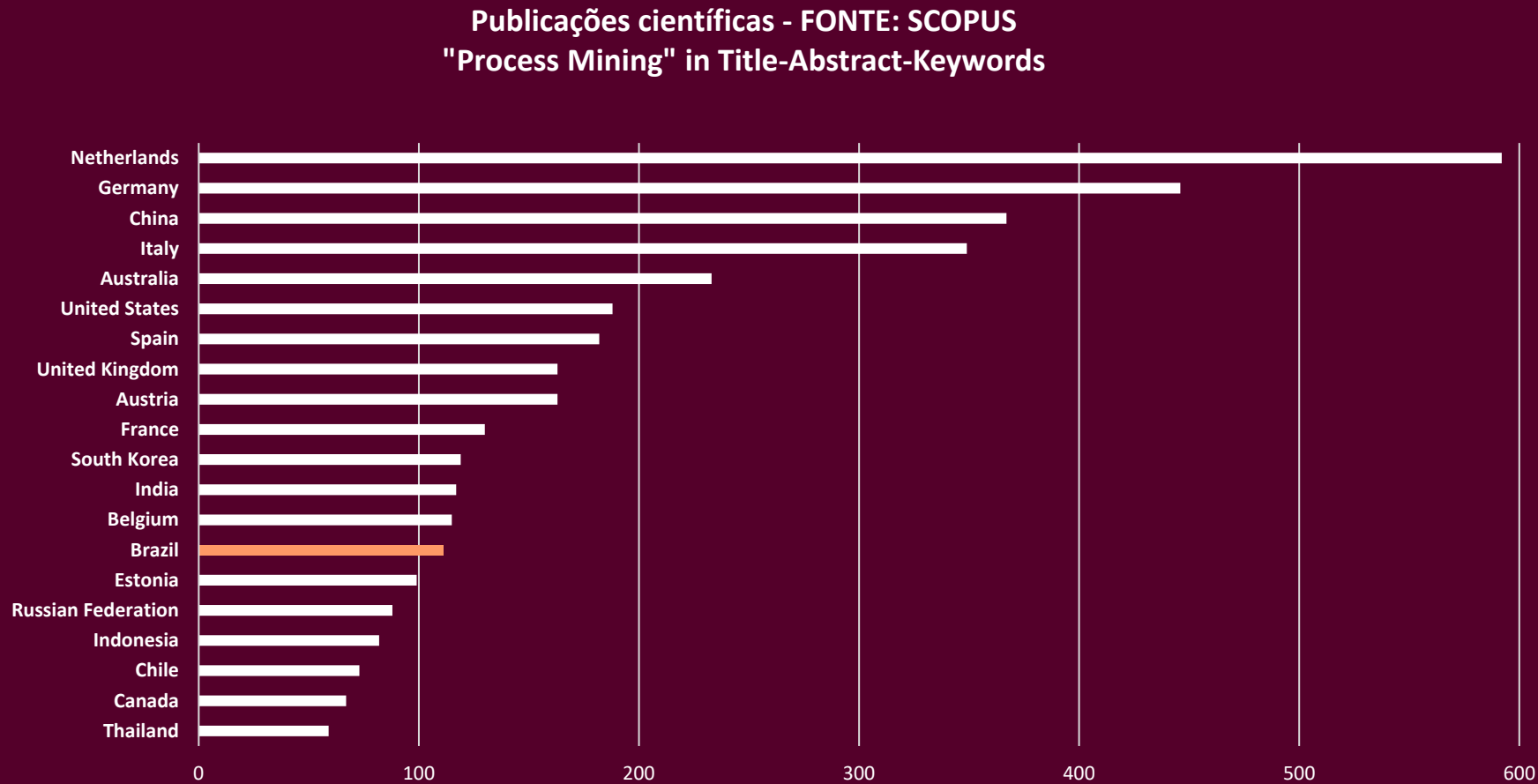
Cientista de dados



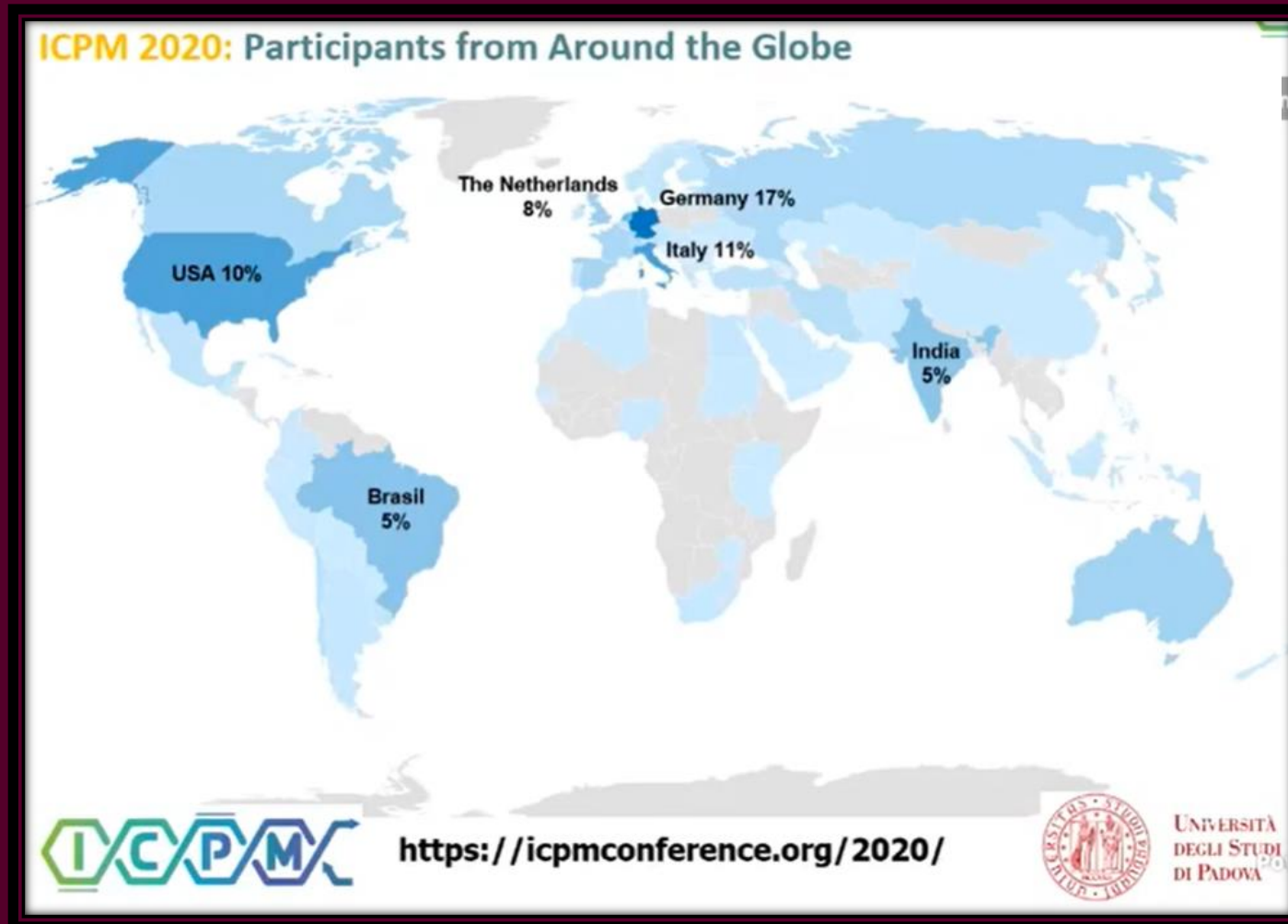
Analista de processos

Analista de negócios

# Brasil e a mineração de processos – visão acadêmica



# Brasil e a mineração de processos – International Conference on Process Mining



## Cursos na USP – Mineração de processos

- **Disciplina de pós-graduação**

**PPgSI – EACH – USP**      <http://ppgsi.each.usp.br/>

**Oferecimentos**      <http://ppgsi.each.usp.br/historico-de-oferecimento/>

**Matrículas (especiais)**      <http://ppgsi.each.usp.br/solicitacao/>

- **Curso de difusão**

**CCEX – EACH - USP**

**Programa de verão / Programa de inverno**

**Consultar oferecimento**

<http://www5.each.usp.br/cultura-e-extensao-ccex-nova/>

Profa. Dra. Sarajane Marques Peres

[sarajane@usp.br](mailto:sarajane@usp.br)

[each.usp.br/processmining](http://each.usp.br/processmining)

[www.linkedin.com/in/sarajane-marques-peres](https://www.linkedin.com/in/sarajane-marques-peres)

**USP LESTE**

Escola de Artes, Ciências e Humanidades (EACH)

Universidade de São Paulo



Natália Dourado / EACH/ USP Imagens