



chist-era



CHIST-ERA Projects Seminar 2021

2017-BDSI: Project PACMEL

***Grzegorz J. Nalepa (AGH, Poland), Szymon Bobek (AGH),
Victor Rodriguez-Fernandez (UPM, Spain)***

March 29, 2022



NARODOWE CENTRUM NAUKI



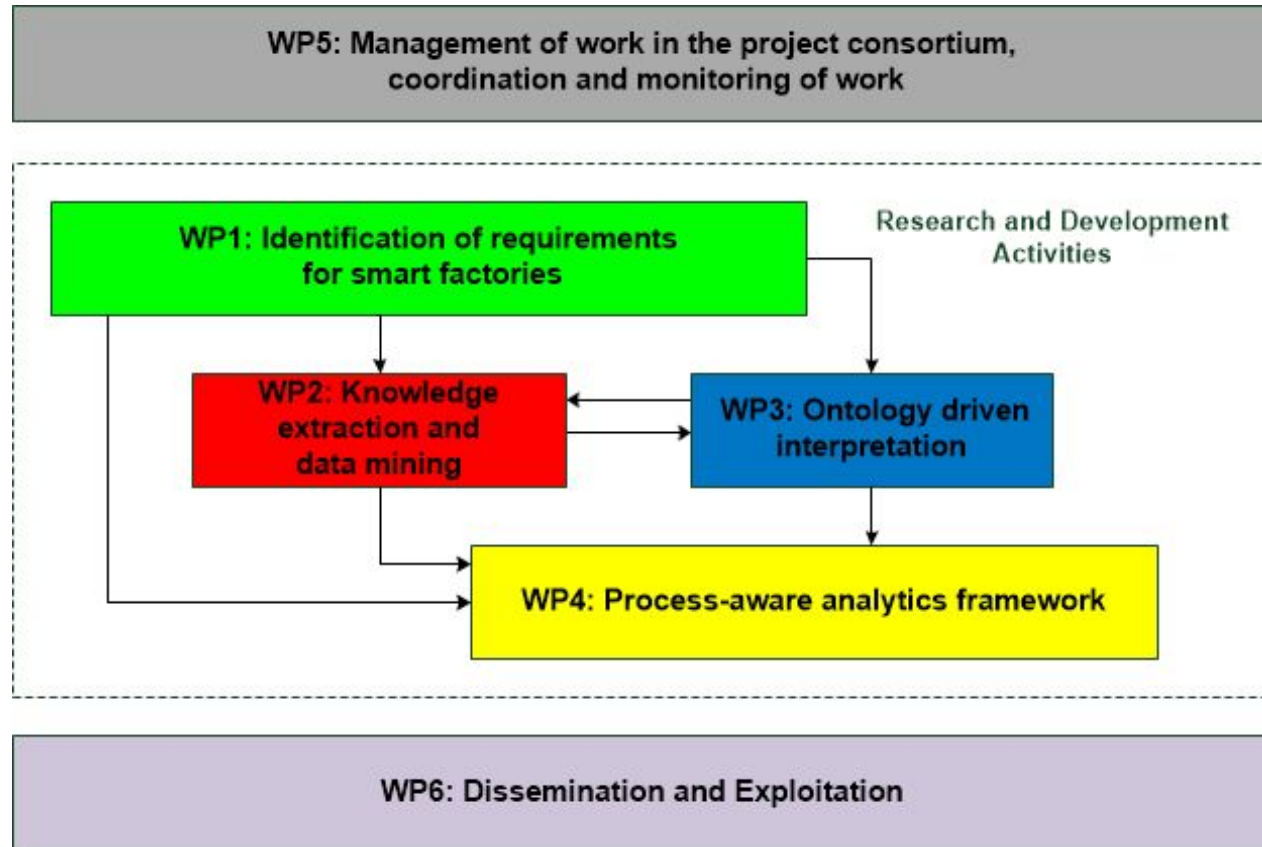
Programme co-funded by the
EUROPEAN UNION

Scientific background: PACMEL Project Challenges&Impacts

Process-aware Analytics Support based on Conceptual Models for Event Logs

1. How can we extract relevant events from *different dataset formats* (unstructured/structured) to discover the *industrial/business processes* being executed?
2. How can we combine relevant events to *discover complex* industrial/business processes operating in a networked ecosystem?
3. How can we efficiently support *process analysis and modeling* for management purposes using sensor data-based event logs?
1. Improved process mining and knowledge extraction techniques enabling flexible multi-level process analytics and process model creation.
2. The integration of various heterogeneous data sources by employing ontologies to support semantically enriched search of information.
3. Process design support by recommendation methods of process modelling notation according to industrial needs and for the purpose of management.
4. Process monitoring support using time-series aware process mining methodologies.

Introduction: PACMEL Project Structure



AGH, Kraków, Poland - G.J.Nalepa

- ❖ Applied Computer Science: S. Bobek, M. Szpyrka
- ❖ Geoengineering: E. Brzychczy
- ❖ GEIST.re research group
- ❖ Expertise: data mining, knowledge representation, eXplainable AI

UniBZ, Bolzano, Italy - D. Calvanese

- ❖ Intelligent Integration and Access to Data (In2Data) research group
- ❖ Research Centre for Knowledge and Data: M. Montali
- ❖ Expertise: Ontology-based data access and integration

UPM, Madrid, Spain - D. Camacho

- ❖ Department of Computer Systems Engineering: V. Rodriguez-Fernandez
- ❖ Applied Intelligence and Data Analysis (AIDA) group
- ❖ Expertise: Machine/Deep learning, Computational intelligence

AGH Uni. of Science and Tech, Kraków - G.J.Nalepa

- ❖ AGH is one of the leading Polish technical universities
 - ❖ Dept. Applied Computer Science: S. Bobek, M. Szpyrka
 - ❖ Dept. Geoengineering: E. Brzychczy
 - ❖ GEIST research group
- 
- The logo of AGH University of Science and Technology, featuring three vertical bars in green, black, and red, with the letters "AGH" in black below them.
- ❖ Project leader: WP 1,5 (6)
 - ❖ Expertise: data mining, knowledge representation, eXplainable AI
 - ❖ Provided two industrial partners
 - ❖ Developed methods and tools for
 - ✓ Incorporation and use of expert knowledge in data mining inc. clustering
 - ✓ Event log analysis
 - ✓ Formal verification and conformance checking

Uni. Politecnica de Madrid, Madrid - D. Camacho

- ❖ One of the leading technical universities in Spain
- ❖ Department of Computer Systems Engineering: V. Rodriguez-Fernandez
- ❖ Applied Intelligence and Data Analysis (AIDA) group



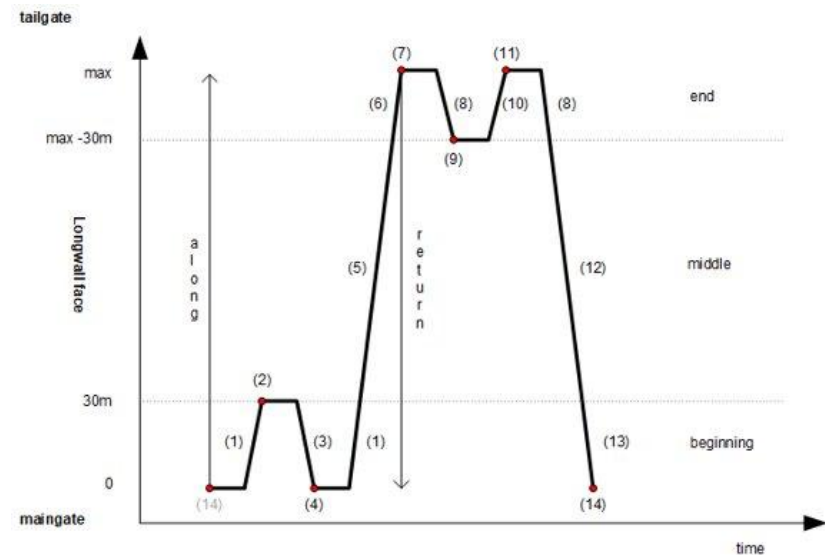
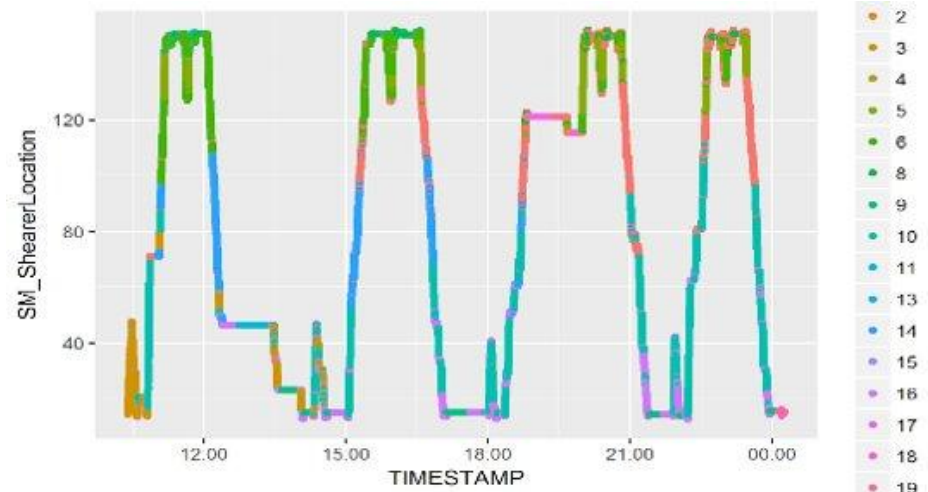
- ❖ Key project partner: WP 2,4
- ❖ Originally project was granted to UAM, then moved to UPM
- ❖ Expertise: Machine/Deep learning, Computational intelligence
- ❖ Developed methods and tools for
 - ✓ Conformance checking
 - ✓ Data analysis and visualization based on deep learning

Use case: underground mining - Famur



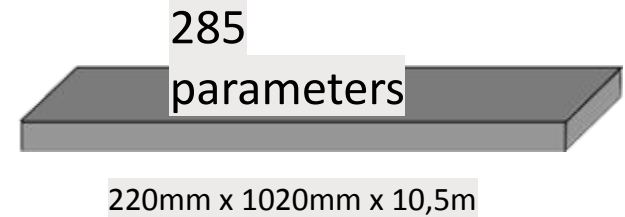
Use case: underground mining - Famur

- ❖ Data was delivered by project partner the Famur S.A. company.
- ❖ Coal shearer moves along longwall excavating coal
- ❖ About 300 different measurements every second
- ❖ Noisy data with lots of missing values
- ❖ Expert rules for extracting basic machinery states

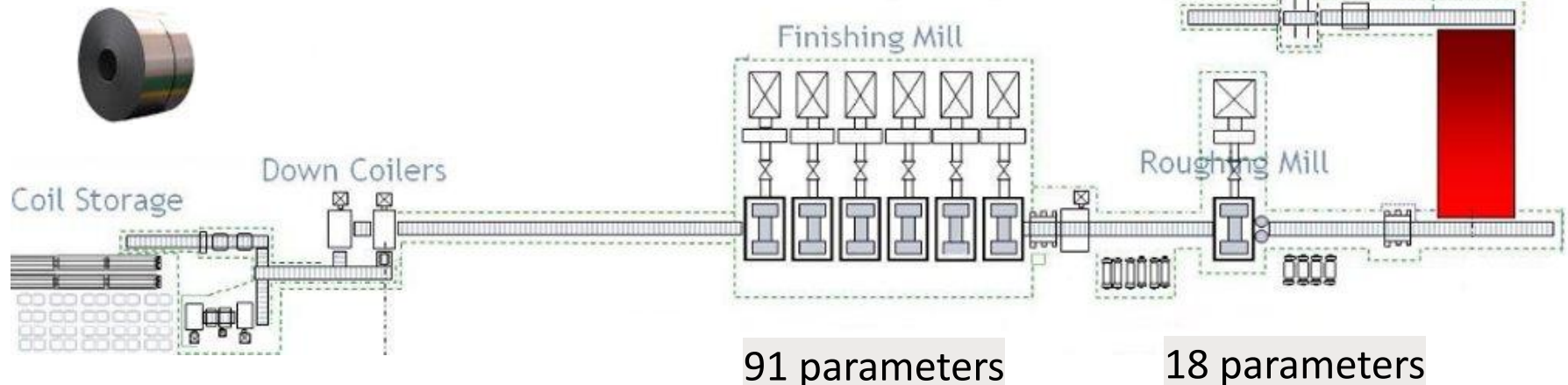


Use case: steel hot rolling ArcelorMittal

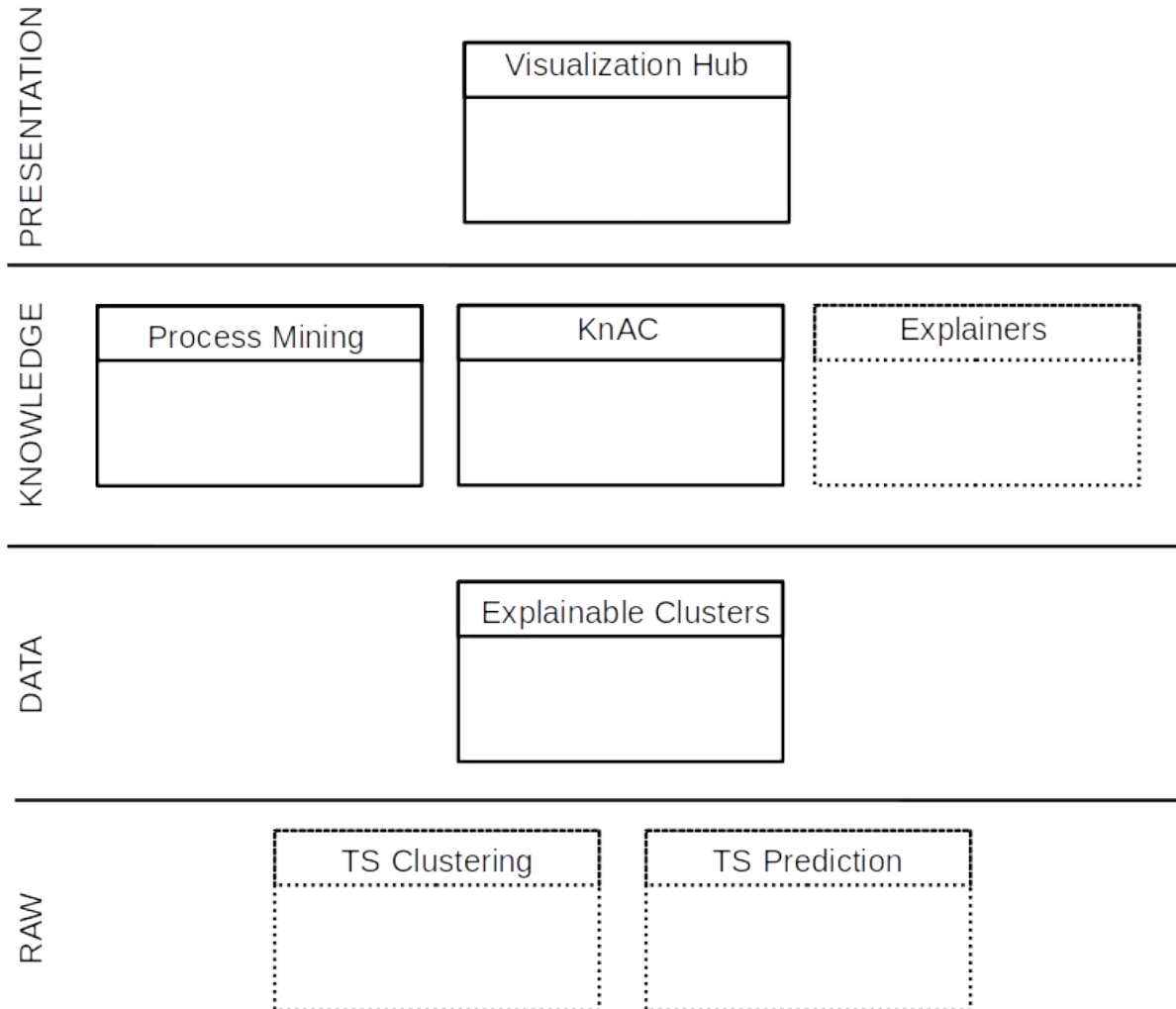
- Fully automated process
- Hundreds of parameters involved
- Approximately 40 000 measurements per product
- Place for improvements:
 - Energy savings
 - Predictive maintenance
 - Defect minimization



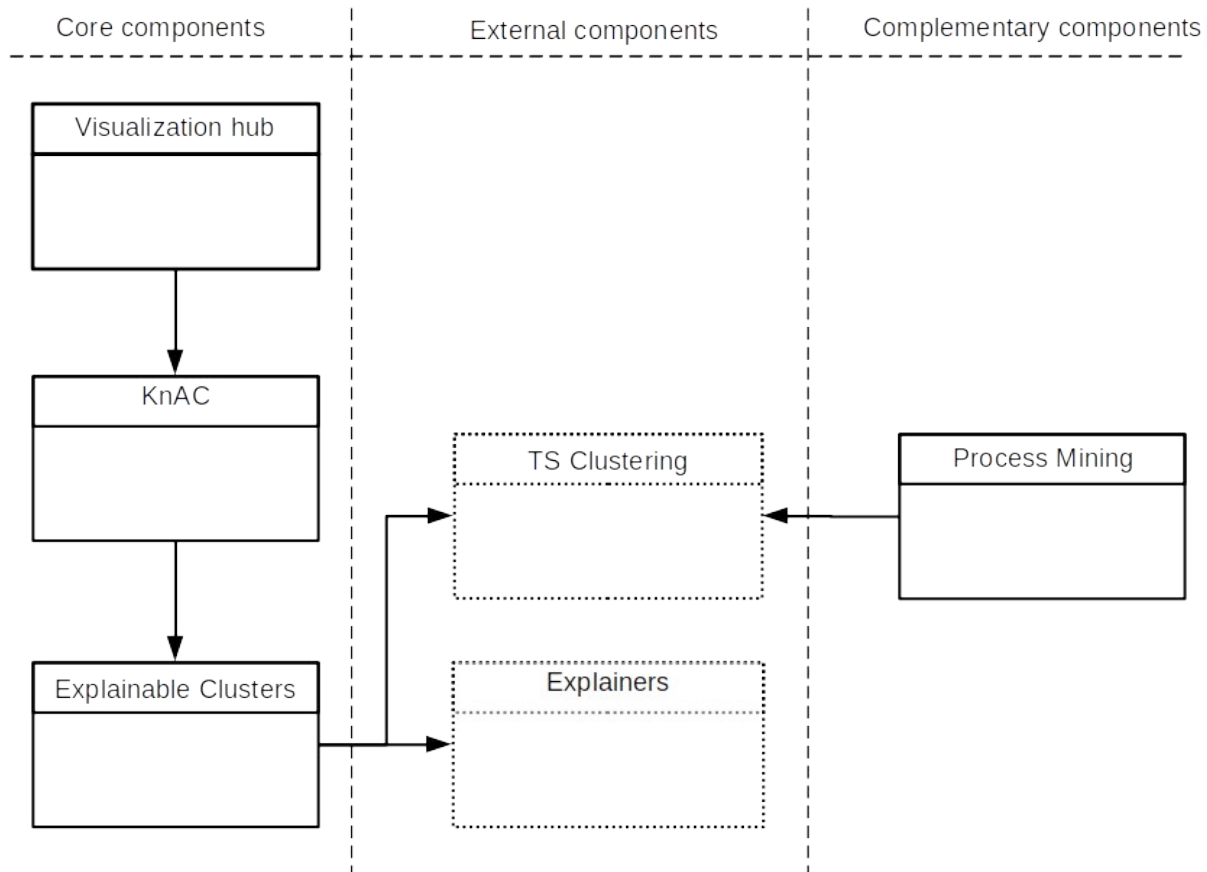
2.3mm x 1020mm x 970m



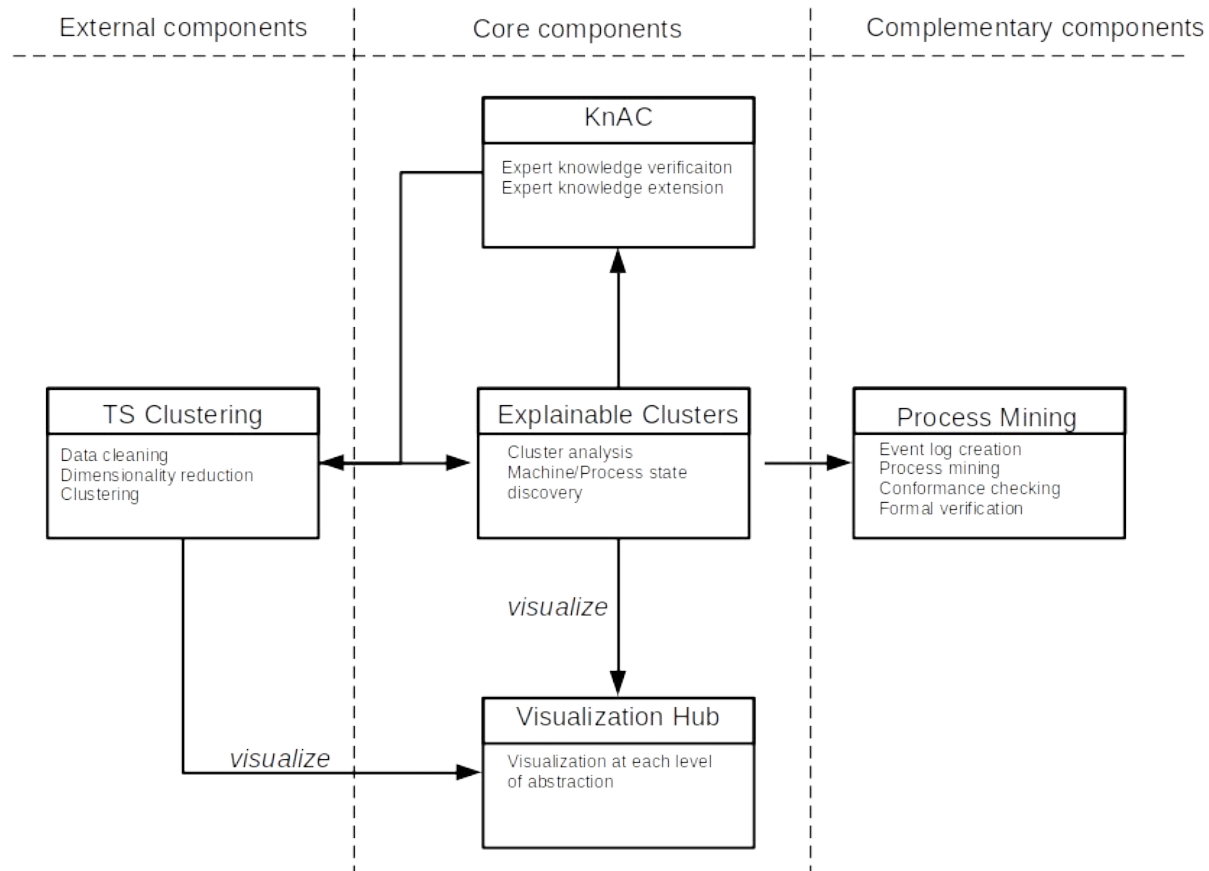
Results - Pacmel Framework



Results - framework architecture



Knowledge discovery perspective



Timecluster hub (UPM)

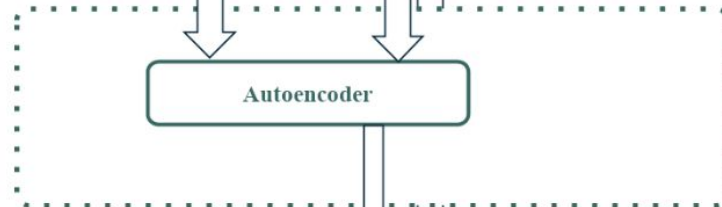
- ❖ Hub for different visual analytics approaches for high-dimensional time series.
- ❖ Inspired by the paper "Timecluster: dimension reduction applied to temporal data for visual analytics"
- ❖ Backend: Deep learning autoencoders + Dimensionality reduction
- ❖ Frontend: Interactive visualizations with the original and embedded space
- ❖ Source code will be publicly available.

Visualization (UPM) - Timecluster hub

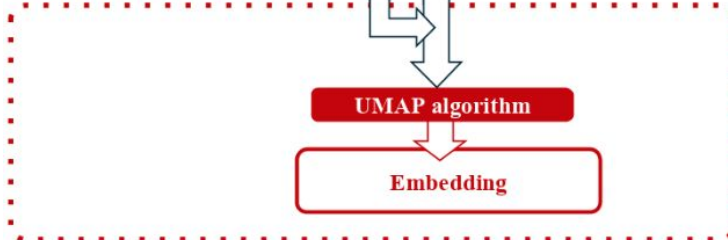
1. Dataset creation



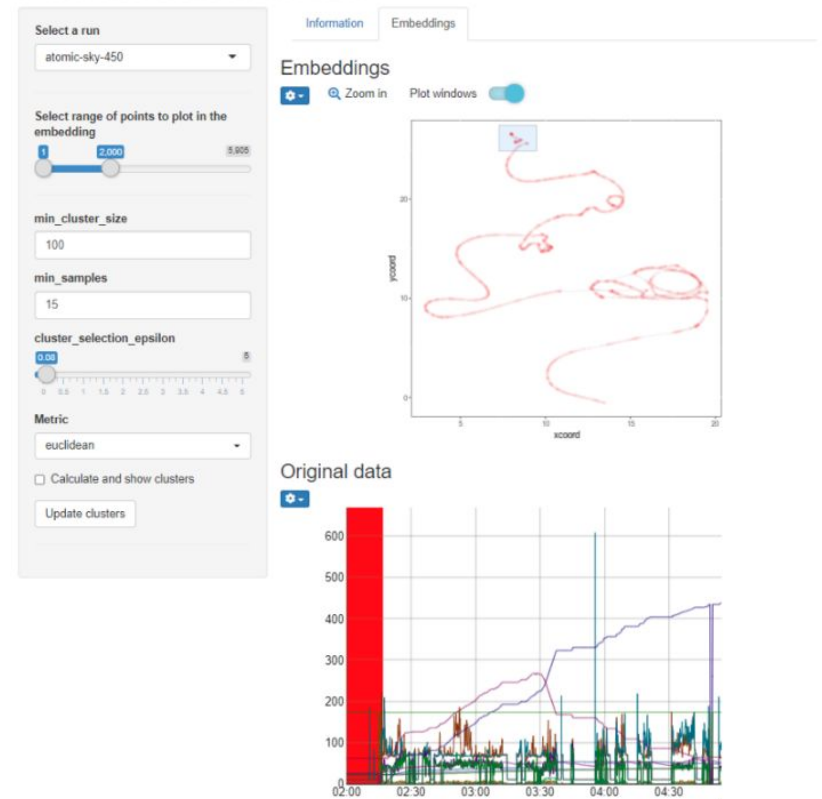
2. DCAE



3. Dimensionality reduction



Timecluster extension visualizer





Data representation (AGH)

- ❖ **Dual representation of data**
 - ✓ Expert based representation
 - ✓ Data-driven representation
- ❖ **Tools supporting exchange process between data-driven and expert-based approaches**
 - ✓ Knowledge Augmented Clustering (KnAC) -- expanding and verifying expert knowledge with automated methods
 - ✓ Explainable Clusters (ExClusters) -- explaining differences between automatically discovered clusters
- ❖ **Formal verification and conformance checking based on event logs**



Main Papers

- ❖ Grzegorz J. Nalepa, Szymon Bobek, Krzysztof Kutt, Martin Atzmueller: Semantic Data Mining in Ubiquitous Sensing: A Survey. *Sensors* 21(13): 4322 (2021)
- ❖ Víctor Rodríguez-Fernández, Agnieszka Trzcionkowska, Antonio González-Pardo, Edyta Brzychczy, Grzegorz J. Nalepa, David Camacho: Conformance Checking for Time-Series-Aware Processes. *IEEE Trans. Ind. Informatics* 17(2): 871-881 (2021)
- ❖ Marcin Szpyrka, Edyta Brzychczy, Agnieszka Napieraj, Jacek Korski, Grzegorz J. Nalepa, Conformance Checking of a Longwall Shearer Operation Based on Low-Level Events. *Energies* 2020, 13, 6630
- ❖ Szymon Bobek, Maciej Mozolewski, Grzegorz J. Nalepa: Explanation-Driven Model Stacking. *ICCS* (6) 2021: 361-371
- ❖ Szymon Bobek, Grzegorz J. Nalepa: Augmenting Automatic Clustering with Expert Knowledge and Explanations. *ICCS* (4) 2021: 631-638
- ❖ 3 more papers in review process, 4 more published

Workshops

- ❖ SEDAMI 2021 @IJCAI
- ❖ PRAXAI 2021 @DSAA IEEE
- ❖ Practical Applications of Deep Learning @IDEAL2020
- ❖ Industrial meetings in Poland and with partners



Industrial partners involved

1. **Famur - the maker of coal mining machines with clients in Poland and over the world**
2. **Arcelor Mittal Poland - global steel maker with facilities in Poland**

Exploitation

- ❖ Ongoing discussion with industrial partners
- ❖ Joint papers and industrial meetings
- ❖ Domain-specific challenges (coal mining)
- ❖ Continuation of collaboration (steel production)



Challenges

- ❖ **transnational cooperation:**
 - ✓ national project was never started in Italy
 - ✓ the start of the Spanish one was largely delayed, and ended prematurely
 - ✓ certain misalignment between national and European regulations
- ❖ **readiness of companies for R&D projects**
- ❖ **pandemic situation is challenging wrt**
 - ✓ dissemination,
 - ✓ on-site meetings,
 - ✓ cooperation with industrial partners

Thank you for your attention!
Questions ?



***2017-BDSI PACMEL project
(AGH-PL, UPM-ES)***



NARODOWE CENTRUM NAUKI