KU LEUVEN

FACULTY OF ECONOMICS AND BUSINESS

SAI

studiecentrum voor automatische informatieverwerking vzw

Process Mining in 2024: een stand van zaken en welke tools?

Jochen De Weerdt

29 februari 2024

1. Quick recap on Process Mining

2. The Process Mining tooling landscape

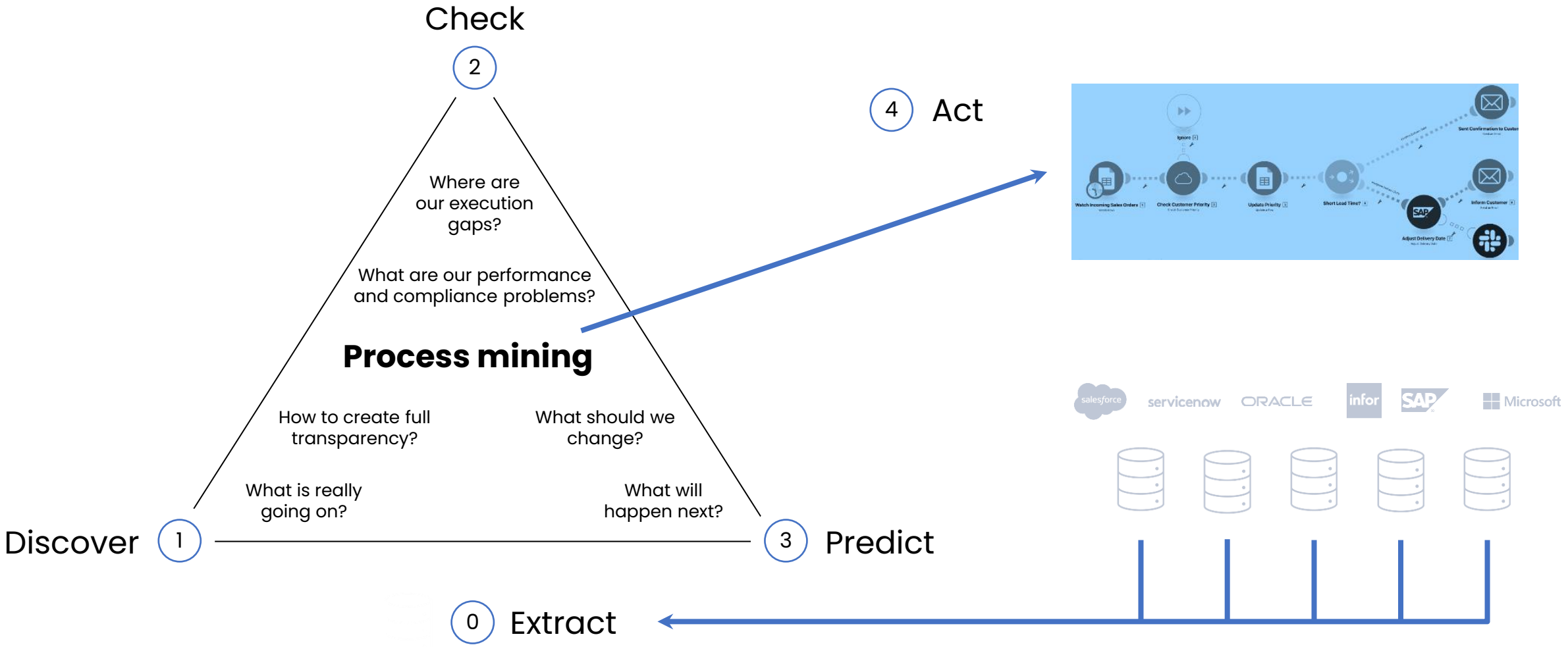
3. Trends in Process Mining

1. Quick recap on Process Mining

2. The Process Mining tooling landscape

3. Trends in Process Mining

High-level view of process mining

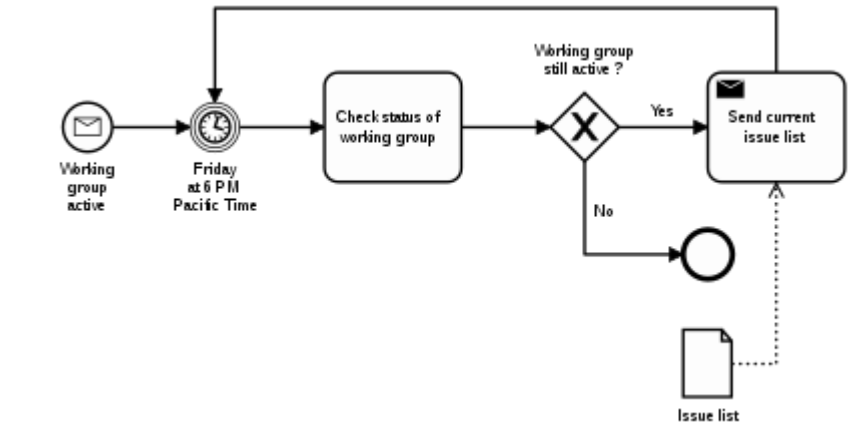


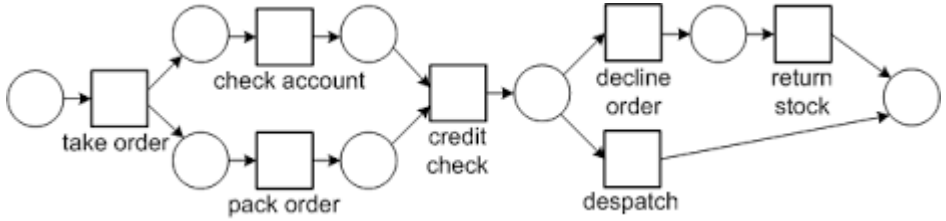
Event log

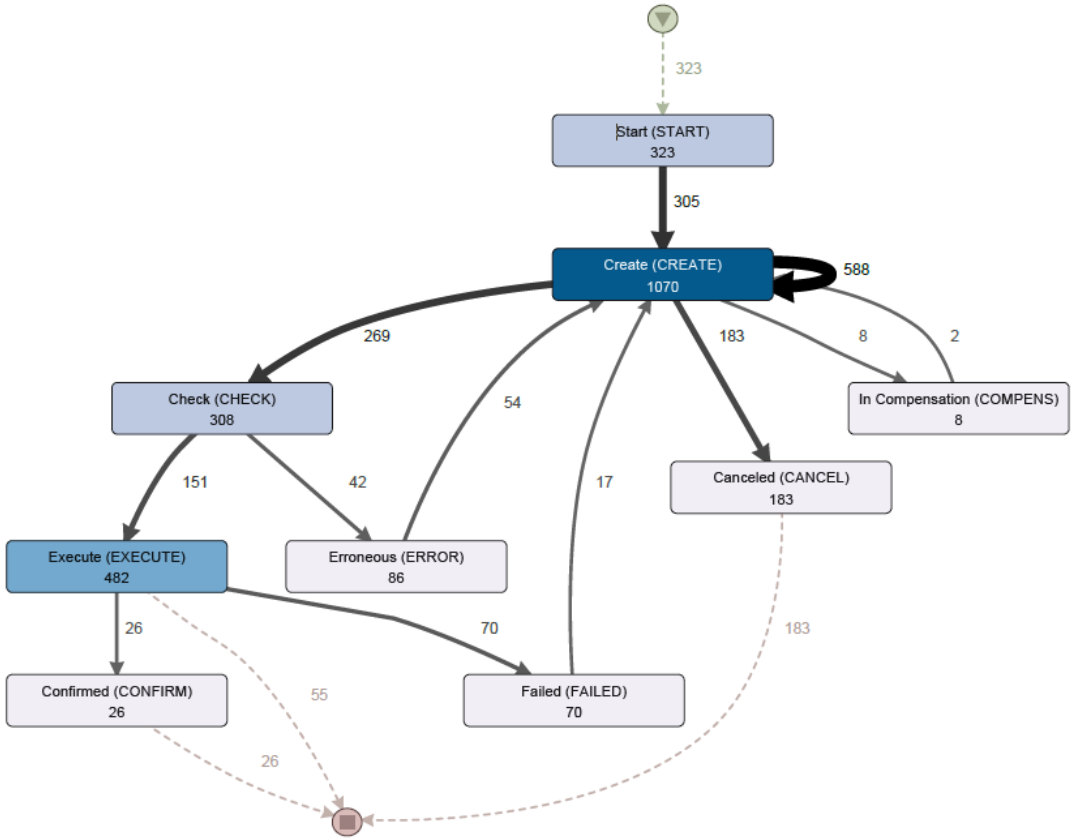
case id	event id	properties				
		timestamp	activity	resource	cost	...
1	35654423	30-12-2010:11.02	register request	Pete	50	...
	35654424	31-12-2010:10.06	examine thoroughly	Sue	400	...
	35654425	05-01-2011:15.12	check ticket	Mike	100	...
	35654426	06-01-2011:11.18	decide	Sara	200	...
	35654427	07-01-2011:14.24	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	...
4	35654641	06-01-2011:15.02	register request	Pete	50	...
	35654643	07-01-2011:12.06	check ticket	Mike	100	...
	35654644	08-01-2011:14.43	examine thoroughly	Sean	400	...
	35654645	09-01-2011:12.02	decide	Sara	200	...
	35654647	12-01-2011:15.44	reject request	Ellen	200	...
...	...	...	...	...	...	...

- A process creates cases or process instances/excecutions/traces
- A case consists out of events
  - Each event belongs to a case (**case id**)
  - Events are ordered (**timestamp**)
  - Each event reflects the execution of a particular task in the process (**activity label**)
- Events and cases can contain attributes.

Process models







1. Quick recap on Process Mining

2. The Process Mining tooling landscape

3. Trends in Process Mining

Industry adoption

- Process mining has also been recognized as **one of the most important trends in big data** by the Financial Times*
- In the recent Global Process Mining Survey by Deloitte** it was found that **process mining is already used in 83% of the companies questioned** on a global scale. Furthermore, 84% believes that process mining delivers value with 87% of the non-adopters eager to take up activities in the future.
- Gartner estimates that RPA is already a 2 billion+ dollar market***
- Fortune Business Insights reports that process mining software market size was valued at USD 1.13 billion, and is expected to grow to USD 27.72 billion by 2030 ****

* https://www.ft.com/content/402553f4-c4a4-11e7-b30e-a7c1c7c13aab
** https://www2.deloitte.com/de/de/pages/finance/articles/global-process-mining-survey-2021.html
*** https://www.gartner.com/en/newsroom/press-releases/2020-09-21-gartner-says-worldwide-robotic-process-automation-software-revenue-to-reach-nearly-2-billion-in-2021
**** https://www.fortunebusinessinsights.com/process-mining-software-market-104792

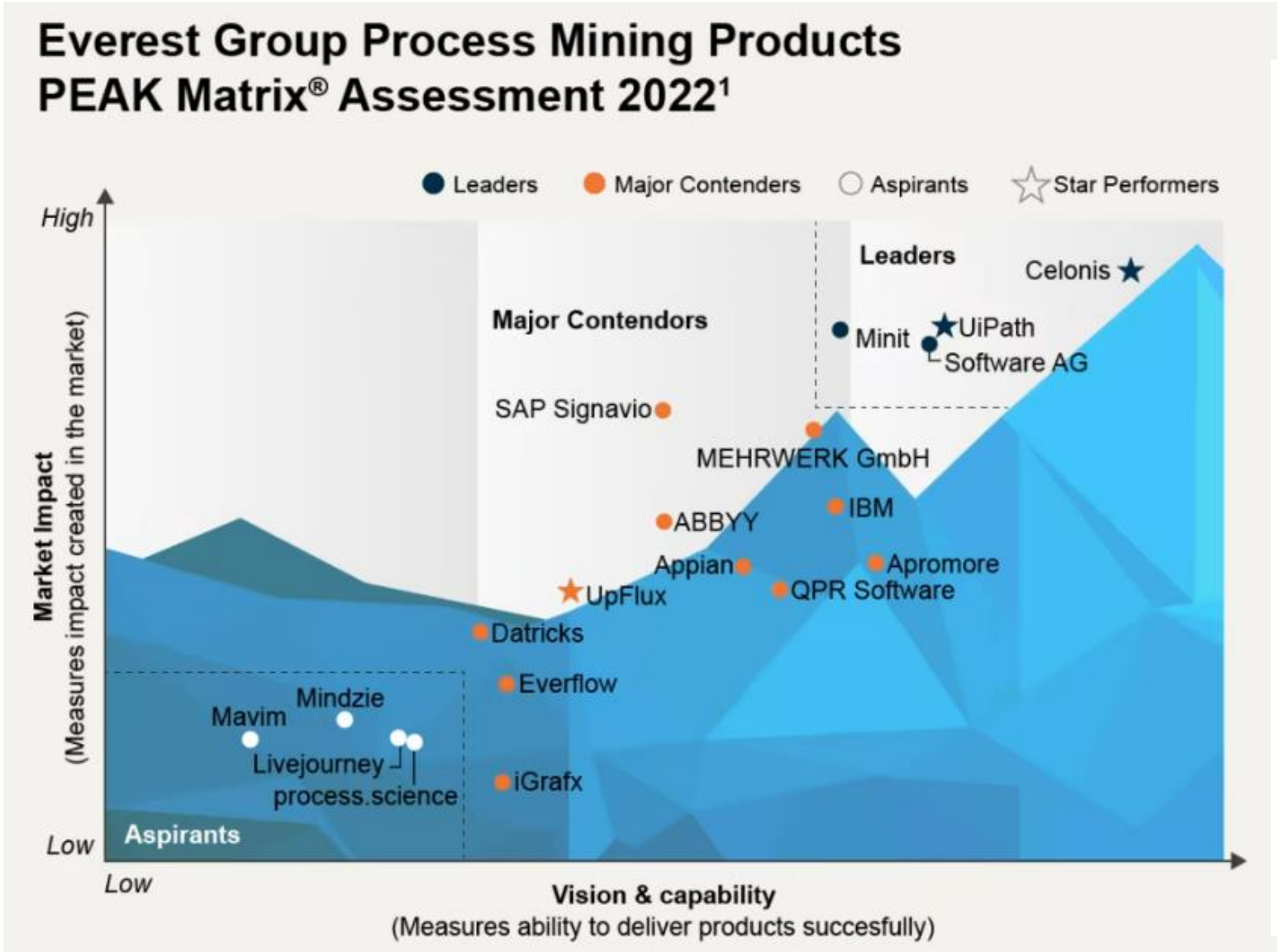


Figure 1: Magic Quadrant for Process Mining Tools

Quite some consolidation activity in 2019-2023

DATE	PROCESS MINING VENDOR	ACQUIRING COMPANY	MARKET/INDUSTRY
May 2022	Everflow	Pegasystems	Business process management (BPM)
March 2022	Minit	Microsoft	Automation/business intelligence
March 2022	PAFnow	Celonis	Process mining
January 2022	Logpickr	iGrafx	BPM
December 2021	FortressIQ	Automation Anywhere	Robotic process automation (RPA)
August 2021	Lana Labs	Appian	BPM
April 2021	MyInvenio	IBM	BPM
January 2021	Signavio	SAP	ERP
October 2019	ProcessGold	UiPath	RPA
May 2019	TimelinePI	Abbyy	Intelligent document processing

Commercial vendor tooling

- Strengths
  - Useful analysis and dashboarding tools
  - Strong focus on (automated) data sourcing
  - Especially powerful on “known, standardized” processes (P2P, O2C, C2R)

- Weaknesses
  - Expensive
  - Open-source alternatives exist PM4PY BUPAR
  - No silver bullet for data quality issues
  - Simplistic process models (Directly-follows graphs*)
  - Limited predictive/prescriptive capabilities

* van der Aalst, W. (2019). A practitioner’s guide to process mining: limitations of the directly-follows graph. *Procedia Computer Science*, *164*, 321-328.

1. Quick recap on Process Mining

2. The Process Mining tooling landscape

3. Trends in Process Mining

Moving beyond descriptive Process Mining

- Classical process mining
  - Process Discovery & Conformance Checking
  - Key challenges are not techniques, nor tooling

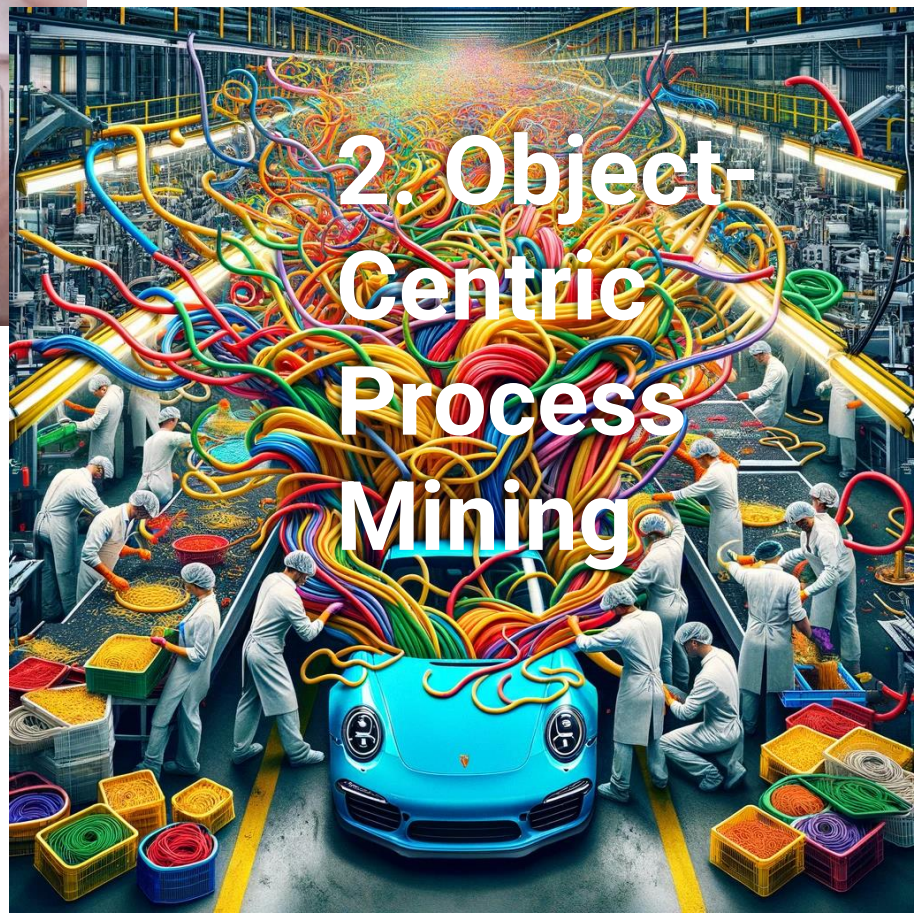
- Real challenges include
  - Unlocking the data
    - Data capture, integration, & quality
    - Data governance (privacy, fairness, …)
  - Added value & insights
    - Addressing variability
  - Organizational readiness for true process optimization
    - Move away from indirect two-step approach

Trend 1: Task mining

Trend 2: Object-Centric Process Mining

Trend 3: AI-driven Automated Optimization

1. Task Mining



2. Object-Centric Process Mining

3. AI-driven Automated Optimization

1. Task mining

- Data types
  - user interaction data = desktop data
  - keystrokes, mouse clicks and data entries that occur as part of completing a given operation
- Technology
  - Optical character recognition (OCR)
  - Natural language processing (NLP)
  - ML/AI
  - IoT
- Part of process mining/process discovery
  - Applied to user interaction data separately to identify operational inefficiencies
  - Or, integrated into a broader process analytics exercise, identifying activities that are part of a more comprehensive business process (activity recognition)
- Closely linked to RPA
- Vendors with prominent task mining functionalities: Celonis, Apromore, UiPath

2. Object-Centric Process Mining

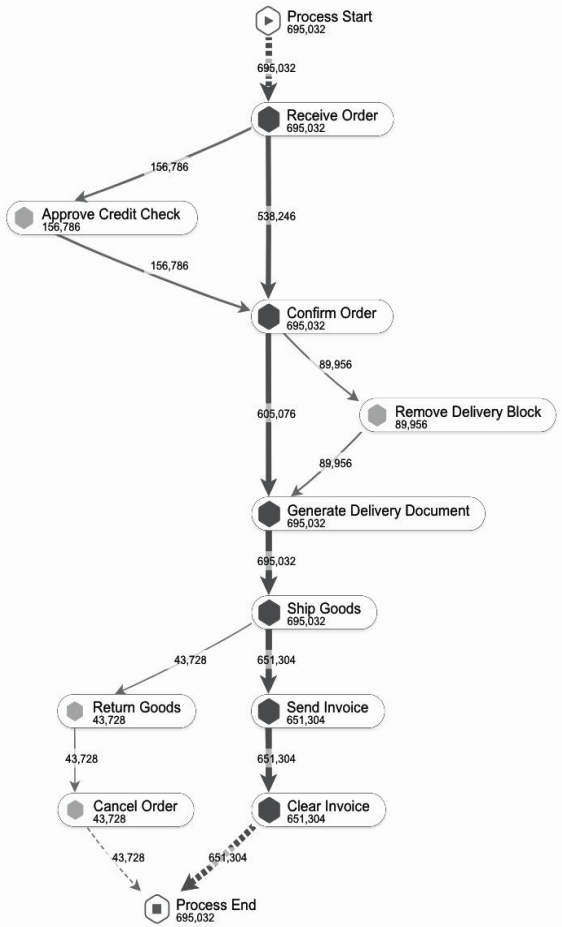
This is what we expect …


This is what we find …
Understanding this is the key to process improvement!

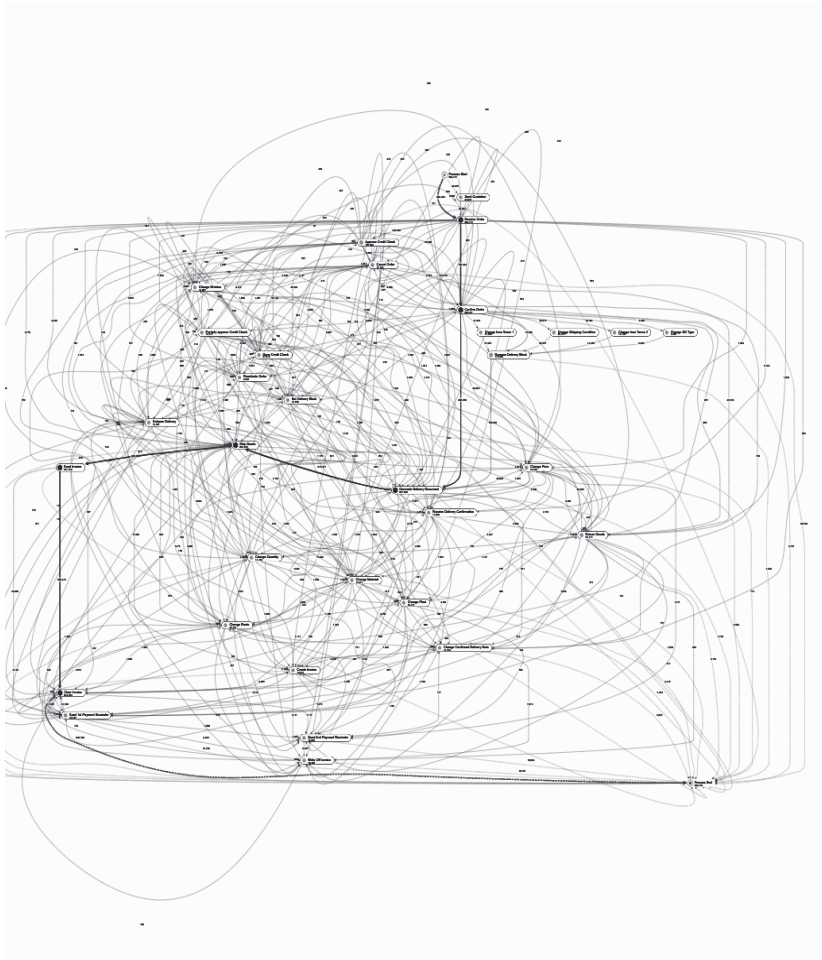
Actual processes are very different from what stakeholders think!

Happy path

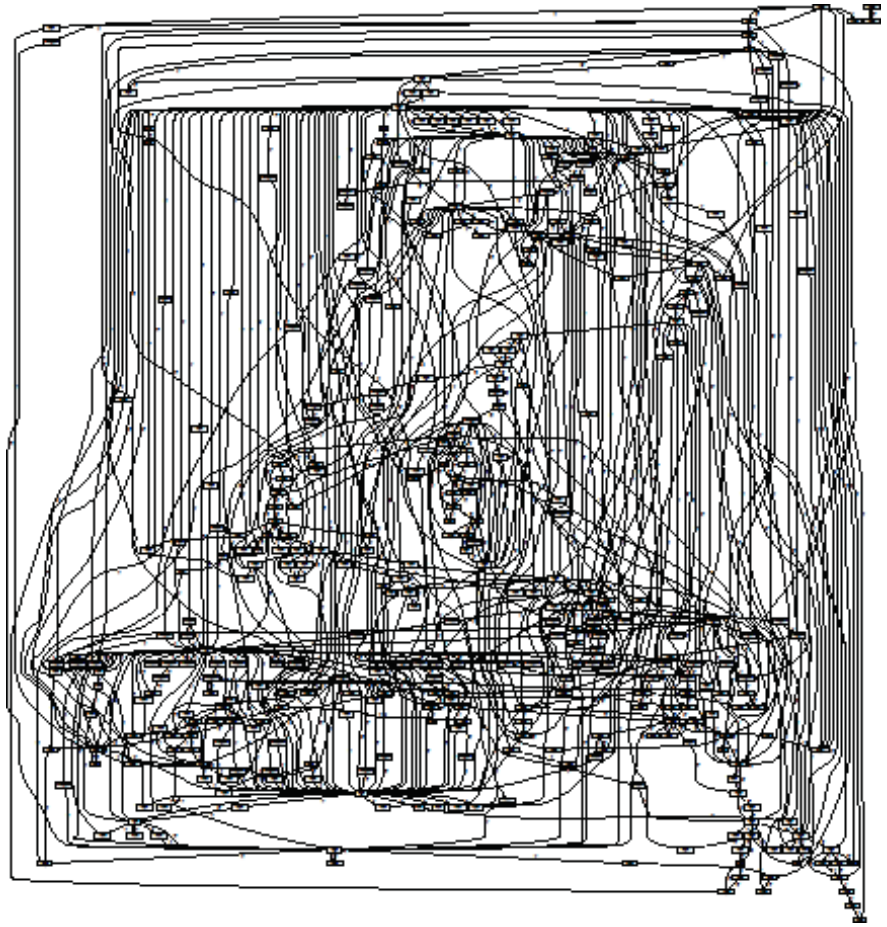
Expected paths



Reality



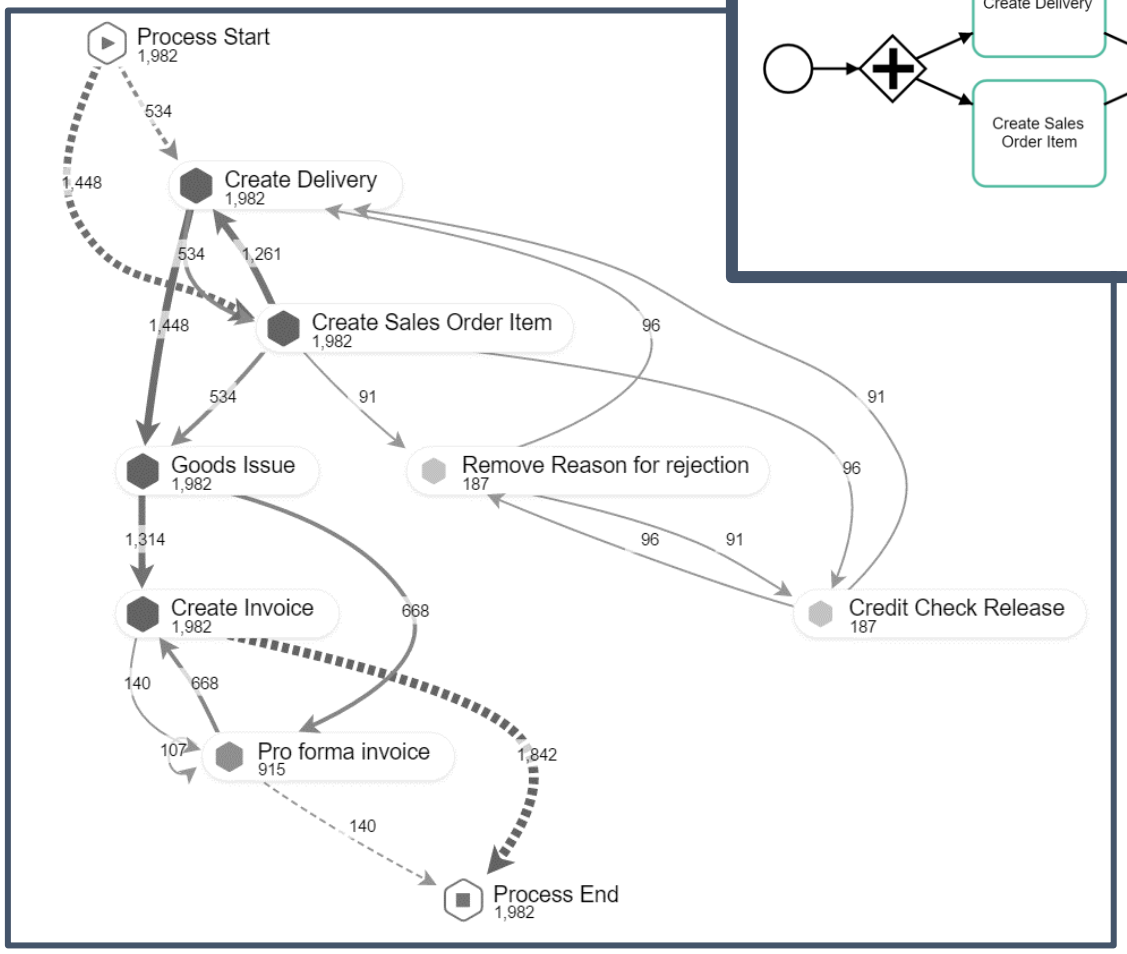
Processes have more variability than expected, but is this all ?

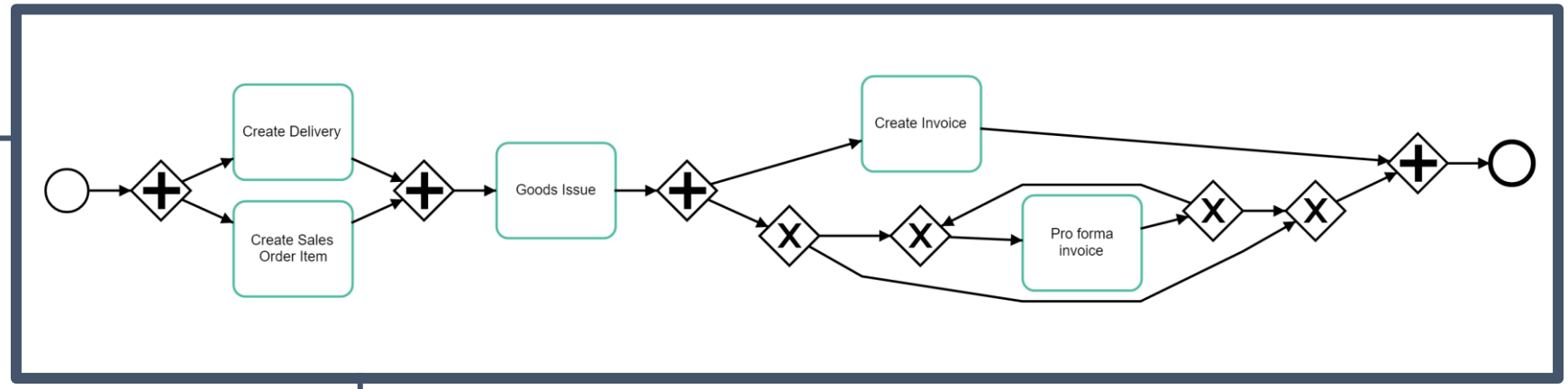


No!

→ Concurrency
→ Intertwined objects

Old news: Concurrency matters!





3,308 cases

501 variants

both cover **60%** of the cases

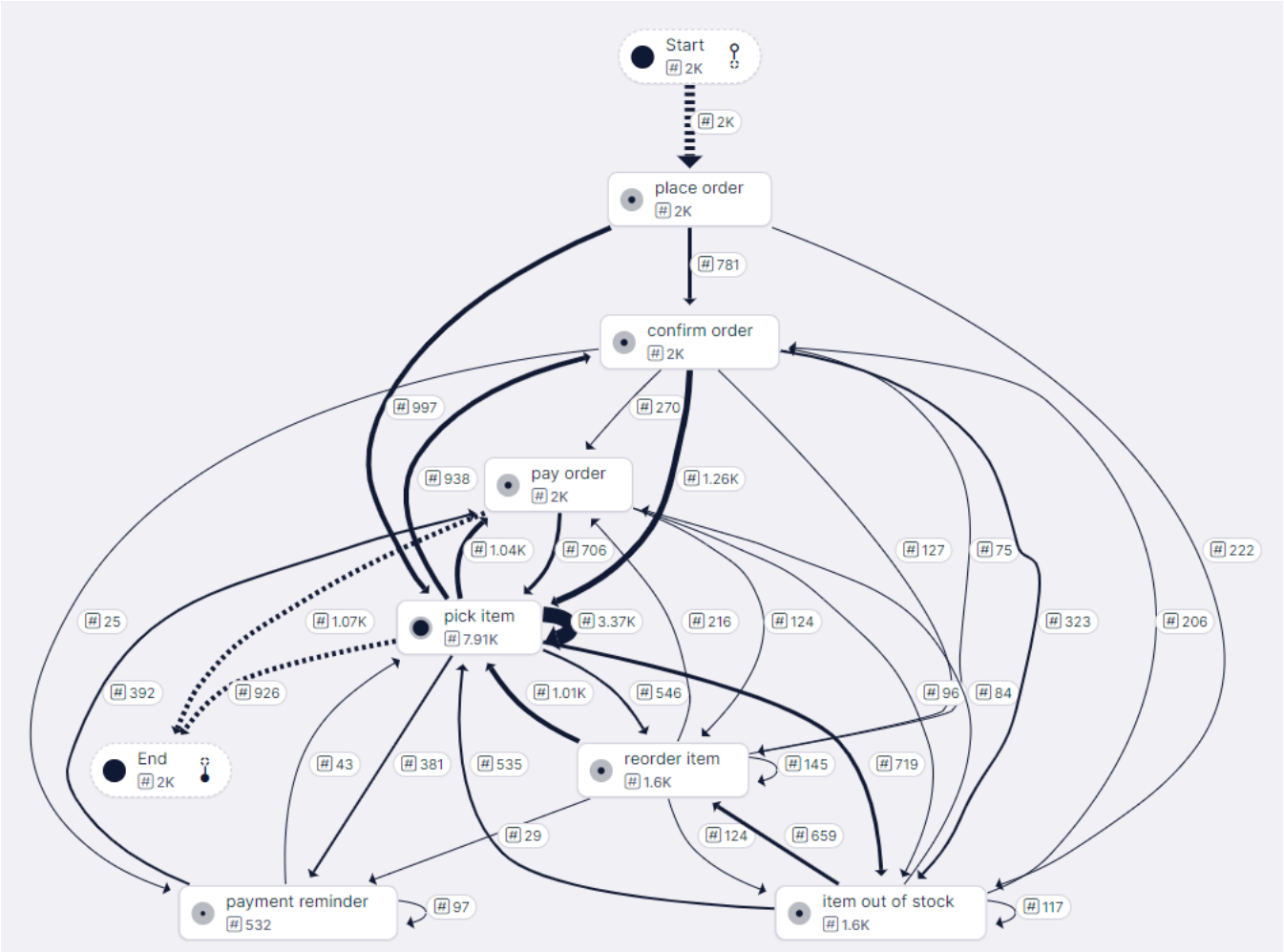
Activities not executed in a fixed order create loops leading to Spaghetti-like underfitting models.

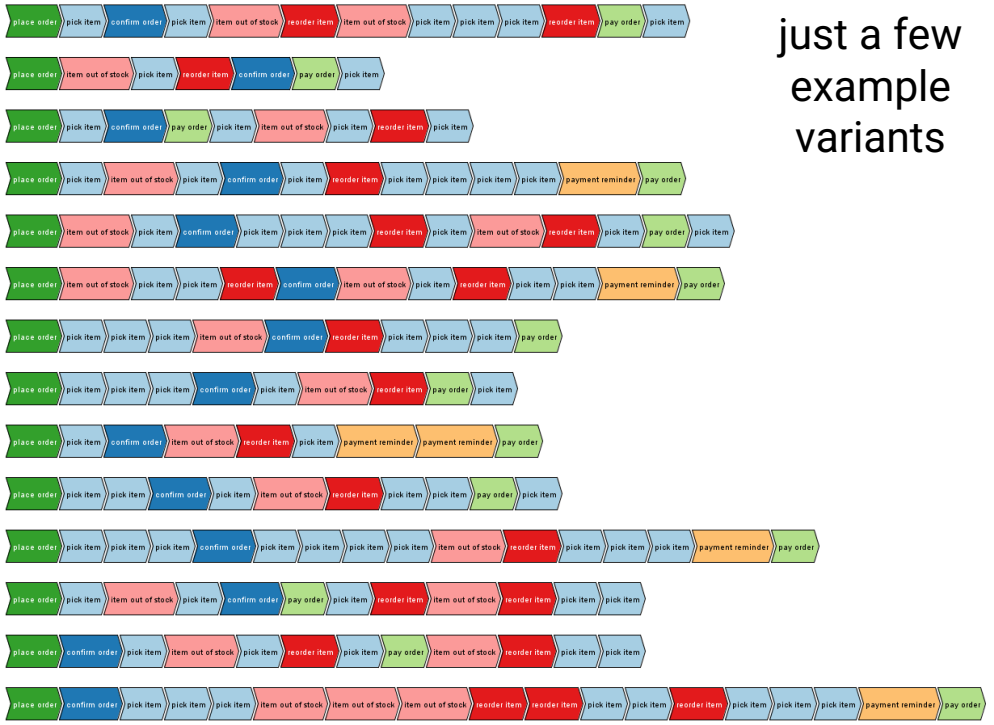
There is more: We are actually dealing with Rainbow Spaghetti

Example object types:

- Orders
- Items
- Packages
- Machines
- Employees
- Patients
- Customers
- Machines
- Containers
- Payments
- Vehicles
- Rooms
- Etc.

Event log describing two object types: Orders and Items



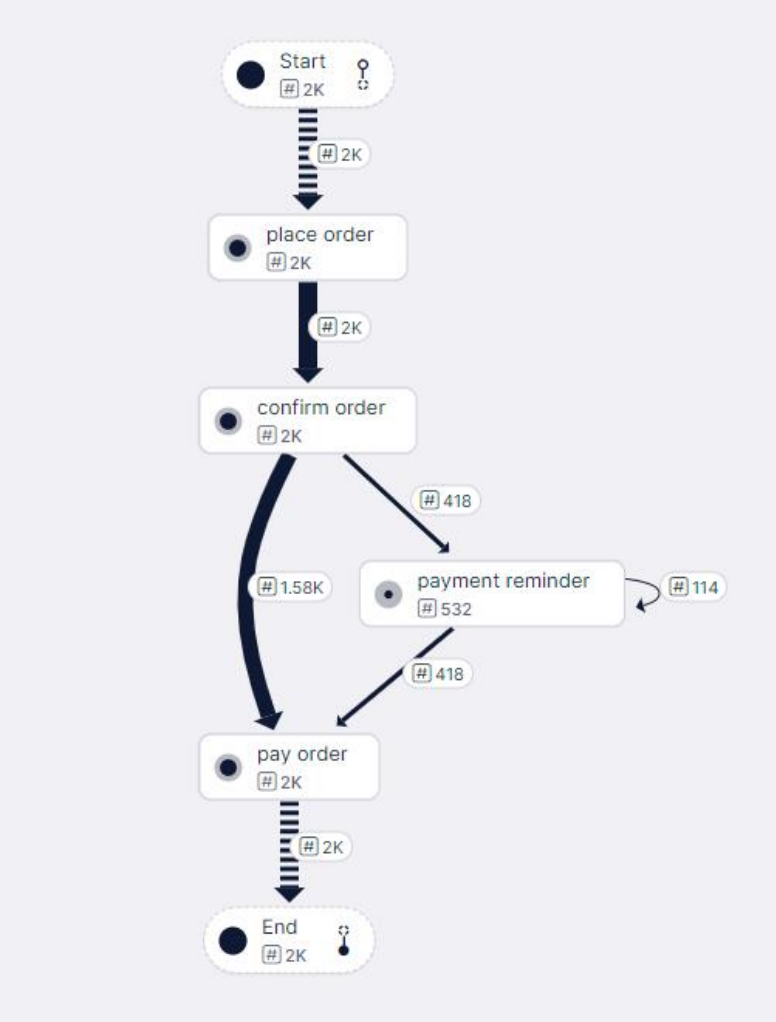


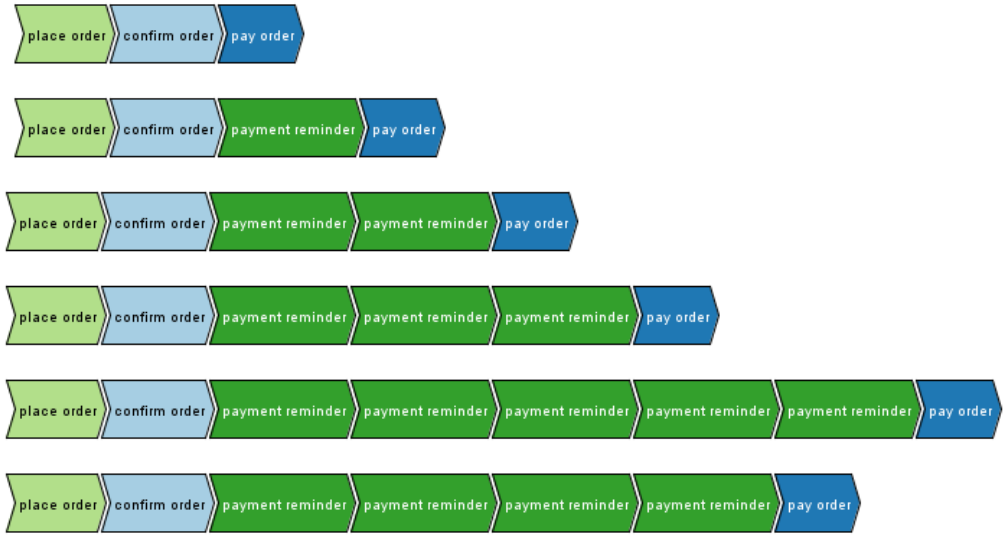
just a few example variants

2,000 orders
7,914 items
17,648 events

2,000 cases (here orders)
1,033 variants

Event log describing just the orders

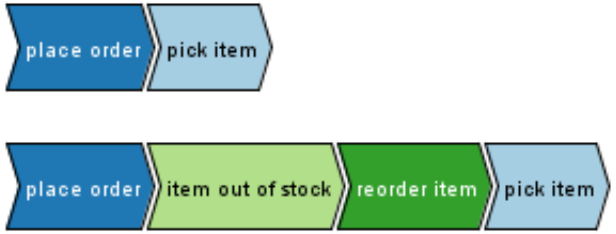




2,000 orders
6,532 events

6 variants

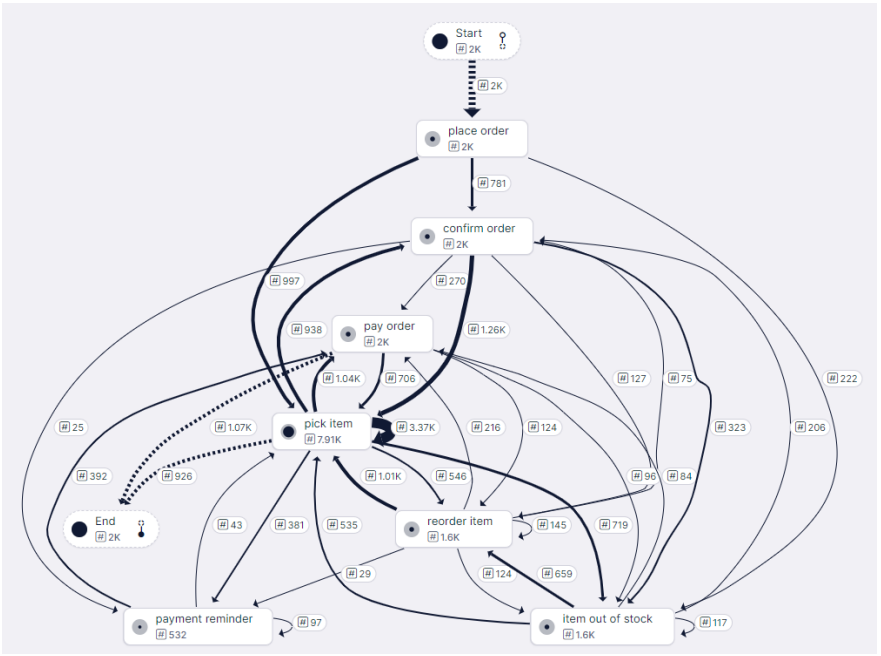
Event log describing two object types: Orders and Items



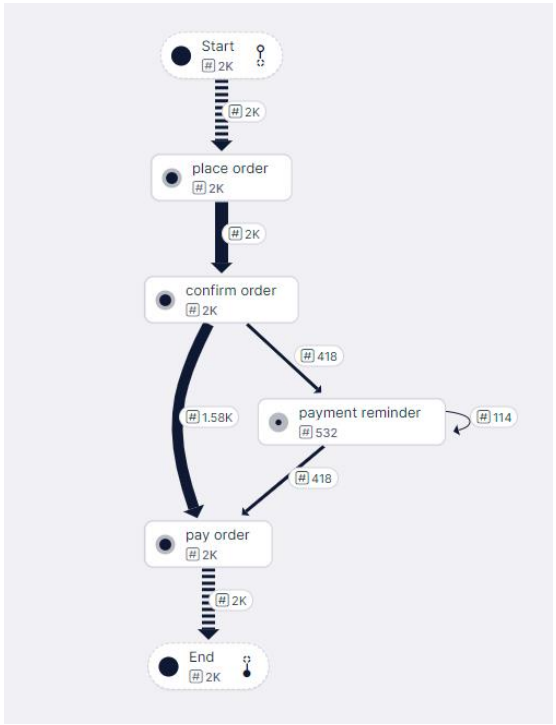
7,914 items
19,030 events

2 variants

Disentangling Spaghetti Bicolore (two objects)



=



×

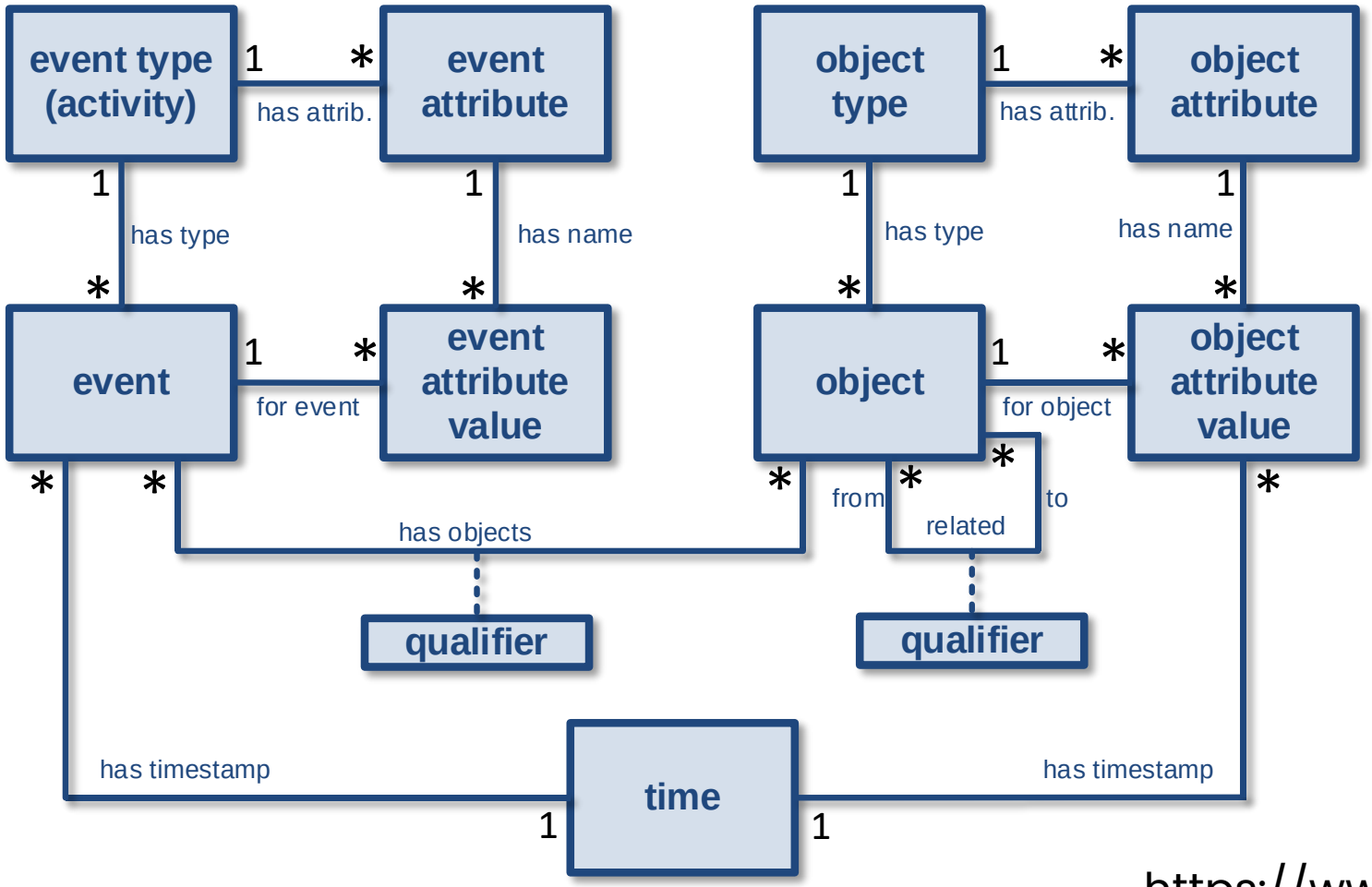
1033 variants **6 variants** **2 variants**

Why OCPM?

- Avoid repeatedly going back to your source systems
- See and understand the interactions between different object types
- Avoid distortions due to the single-case assumption
  - circumventing convergence and divergence problems

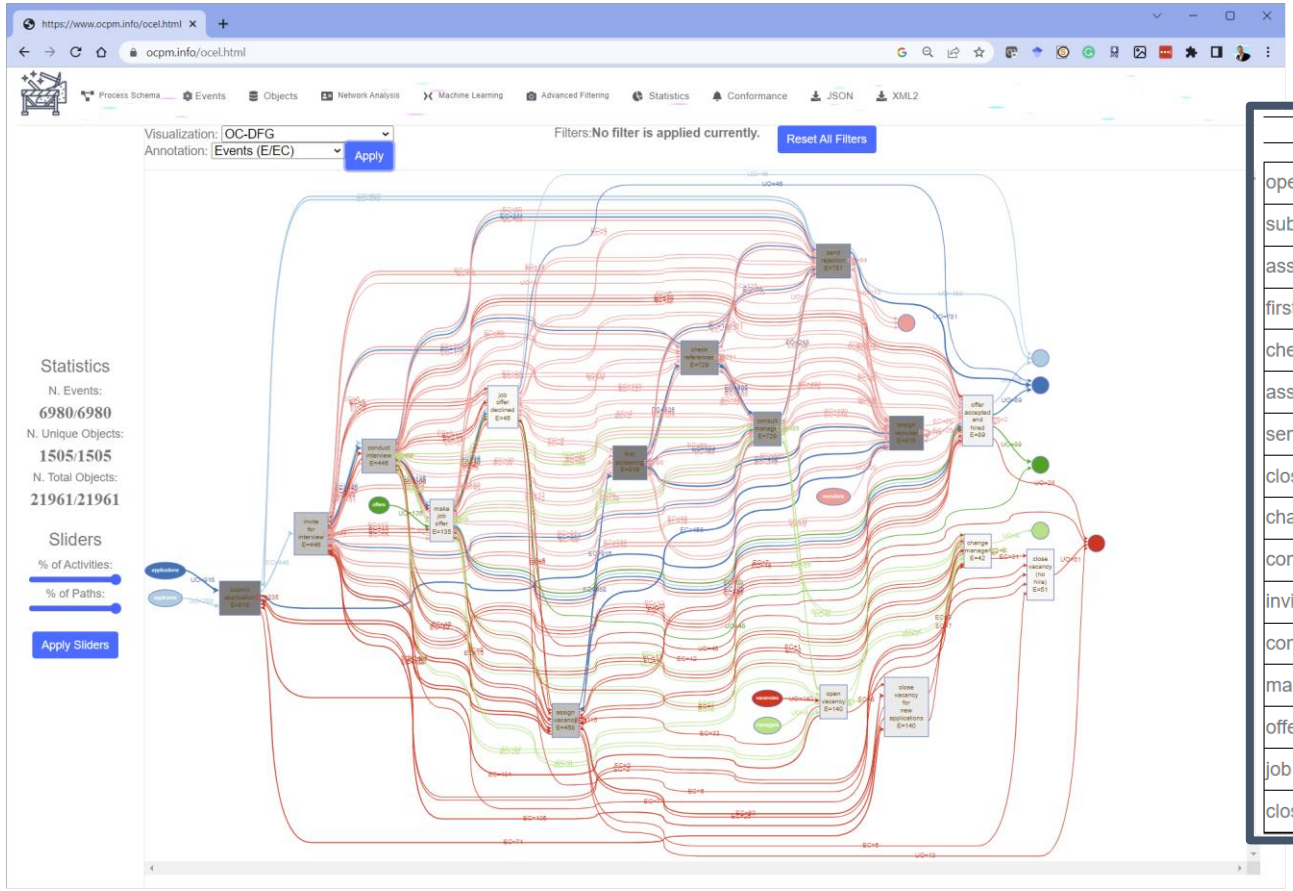
- But this requires rethinking and redeveloping process mining techniques
  - Data storage standard
  - Process discovery

Object-Centric Event Data: OCEL 2.0



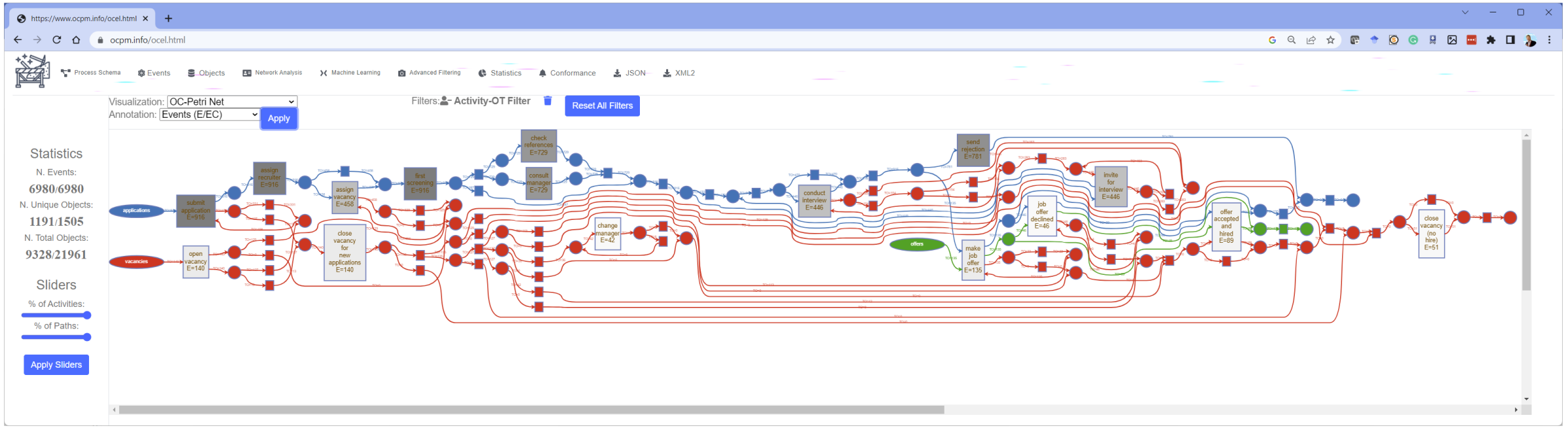
https://www.ocel-standard.org/

Example OC-DFG



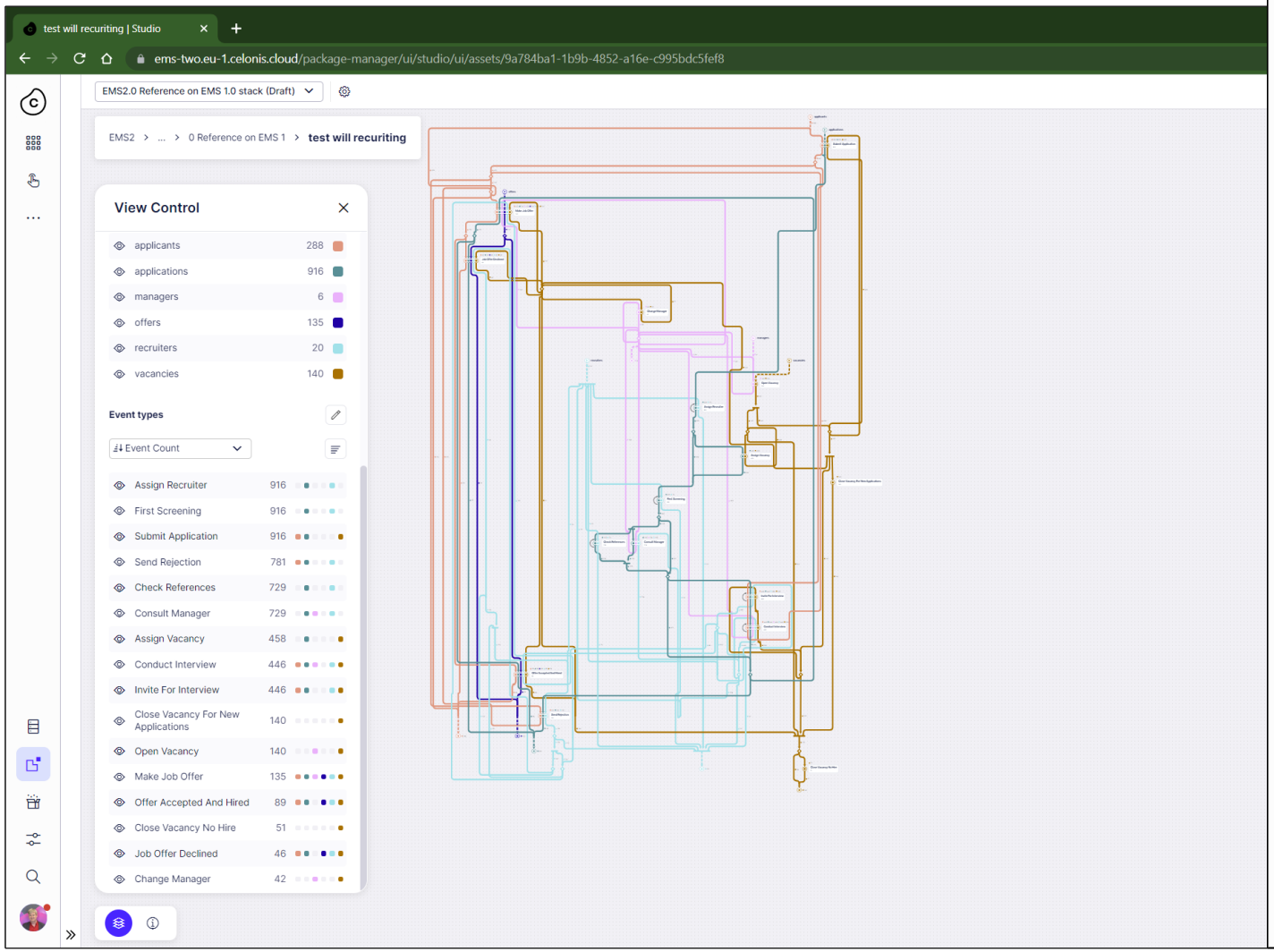
	vacancies	managers	applicants	applications	recruiters	offers
open vacancy	UO=140	UO=140				
submit application	UO=458		UO=916	UO=916		
assign recruiter				UO=916	UO=2748	
first screening				UO=916	UO=916	
check references				UO=729	UO=1458	
assign vacancy	UO=458			UO=458		
send rejection			UO=781	UO=781	UO=781	
close vacancy for new applications	UO=140					
change manager	UO=42	UO=84				
consult manager		UO=729		UO=729	UO=729	
invite for interview	UO=446		UO=446	UO=446	UO=892	
conduct interview	UO=446	UO=446	UO=446	UO=446	UO=446	
make job offer	UO=135	UO=135	UO=135	UO=135	UO=135	UO=135
offer accepted and hired	UO=89		UO=89	UO=89	UO=89	UO=89
job offer declined	UO=46		UO=46	UO=46	UO=46	UO=46
close vacancy (no hire)	UO=51					

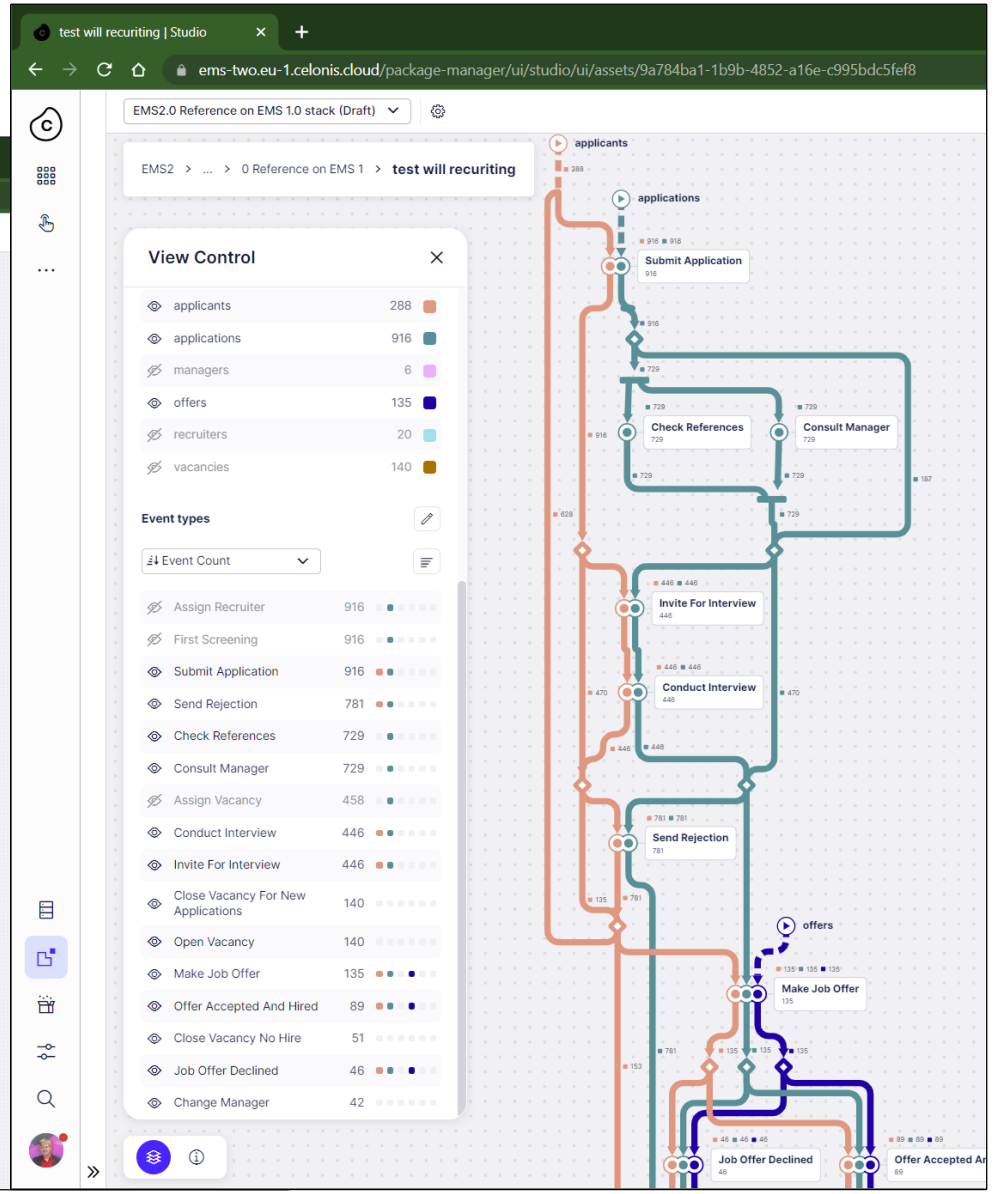
Example OC-PN



ocpm.info

OC-BPMN in Celonis



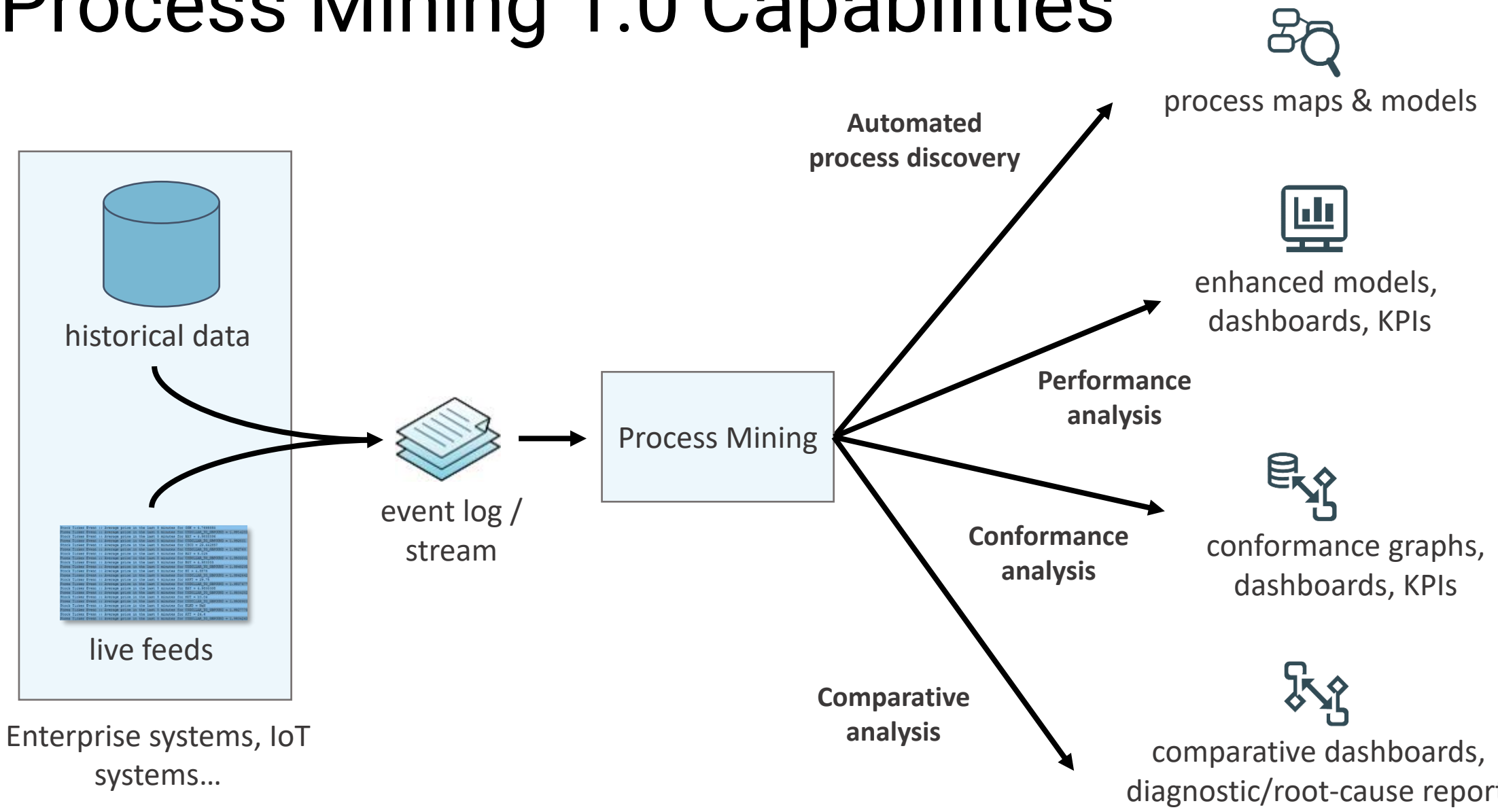


Learn more about OCPM

- W. van der Aalst, Object-Centric Process Mining: Unraveling the Fabric of Real Processes. Mathematics 2023, 11, 2691. https://doi.org/10.3390/math11122691
- W. van der Aalst, Object-Centric Process Mining: The Next Frontier in Business Performance. 2023. Available online: celon.is/OCPM-Whitepaper
- Goossens, A., Verbruggen, C., Snoeck, M., De Smedt, J., & Vanthienen, J. (2023). Aligning Object-Centric Event Logs with Data-Centric Conceptual Models. In *International Conference on Business Process Modeling, Development and Support* (pp. 44-59). Cham: Springer Nature Switzerland. https://link.springer.com/chapter/10.1007/978-3-031-34241-7_4
- Goossens, A., De Smedt, J., & Vanthienen, J. (2024). Extracting Process-Aware Decision Models from Object-Centric Process Data. arXiv preprint arXiv:2401.14847.

3. AI-driven Process Optimization

Process Mining 1.0 Capabilities



Process Mining is the beginning of a journey…

How do my processes look like?
Where are the bottlenecks, wastes, compliance violations, positive & negative deviance?

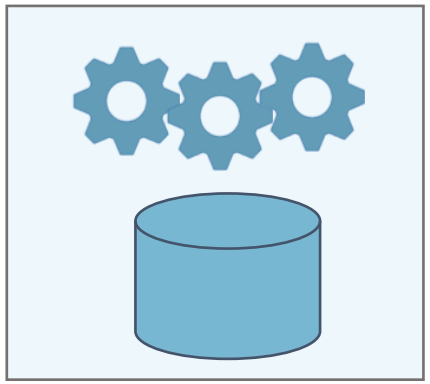
(Descriptive) Process Mining

Automated Process Discovery
Conformance Checking
Performance Mining
Comparative Variant Analysis

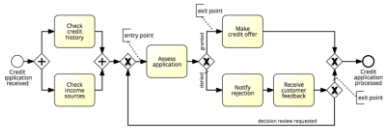
Next step: predict the future

Data-Driven Process Simulation

Enterprise System

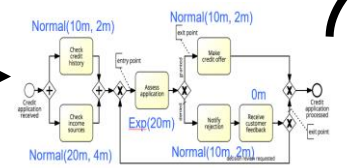


Process Constraints / Process Model

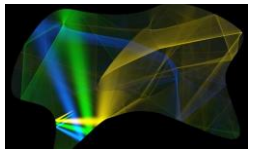


Model Discovery

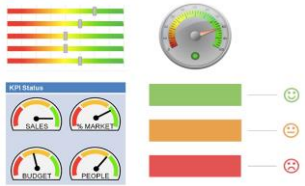
Change Specification



Simulation Model **(Digital Process Twin)**



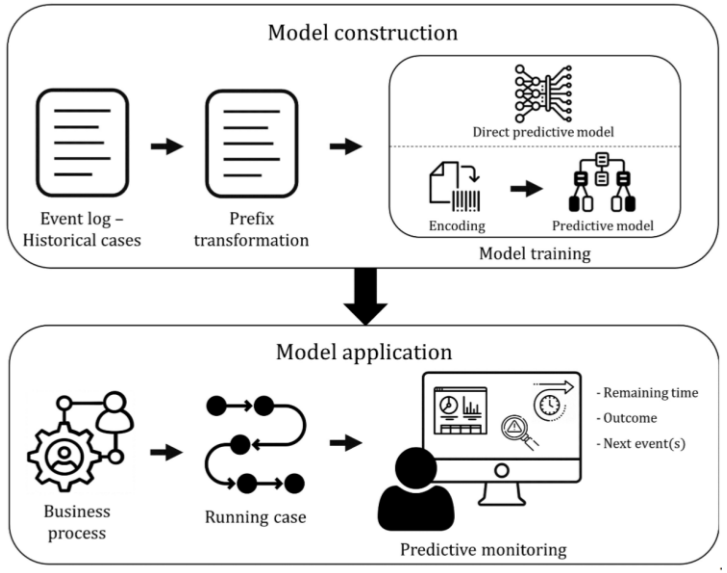
Change Simulation

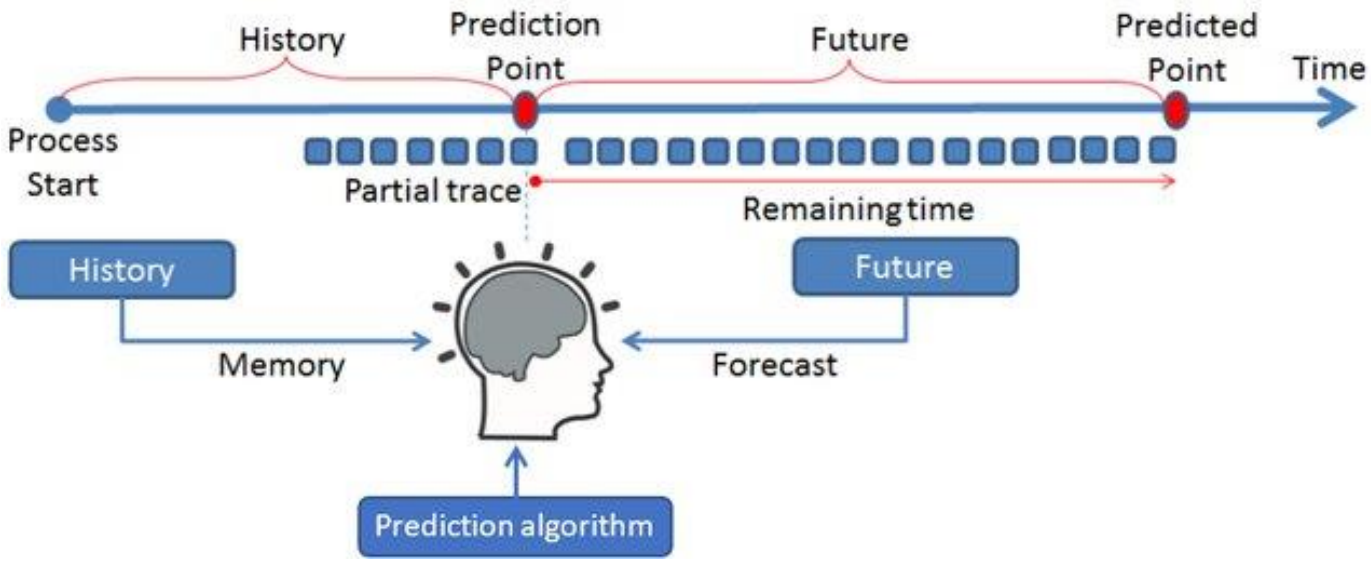


Predicted Performance Profile

https://github.com/AutomatedProcessImprovement/Simod

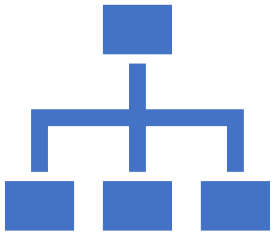
Predictive Process Monitoring





Predict **Process Outcome**	Predict **Process Performance**	Predict **Future Events**
Is this loan offer going to be rejected?	Will this claim take more than 5 days to be handled? How much longer will it take?	What activity is likely to be executed next? And after that?

Pitfalls of predictive process optimization



Organizational Pitfalls

Lack of buy-in from operations

Lack of trust in predictions

Prescriptions not linked with actions

Lack of processes to validate, monitor, and maintain prescriptive models

Technical Pitfalls

Insufficient data availability & quality

Lack of uncertainty modeling

Drifts and out-of-distribution predictions

Correlation ≠ Causation

Accurate Predictions, Invalid Recommendations: Lessons Learned at the Dutch Social Security Institute UWV

Marcus Dees, Massimiliano de Leoni, Wil M. P. van der Aalst & Hajo A. Reijers

Chapter | First Online: 05 August 2021

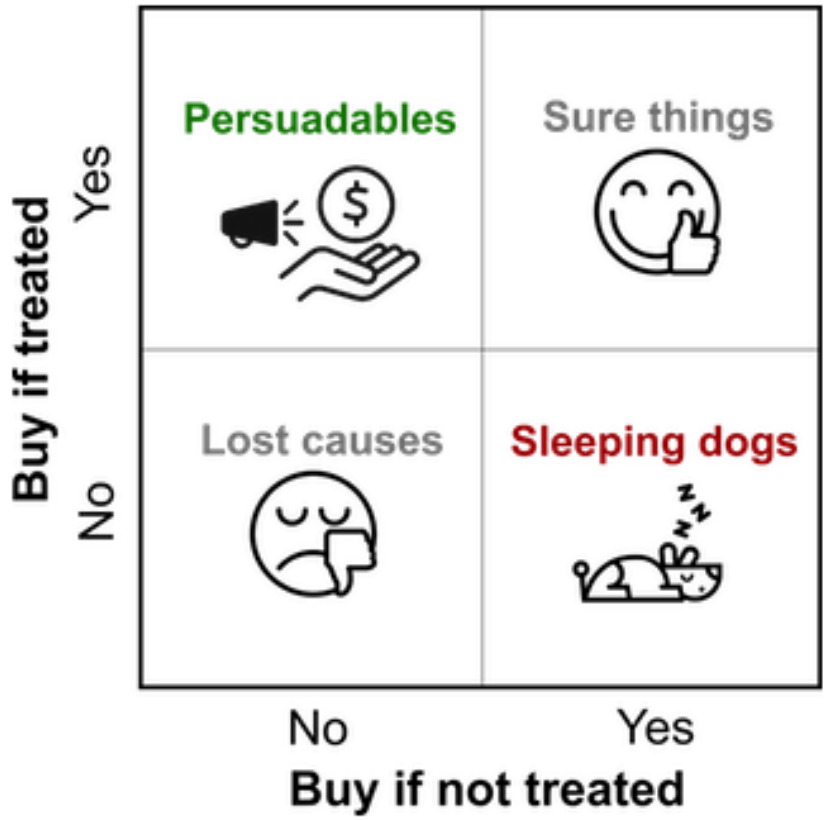
1785 Accesses | 1 Citations

Abstract

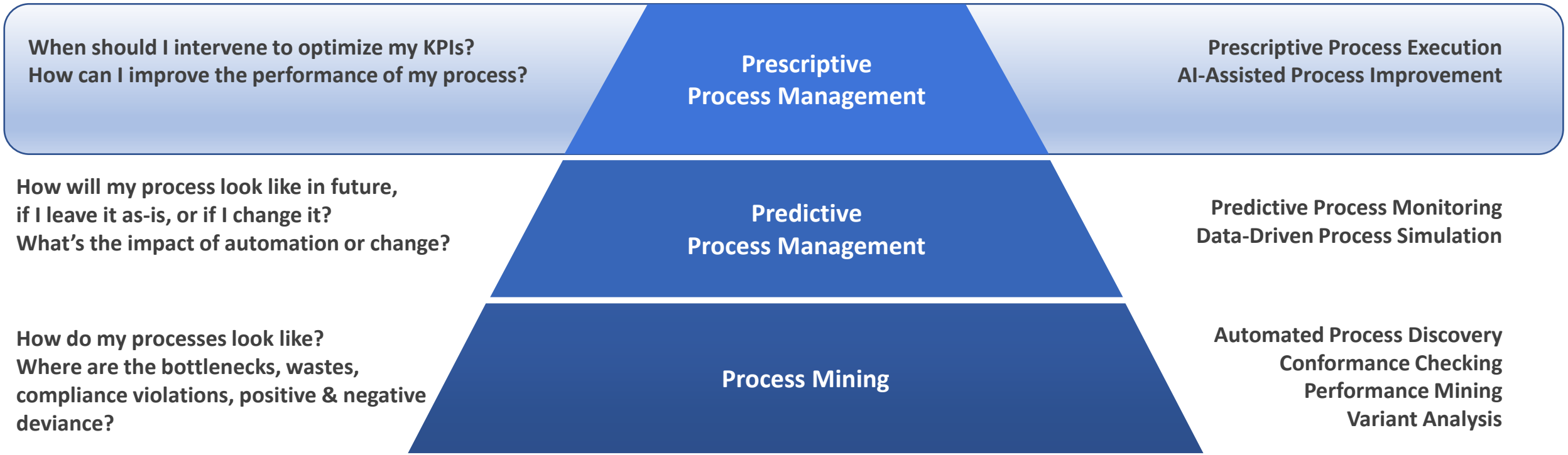
(a) *Situation faced*: The Dutch social security institute, UWV, is responsible for providing benefits to people who have lost their jobs. When such a person provides UWV with incorrect information about their income situation, whether intentionally or not, he or she sometimes receives too much in benefits, in which case the person must repay the excess amount. Handling these situations takes up UWV resources and can be problematic for the benefit recipient. UWV wants to prevent customers from having to repay part of their benefits when the error in reporting income was unintentional.

(b) *Action taken*: An intervention was selected, based on human judgment and subjective opinions, to inform benefit recipients ("customers" hereafter) via e-mail how to provide

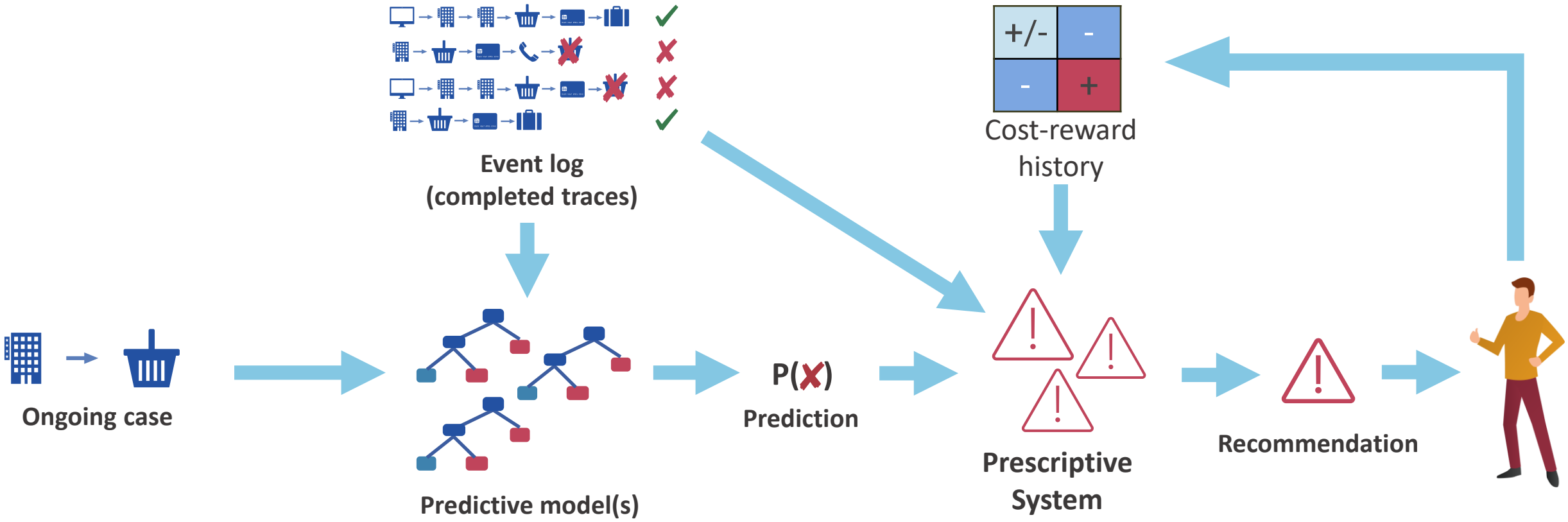
Results achieved: Even though the prediction of the most likely customers to report incorrectly was reasonably accurate, the intervention did not have a preventive effect. No root cause was identified to explain why the intervention did not have the desired effect.



Predicting is useful, preempting is better!



Prescriptive Process Execution



Where does GenAI fit in all this?

- Everywhere!
- GenAI brings context recognition: What types of processes, activities, KPIs are we talking about?
- LLMs enable conversational process optimization across all layers of the pyramid

Descriptive Process Mining

- Where are the sources of waiting time?
- Where are the rework loops?
- Where are we over-processing?
- What are the sources of variance?
- Which cases require the most touches?
- Where are we violating our KPIs?
- Are we abiding to our business rules and policies?
- …

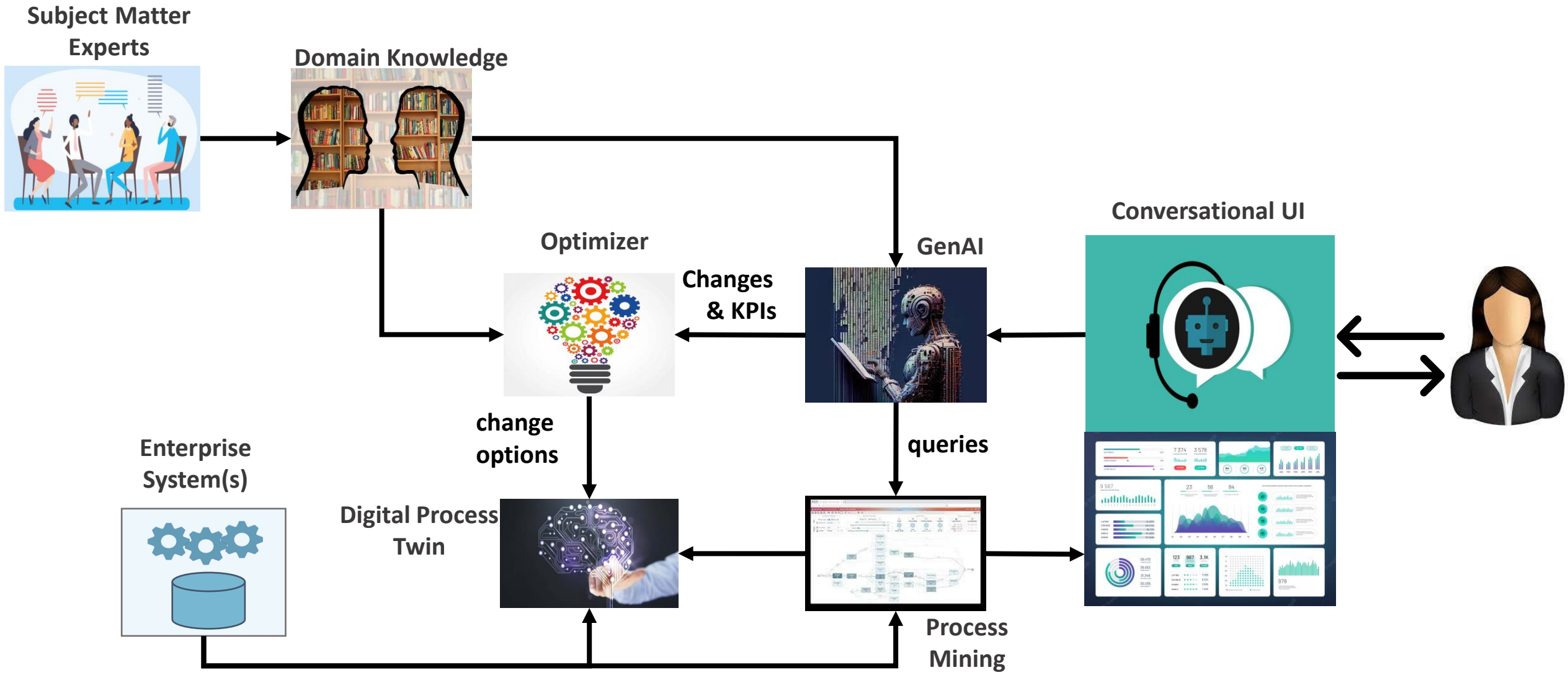
Predictive Process Optimization

- By how much would we reduce order-to-delivery times is we:
  - Shorten inter-batch cycles?
  - Move resources to packaging?
  - Automate verification steps?
- By how much would we reduce costs if we:
  - Consolidate touchpoints?
  - Reorder verification steps to reduce over-processing?
  - Reduce rework rates?
  - …

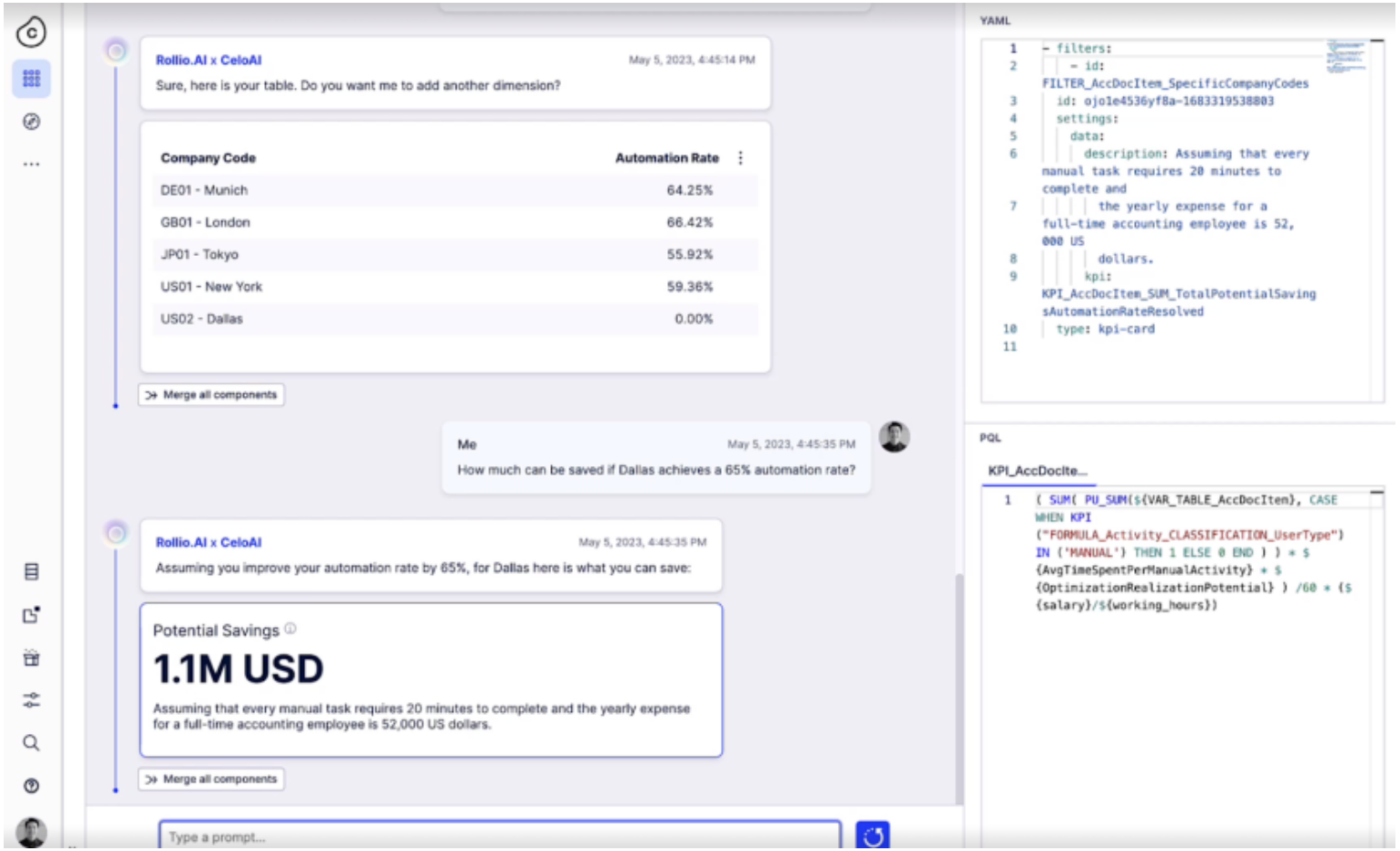
Assisted Process Optimization

- How can we slash KPI violation rates?
- What are the practices of the teams with highest performance?
- What should we change to reduce the number of touches?
- Which checks should we add to reduce compliance violations?
- How can we cut the cycle time at constant cost?
- How to allocate resources to optimize time at constant capacity?
- …

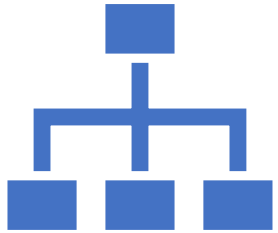
Conversational Process Optimization



First steps in tooling



Pitfalls of prescriptive process optimization



Organizational Pitfalls

Lack of buy-in from operations

Lack of trust in suggestions / prescriptions

Prescription flooding / over-prescription

Prescriptions not linked with actions

Lack of processes to validate, monitor, and maintain prescriptive models

Technical Pitfalls

Insufficient data availability & quality

Neglecting inter-process dependencies

Lack of uncertainty modeling

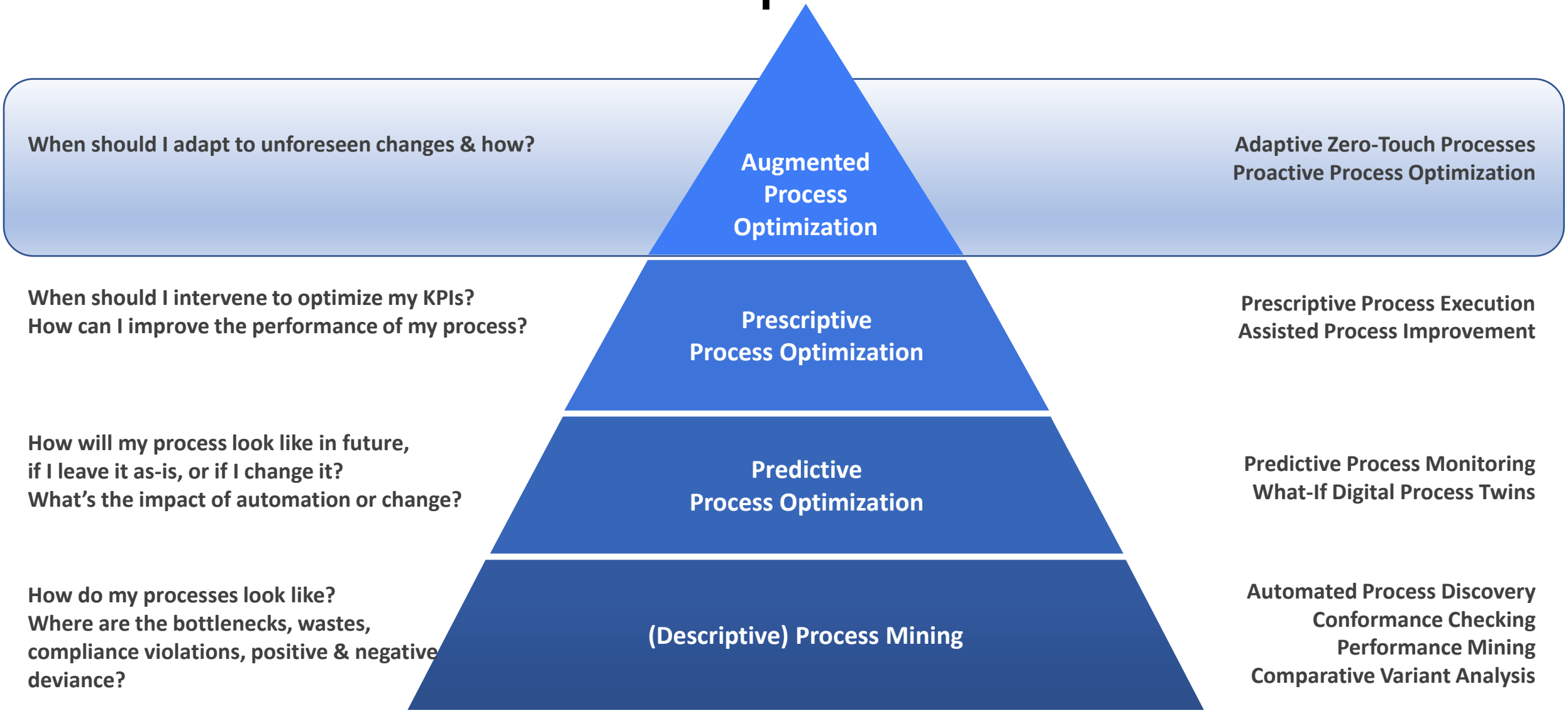
Drifts and out-of-distribution predictions

Unreliable prescriptions

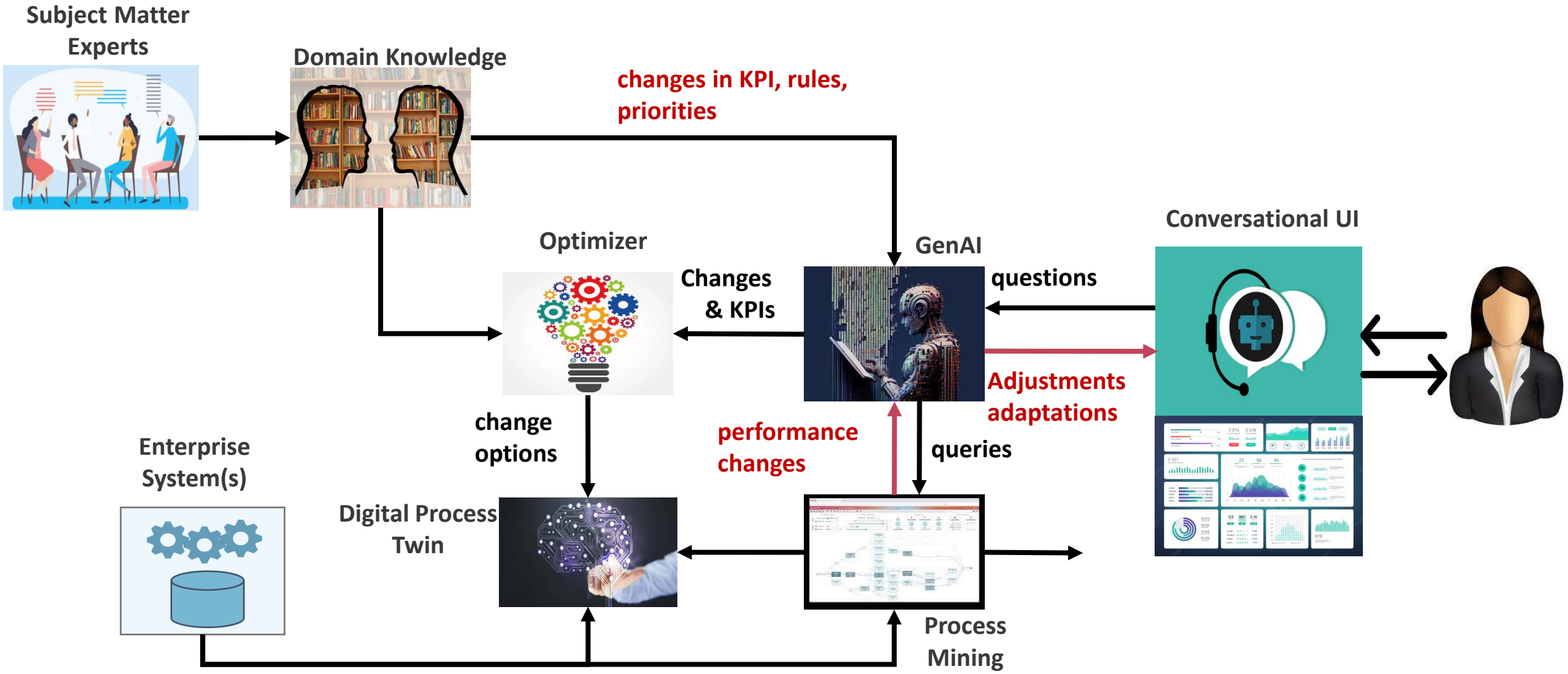
Lack of feedback loop

No amount of organizational and technical readiness will save you from the sin of using AI-driven improvement recommendations without validation and pilot testing or prescriptive models without A/B testing

AI-Driven Process Optimization

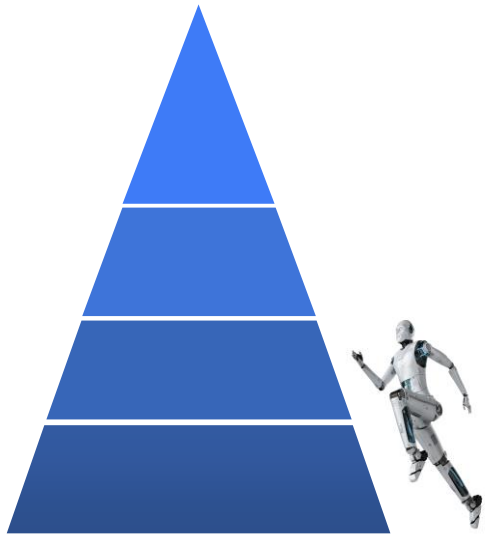


Proactive Process Optimization



Climbing the AI-Driven Process Optimization Pyramid

- Lay the foundations, start climbing, keep climbing, don't hold off
  - Getting data for process mining is often a challenge. But there are both short-term benefits (bottom of the pyramid) and long-term ones (top)
- Don't skip the layers
  - The lower layers of the pyramid provide a foundation to draw business value from the upper layers.
- Align strategically and build governance incrementally
  - Apply these capabilities first and foremost to processes that matter
  - Adopt these capabilities incrementally, one process at a time
  - Build success stories internally, ensure each layer of the pyramid yields value



Research at KU Leuven

- Industry collaborations with

- Research themes
  - Seq2Seq LSTM and transformer models for predictive process monitoring
  - Process model forecasting
  - XAI in predictive process monitoring
  - Intercase featurization for predictive process monitoring
  - Timed interventions in prescriptive process monitoring
  - Testbed development for prescriptive process monitoring
  - IoT & process mining

Interessante links

- ICPM conference 2024 (Copenhagen, 14-18 oktober)
  - https://icpmconference.org/2024/
- BPM conference 2024 (Krakow, 1-6 September)
  - https://bpm2024.agh.edu.pl/
- IEEE Taskforce on Process Mining
  - https://www.tf-pm.org/
  - Become a member!

Bedankt!

Vragen?

Contact: jochen.deweerdt@kuleuven.be