Simona Fioretto

A Process-Aware Perspective on the Challenges and Solutions of Public Administration Digital Transformation

Tutor: Elio Masciari

Cycle: XXXVIII   Year: III

Candidate's information

- MSc degree in Management Engineering @DII – Federico II

- DIETI Research group/laboratory:
  - PICUS Lab

- PhD start date – end date:
  - 01/11/2023 – 31/10/2025

- Scholarship type:
  - PNRR - DM 351 Mis.: I.4.1 Dottorati Pubblica Amministrazione

- Periods abroad:
  - 03/06/2024 – 02/12/2024

- Abroad research institution:
  - Centre de coopération internationale en recherche agronomique pour le développement (CIRAD)

Summary of study activities

- Ad hoc PhD courses / schools:
  - Machine Learning for Science and Engineering Research - Proff. Anna Corazza, Roberto Prevete, Carlo Sansone, Antonio Lieto
  - I pilastri della trasformazione digitale - Prof. Francesco Tortorelli
  - Statistical Data Analysis for science and engineering research Prof. Roberto Pietrantuono
  - How to Boost Your PhD - Prof. Antigone Marino
  - Strategic Orientation for STEM Research & Writing Dr Chie Shin Fraser
  - Summer School: Software Engineering @Salerno Week

- Attended Conferences:
  - 31st Symposium on Advanced Database System, Padua, Italy, Jul. 2023
  - European Conference on Advances in Databases and Information Systems (ADBIS 2023), Barcellona, Spain, Sep. 2023
  - 2024 International Symposium on Methodologies for Intelligent Systems (ISMIS 2024), Poitiers, France, Jun. 2024

Research area(s)

The research is positioned within the **Information Systems** field, addressing the **digital transformation** of **Public Administration** by leveraging **process-aware** and **data-driven** methods to realign public service processes with modern digital technologies.

Research results

- **Business Process Management (BPM) Custom Application to Public Administration**

- **Systematic Literature Review on Predictive Process Monitoring (PPM)**

- **Process Mining on Land Matrix Data**

Research products

[P1]	A. Anniciello, S. Fioretto, E. Masciari, E.V. Napolitano *Covid-19 impact on health information technology: the rapid rise of e-Health and Big Data driven innovation of healthcare processes,* **International Conference on IEEE International Conference on Bioinformatics and Biomedicine (BIBM),** Las Vegas, NV, USA, Dec. 2022, pp. 2759-2764, IEEE, DOI: 10.1109/BIBM55620.2022.9995515
[P2]	F. Amato, S. Fioretto, E. Forgillo, E. Masciari, N. Mazzocca, S. Merola, E. V. Napolitano *Introducing AI-Based Techniques in the Justice Sector: A Proposal for Digital Transformation of Court Offices,* **31st Symposium on Advanced Database System,** Padua, Italy, Jul. 2023, vol. 3478, pp. 497-504, CEUR-WS, EID: 2-s2.0-85173489342
[P3]	S. Fioretto *Business Process Extraction from Documents with AI,* **31st Symposium on Advanced Database System,** Padua, Italy, Jul. 2023, vol. 3478, pp. 697-702, CEUR-WS, EID: 2-s2.0-85173553023
[P4]	S. Fioretto, E. Masciari, E.V. Napolitano *Can the Study of Trajectories Help to Extract Information from Business Processes?* **2nd International Workshop on Process Management in the AI era,** Macao, SAR, Sep. 2023, vol. 3569, pp. 13-24, CEUR-WS, EID: 2-s2.0-85180158158

Research products

[P5]	A. Anniciello, S. Fioretto, E. Masciari, E.V. Napolitano *Digital Twins for Traffic Congestion in Smart Cities: a novel solution using Data Mining techniques,* **Proceedings of the 15th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2023),** Rome, Italy, Nov. 2023, vol. 3, pp. 241.248, DOI: 10.5220/0012208100003598
[P6]	A. Anniciello, S. Fioretto, E. Masciari, E.V. Napolitano *How Pandemic Affected the Adoption of e-Health Systems,* **IDEAS '23: Proceedings of the 27th International Database Engineered Applications Symposium** Heraklion, Crete, Greece, May 2023, pp. 94-98, DOI: 10.1145/3589462.3589496
[P7]	S. Fioretto, E. Masciari, E.V. Napolitano *Dossier classification to support workflow management optimization,* **The 4th International Conference and Summer School "Numerical Computations: Theory and Algorithms" (NUMTA 2023)** Pizzo Calabro, Italy, Jun. 2023, vol. 14478, pp. 286-292, DOI: 10.1007/978-3-031-81247-7_25
[P8]	F. Amato, S. Fioretto, E. Forgillo, E. Masciari, N. Mazzocca, S. Merola, E. V. Napolitano *Evolving Justice Sector: An Innovative Proposal for Introducing AI-Based Techniques in Court Offices,* **12th International Conference on Electronic Government and the Information Systems Perspective (EGOVIS 2023)** Penang, Malaysia, Aug. 2023, pp. 75-88, DOI: 10.1007/978-3-031-39841-4_6

itee PhD
information technology
electrical engineering

Research products

[P9]	S. Fioretto *Process Mining Solutions for Public Administration,* **European Conference on Advances in Databases and Information Systems (ADBIS 2023),** Barcellona, Spain, Sep. 2023, vol. New Trends in Database and Information Systems, pp. 668-675, DOI: 10.1007/978-3-031-42941-5_60
[P10]	S. Fioretto, E. Masciari, E.V. Napolitano *A Joint Analysis of Trajectory Mining and Process Mining for Smartphone User Behaviour,* **Joint European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases. ECML PKDD 2023,** Turing, Italy, Sep. 2023, vol. CCIS, 2135, Springer, Cham, DOI: 10.1007/978-3-031-74633-8_43
[P11]	S. Fioretto, E. Masciari, E.V. Napolitano *Machine Learning for KPI Development in Public Administration,* **Proceedings of the 13th International Conference on Data Science, Technology and Applications, DATA 2024** Dijon, France, Jul. 2024, pp. 522-527, SciTePress, DOI: 10.5220/0012820300003756
[P12]	S. Fioretto, D. Ienco, R. Interdonato, E. Masciari *Integrating predictive process monitoring techniques in smart agriculture,* **2024 International Symposium on Methodologies for Intelligent Systems (ISMIS 2024)** Poitiers, France, Jun. 2024, LNAI,vol. 14670 pp. 306-313, Springer, Cham, DOI: 10.1007/978-3-031-62700-2_27

Research products

[P13]	S. Fioretto, E. Masciari *A Conceptual Framework for Predictive Process Monitoring in Public Administration,* **2024 International Conference on Complex, Intelligent, and Software Intensive Systems (CISIS 2024)** Taichung, Taiwan, Sep. 2024, LNDECT,vol. 87, pp. 423-431, Springer, Cham, DOI: 10.1007/978-3-031-70011-8_40
[P14]	M.L. Conza, S. Fioretto, E. Masciari, E.V. Napolitano *Enhancing Employee Health Through an Experimental Diet: Insights from Machine Learning Analysis,* **33rd Euromicro international Conference on Parallel, Distributed, and Network-Based Processing (PDP)** Turin, Italy, Mar. 2025, pp. 283-290, IEEE, DOI: 10.1109/PDP66500.2025.00046
[P15]	S. Fioretto, E. Masciari, *A comparative analysis of predictive process monitoring: object-centric versus classical event logs,* **International Journal of Knowledge and Information Systems,** vol. 67, pp. 7355-7398, 2025, DOI: 10.1007/s10115-025-02461-y.
[R16]	S. Fioretto, D. Ienco, R. Interdonato, J. Bourgoin, E. Masciari *Part of the Process: Analyzing Land Acquisition Dynamics Through Process Mining* **Under review in IEEE Access**

itee**PhD**
information technology
electrical engineering

PhD thesis overview

Public administrations struggle to deliver efficient services,
limited by confusion between digitization, digitalization, and true digital transformation.

Digitization

conversion of analog information into digital form, enabling its storage and processing by computers.

Digitalization

application of digital technologies to improve workflows, automating and streamlining existing processes.

Digital Transformation

involves the strategic use of digital technologies and process knowledge to reconfigure operations, enabling end-to-end proess integration.

Objective

Adopt a **process-aware** and **data-driven** approach, utilizing top-down and bottom-up approaches:
* Business Process Management (BPM), Process Mining (PM) and Predictive Process Monitoring (PPM)

PhD thesis overview: Methodology

Application of PM and PPM to Land Matrix Data
Assess the applicability of PM and PPM in noisy, unstructured, and highly complex process environments.

O3

**Systematic Literature Review on Predictive Process Monitoring (PPM)**
Examine applications on **classical** and **object-centric event logs** to identify trends in approaches.

O2

Empirical Analysis of Public Administration
Exploration of the PA context through top-down **BPM** to analyze:
•Processes
•Data and information systems in use

OI

itee PhD
information technology
electrical engineering
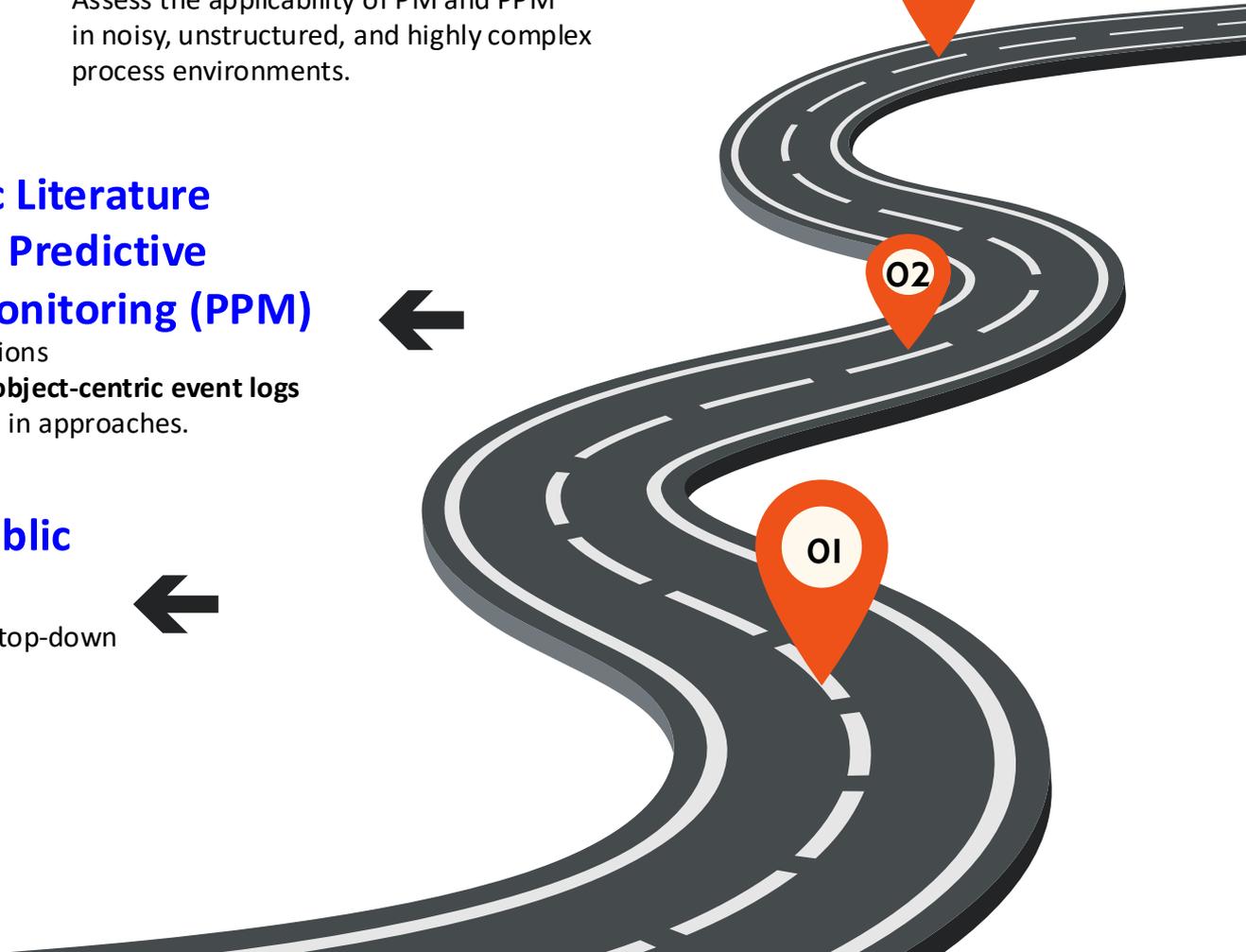
Research Line 1

An Empirical Analysis of Digital Transformation in Judicial Offices Using Business Process Management (BPM) Methodologies

Motivation & Context

Context - Court of Appeal of Naples

- Complex judicial body managing high case volumes across civil, criminal, and labour divisions.

- Daily administrative activities — **staff scheduling, leave management, and workload distribution** — directly impact overall efficiency.

Motivation

- investigate how BPM methods can support digital transformation in environments characterized by **low technological maturity, fragmented data, and legacy infrastructures**

Methodology

No.	Name	Type	Description
1	Leave Management	Process	Describes how leave requests are inserted, verified, approved, and recorded.
2	Workload Assignment	Process	Records and assignment updates workload data, and monitors the availability of judges

Process Identification

Process Discovery

Process Analysis

Process Redesign

Process Implementation

Process Monitoring

BPM life cycle

technical and institutional constraints

Problem (As-Is)	Redesign Action (To-Be)
Manual management of operational data	Introduction of digital forms and automated validation for leave requests and workload updates
Dispersed and non-integrated data sources	Creation of a shared database connecting leave, workload, and section data
Limited visibility of availability and workload	Develop interactive dashboards to provide real-time and planned availability.
Absence of automated allocation support	Introduction of digital tools assisting case distribution, based on up-to-date data on workload and availability
Lack of integration between related processes	Connection of Leave Management and Workload Assignment modules through shared and synchronized data

Simona Fioretto

14

Proposed Solution & Results

- **Adapted BPM Framework**

- **Demonstrated** the applicability of BPM in low-digital-maturity contexts.

- Identified an **object-centric structure** where *Leave Management* and *Workload Assignment* share the same core entity — *judge availability*.
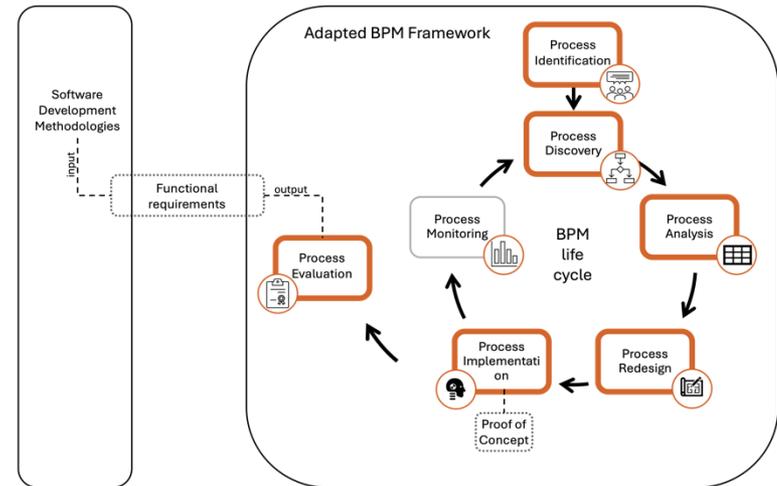
| Leave Management Process	Workload Assignment Process
Digital data-entry form for structured registration of leave requests.	**Structured personnel registry** for judges' profiles, roles, and assignments.
Automated validation **checks** to detect duplicates and overlapping periods.	**Real-time workforce overview** showing active, assigned, and unavailable judges.
Centralized historical archive of all recorded leaves for traceability and reporting.	**Assignment management module** with task creation, update, and completion tracking.
Summary reporting functions to monitor extended or accumulated leaves.	**Integrated monitoring tables** for ongoing and completed assignments.
Data synchronization across sheets	**Reporting and summary visualization**
Configurable calendar view for period-based visualization and planning.	—

Research Line 2

Comparative Analysis of Predictive Process Monitoring Approaches (Classical vs. Object-Centric Event Logs)

Motivation

- ## **Predictive Process Monitoring**
  - area of Process Mining which predict the future behaviour of an ongoing process instance
  - rely on **single-ID Event Logs**, which fail to capture the complexity of real-world, multi-entity processes.
  - modern systems produce **Object-Centric Event Logs (OCEL)** linking multiple related objects and interactions.
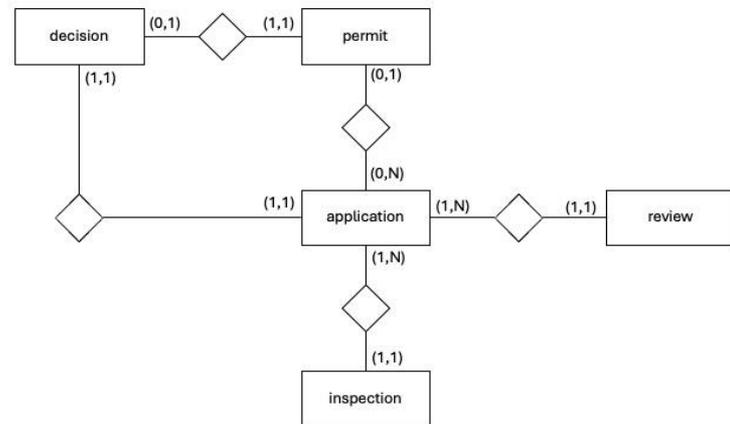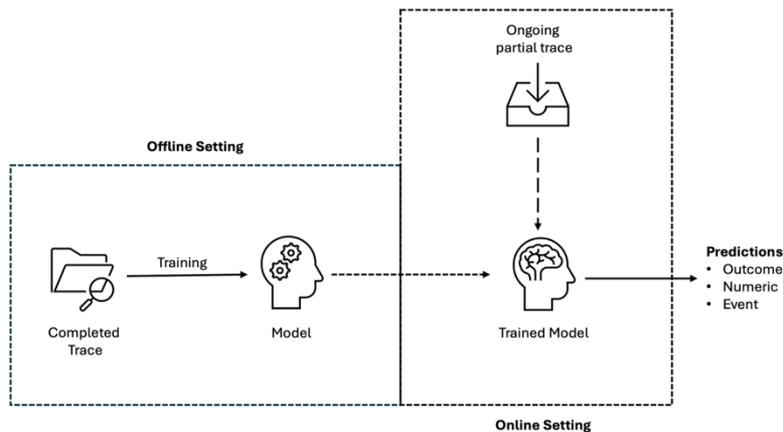
Methodology

- **Approach:** Systematic Literature Review
- **Databases:** Scopus, IEEE Xplore, Google Scholar (2019–2024)
- **Research Questions:**
  - RQ1: *What is the nature of the input Event Logs in PPM approaches?*
  - RQ2: *How do the prediction tasks, the methods used, and the application domains differ based on the type of Event Log?*
  - RQ3: *To which phase of the PPM framework does the scientific contribution of each research proposal specifically relate?*
- **47 papers selected** after inclusion/exclusion screening
- **Dimensions analyzed:** Event Log Type, Prediction Task, Method, Contribution Area, Domain

Results

Dimension	Key Insights
Event Log Type	**Classical:** single-case logs; standardized, efficient, widely supported. **OCEL:** multi-object logs; capture interactions, richer context but higher complexity. Three strategies: **Flattening** (simple, compatible, good with feature engineering), **Graph-based** (GNNs, embeddings accurate, costly), **Hybrid/Alternative** (Petri Nets, conceptual).
Prediction Task	• Classical: next event / remaining time • OCEL: multi-target (next event + KPI + anomaly)
Method	• OCEL: GNNs, LSTM, CatBoost (accuracy vs cost) • Classical: LSTM, Transformers, CNN
Contribution Area	• OCEL: focus on data modeling & encoding • Classical: focus on model tuning & benchmarks
Domain	• OCEL: complex sectors (Public Admin, Healthcare, Finance) • Classical: structured domains with public datasets Common domain: Finance, useful for cross-method benchmarking.

Research Line 3

Application of Process Mining and Predictive Process Monitoring in Large Scale Land Acquisitions

Motivation

Context

- Large-Scale Land Acquisitions (LSLAs) raise issues of transparency, governance, and local rights, while their dynamics remain highly variable and opaque.

Objective

- Use **Process Mining** and **ML** to explore LSLA behavior and test the applicability of PM/PPM in unstructured, noisy, real-world data environments.

Methodology

Phase	Objective	Techniques	Outcome
1. Data Preprocessing & **Event Log Creation**	Transform raw Land Matrix tables (Deals, Investors) into structured event logs	Extraction of Negotiation/Implementation status sequences, cleaning, timestamp parsing.	2,198 cases · 5,138 events · 13 activities → usable event log.
2. **Process Discovery & Inspection**	Model land acquisition behavior and identify process variability.	Directly-Follows Graphs (DFG) via PM4Py, variant analysis, temporal inspection.	Highly unstructured spaghetti process → need for complexity reduction.
3. **Complexity Reduction**	Simplify and segment behavior while preserving rare but meaningful cases.	Log decomposition (by end-status) + K-Means & Agglomerative Hierarchical clustering (index-based encoding).	Sub-logs capturing homogeneous behavioral clusters.

Deal ID	Intention of investment	Negotiation status	Implementation status
10362	2022#current#7818#For carbon sequestration/REDD	2016#current#Concluded (Contract signed)	2022#current#In operation (production)
10363	2021-06-01#current#2863#For carbon sequestration/REDD	2014#current#Concluded (Contract signed)	2021-09#current#In operation (production)
10364	2014#current#165000#For carbon sequestration/REDD	2014#current#Concluded (Contract signed)	2014#current#In operation (production)

Deal ID	Activity	Timestamp
9958	Failed (Negotiations failed)	2023-01-01
9958	Project abandoned	2023-01-01
9962	Intended (Memorandum of understanding)	2019-05-22
9962	Project not started	2020-01-01
9990	Failed (Contract cancelled)	2022-01-01
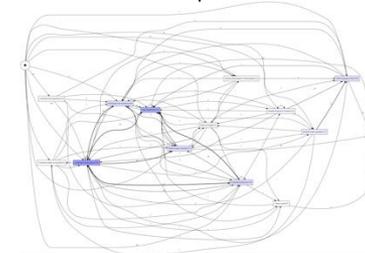
	event log						DFG		
	# traces	# variants	variants / traces	# events	min. trace length	max. trace length	#vertices	#edges	|E| / |V|
Failed Sub-Log	262	86	0.33	760	1	8	13	72	5.54
C0_AHC	98	18	0.18	199	2	3	10	18	1.8
C1_AHC	13	13	1	76	5	7	12	34	2.83
C2_AHC	93	25	0.27	279	3	3	13	32	2.46
C3_AHC	47	27	0.57	188	4	4	13	42	3.23
C0_KM	59	39	0.66	260	4	7	13	55	4.23
C1_KM	96	28	0.29	288	3	3	13	35	2.69
C2_KM	96	16	0.17	194	2	4	11	17	1.55

Methodology & Results

Phase	Objective	Techniques	Outcome
4. Predictive Analysis	Assess whether process information improves outcome prediction.	Random Forest classifier trained on (i) attributes only vs (ii) attributes + activities.	Control-flow info gives marginal gain due to data noise and class imbalance.
5. Encoding Comparison	Test impact of different encoding strategies on performance.	Ordinal Encoding, BERT Embeddings, MLP Transformed Features.	BERT + MLP yields richer representations but limited accuracy gain.

Model	Accuracy	Balanced Accuracy[1]	Precision	Recall	F1 Score[2]
O_attr	0.58	0.55	0.58	0.58	0.57
O_act_attr[3]	0.55	0.51	0.57	0.55	0.54
BERT+O_act_attr[4]	0.62	0.59	0.62	0.62	0.61
BERT+MLP_act_attr[5]	**0.63**	**0.61**	**0.63**	**0.63**	**0.63**

- Empirical demonstration that control-flow info, when properly encoded (BERT+MLP), enhance outcome prediction, even in highly unstructured data

Conclusion

- Business Process Management can effectively support the analysis and redesign of complex workflows within the Public Administration domain.

- Process Mining techniques can extract actionable insights even from incomplete, heterogeneous, and unstructured data sources.

- Predictive Process Monitoring can be applied to real-world, low-quality datasets, enabling early outcome prediction despite data irregularities.

Thank You For Your Attention