# PRESED -
# Predictive Sensor Data Mining for Product Quality Improvement

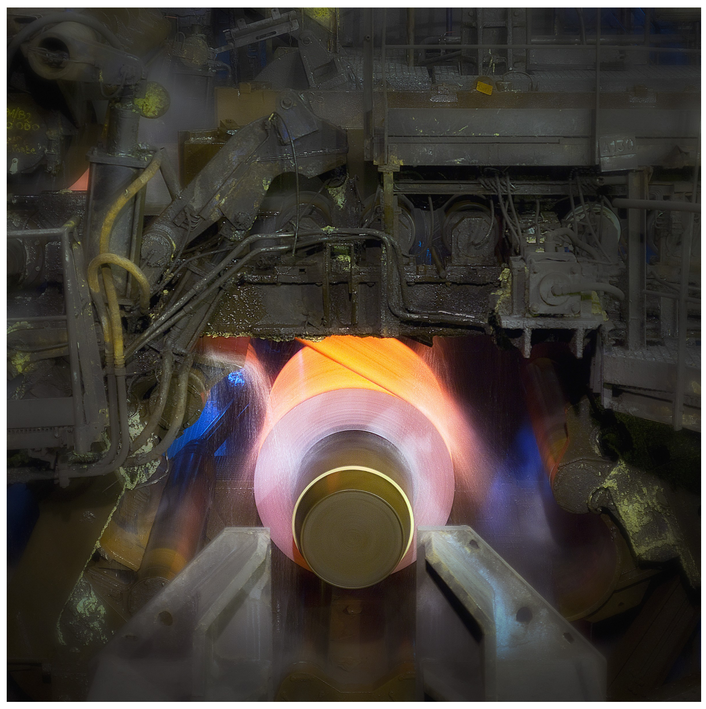
**David Arnu**[1], Gabriel Fricout[2], Marcus Neuer[3], Partick Gallinari[4], Jean-Baptiste Leger[5], Valentina Colla[6]

## Project Overview

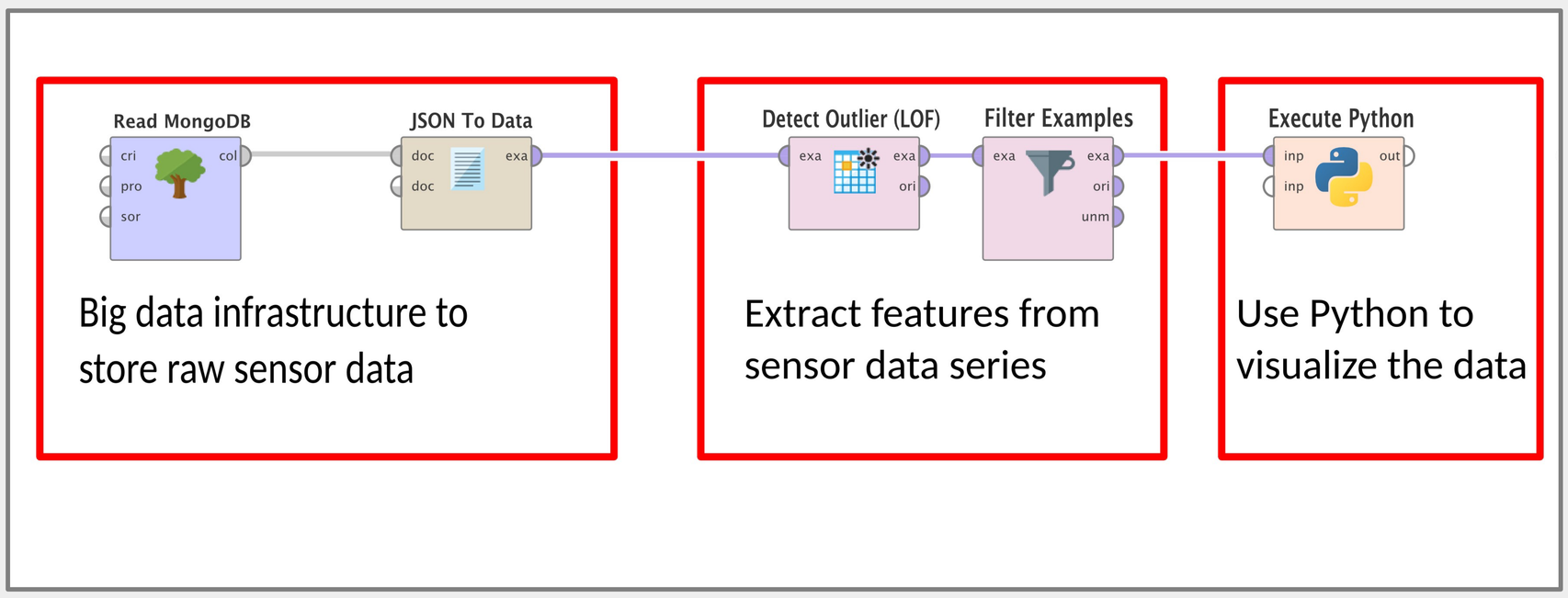
Manufacturing high quality steel products is a very complex process with a lot of intermediate steps.
Modern steel plants also produce huge amounts of complex data from sensors and production parameters.
The goal of our project is to develop new methodologies to help steel plants to improve their products and reduce manufacturing costs.

## Use Cases

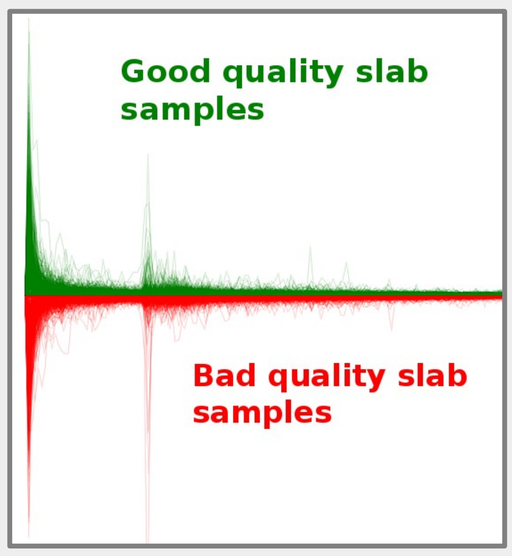
- Identifying causes for low product quality
- Predict the quality of a product during production as early as possible
- Integrate expert knowledge for different production sites



Fos-sur-Mer, France. Photo by and all rights by: Alain Sauvan / ArcelorMittal Fos-sur-Mer



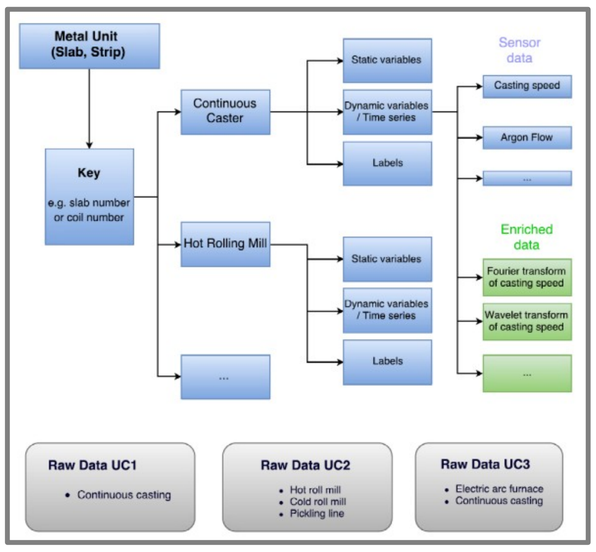

RapidMiner process showing the access and transformation of the raw data




Overlay of Fourier transformed signals

## Challenges

- Design a data structure to track the material over the complete production process
- Tracking is difficult as the product changes shape and quantity during processing (rolling, cutting)
- Huge amount of raw data (several hundred parameters, with a frequency of 1-10Hz over a 2-3 year period)




Generic data model for the PRESED use cases

## Technologies used

- *MongoDB* as NoSQL database to handle the production data
- Data visualization tool box written in *Python* and integrated into *RapidMiner* [1]
- *Representation learning* to process time series information [2]
- Time series shapelets for pattern discovery [3]
- Outlier detection methods for time series [4]
- Data and model management with *RapidMiner*
- Dashboards for interactive exchange between plants and data mining experts
- Ontology management with *KASEM* [5]

## References


- [1] Ingo Mierswa, Michael Wurst, Ralf Klinkenberg, Martin Scholz, and Timm Euler. 2006. YALE: rapid prototyping for complex data mining tasks. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD 2006)
- [2] A. Ziat, G. Contardo, N. Baskiotis, and L. Denoyer, “Car-Traffic Forecasting: A Representation Learning Approach.,” MUD@ICML, pp. 85–87, 2015
- [3] Xavier Renard, Maria Rifqi, Walid Erray, and Marcin Detyniecki. Random-shapelet: an algorithm for fast shapelet discovery. IEEE International Conference on Data Science and Advanced Analytics, pages 1–10, 2015
- [3] Cateni Silvia, Valentina Colla, and Gianluca Nastasi. "A multivariate fuzzy system applied for outliers detection." Journal of Intelligent & Fuzzy Systems 24.4, 2013
- [5] Monnin, M, Leger, J-B., Morel, D. (2011). KASEM®: e-Maintenance SOA Platform, in Proceedings of 24th International Congress on Condition Monitoring and Diagnostics Engineering Management, 29th May – 1st June, Stavanger, Norway


## Project partners

ArcelorMittal, BFI, PREDICT, IVA, rapidminer, Institute of Communication, Information and Perception Technologies, Scuola Superiore Sant'Anna, UPMC Sorbonne Universités

## Funded by

European Commission, Research Fund for Coal & Steel

Grant Agreement No: RFSR-CT-2014-00031
Project duration: 07/2014 – 01/2018
**www.presed.eu**

**1: RapidMiner GmbH (research@rapidminer.com)**, 2: ArcelorMittal Maizières Research SA, 3: VDEH-BETRIEBSFORSCHUNGSINSTITUT GMBH, 4: UNIVERSITE PIERRE ET MARIE CURIE - PARIS 6, 5: PREDICT SAS, 6: Scuola Superiore Sant'Anna