DT-SPIRE-06-2019 (870130)

Cognitive plants through proactive self-learning hybrid digital twins

Deliverable Report

Deliverable ID	D2.4	**Version**	V1.1
Deliverable name	A complete digital twin enabled with cognitive elements for the Steel pilots		
Lead beneficiary	SAG (SAARSTAHL AG)		
Editor(s)	Ulrike Faltings (SAG)		
Contributors	Ulrike Faltings (SAG), Maria Luschkova (DFKI), Tilbe Alp (Teknopar), Tamara Rodríguez Durán (Sidenor), Nenad Stojanovic (Nissatech), Branislav Jovicic (Nissatech), Sailesh Abburu (SINTEF), Stein T. Johansen (SINTEF), An Lam (SINTEF), Aylin Demircioğlu (Noksel), Özlem Albayrak (Teknopar), Perin Ünal (Teknopar), Ljiljana Stojanovic (Fraunhofer)		
Due date	28.02.2023		
Date of final version	Revised version – 24.11.2023		
Dissemination level	Demonstrator – Public		
Document approval	Frode Brakstad	24.11.2023	

The COGNITWIN project has received funding from the European Union's Horizon 2020 research and innovation programme under GA No. 870130

Executive Summary

This document is the final public deliverable “D2.4: A complete digital twin enabled with cognitive elements for the Steel pilots” of the project COGNITWIN. This document accompanies the demonstrator and is the result of the final stage (M33-M42) of the project development.

Saarstahl AG Pilot: Two main points for action have been focused on: **improving steel billet identification upon entrance into the mill train and enable seamless tracking of billets in the blooming train, a part of the mill train**. The first point is tackled by installing an improved billet identification system and the latter is to be solved by setting up a computer vision tracking system. The overall goals of SAG’s COGNITWIN approach are to improve rolling line efficiency by 15%, reduce energy consumption and process emissions by 15 % and to set up an automatic error detection.
The final project phase was used to enhance the billet identification system with cognitive elements and continue work on the tracking system development.

SIDENOR Pilot: The Cognitive Digital Twin will help to **reduce refractory wear and increase operational ladle lifetime.** The goal is to increase ladle refractory lifetime to 80 heats for full relining and 40 heats to partial relining. As part of this the ambition is to **reduce the critical refractory depth** for renewing the refractory lining. Sidenor’s initial situation was assessed and the relevant process data were identified. Moreover, the measurement of the ladle profile was also established. The relevant production information was shared with the involved partners so that they defined the predictive models.

NOKSEL Pilot: The purpose of the planned cognitive digital twin regarding **predictive maintenance for the Steel Welding Plant (SWP)** was developed as a real-time monitoring system of the SWP production process that enables **predictive maintenance by integrating data from multiple sensor streams with the existing process monitoring models**. Hybrid digital twin was established. Use cases, user stories and related challenges have been determined and specified. On site tests and validations of the system have been performed and are in progress. The development of a cognitive digital twin has been completed. With the COGNITWIN SWP System, the main targets are to achieve: 10% reduction in energy consumption, and 10% reduction in shifted average duration of downtimes.

Table of Contents

List of Figures

List of Tables

Acronyms

NOKSEL	Noksel Celik Boru Sanayi A.S.
SAG	Saarstahl AG
SIDENOR	SIDENOR Aceros Especiales S.L.
CC	Continuous Casting
CEP	Complex Event Processing
CT	cognitive digital twin
DB	Database
DP	Distribution Panel
DT	digital twin
DWH	Data Warehouse
EAF	Electric Arc Furnace
EBT	Eccentric Bottom Tapping
ESB	Enterprise Service Bus

FC	Factored Contributions
GAN	Generative Adversarial Networks
GUI	Graphical User Interface
IoT	Internet of Things
JSON	JavaScript Object Notation
KeNN	Keras Neural Network
LF	Ladle Furnace
MCC	Main Control Center
MES	Manufacturing Execution System
MFT	Material Flow Tracking
ML	Machine Learning
MLP	Multi-Layer Perceptron
MQTT	Message Queuing Telemetry Transport
OEM	Original Equipment Manufacturer
OPC	Open Platform Communications
OPCUA	Open Platform Communications Unified Architecture
PCA	Principal Component Analysis
PLC	Programmable Logic Controller
PPBM	Physics-Based Modelling
REST	Representational State Transfer
RFID	Radio-Frequency Identification
RTSP	Real-Time Streaming Protocol
RUL	Remaining Useful Life
S3	Simple Storage Service
SAW	Submerged Arc Welding
SCADA	Supervisory Control And Data Acquisition
SM	Secondary Metallurgy
MEWMA	Multivariate Exponentially Weighted Moving Average
SMB	Sidenor Measurements Buffe
SUCM	Sidenor Unbatched Cyclic Measurements Simulation
SWP	Spirally Welding Pipe Machine
UML	Unified Modeling Language
VD	Vacuum Degasser

1 Introduction

In the COGNITWIN project, three different pilots in steel production has been implemented. Each pilot covers a particular segment of a multi-billion euros industry and has its own specific challenges that need to be fully understood in order to develop techniques and methods to overcome them. The main objective of the final phase of the project (M33-42) has been to develop cognitive digital twins based on the hybrid digital twins established in the previous project phase. This document summarises this work.

The state of the models used, input and output data, their accuracy, speed and level of detailing, are clarified up to the limit that the confidentiality of prior backgrounds (IPR) is not compromised. The interaction and collaboration with the technical work packages WP4, WP5 and the usage of toolbox components is laid out.

Each pilot's individually challenges and subsequent solutions to these challenges are detailed.

This document is composed of three main chapters. Each chapter presents one of the pilot cases. In each chapter, the following template has been followed:

- Process description
- Current challenges
- Pilot specific aim
- Innovation
- Description of Data available
- IoT platform and architecture in use
- Cognitive digital twin
- Demonstrator
- Conclusion
- Measurable KPIs and Final impact

In the steel production process, from pig iron to rolled bars or wire rod, a multitude of sensorial and relational data from various sources arises. In order to generate additional value from this, a linkage between data from different sources is needed, but due to the typically harsh conditions in a steel production plant, technologies such as RFID sensors are often unsuitable for this task.
A computer vision-based system can provide a robust alternative as cameras can often be placed at a certain distance from the target to be observed, shielding them from the gravest impact of the harsh environment.

The steel industry is a key driver of new developments in the refractory industry due to the high market share of steel refractories in the range of 60 to 70% and the harsh conditions for refractories in the steel making processes. In fact, the annual refractory consumption in a steel plant that produces more than 750.000 ton of steel per year, can reach around 10.000 ton.

Ladle refractory is a key factor in secondary metallurgy management for all steelworks despite particular differences due to process, installations or product conditions. It has clear economic implications but also affects quality, productivity and safety of the steelmaking installations and people involved. At the same time, the spent ladle bricks are generally sent to the landfill generating an additional cost and a waste of historically considered critical materials like Magnesite. From the refractory utilised in the steel industry, a 49% is dissolved in the process and 36% ends in a landfill.

Spiral Pipe Manufacturing industries has promising features for market uptake of digital technologies and provide a convenient infrastructure for ground-breaking innovations as it is immature yet. Condition monitoring market is expected to witness high growth value like 2.21 billion USD in 2017 and expected to reach 3.50 billion USD by 2024. Digital twin is gaining attraction and digital twin related sales are foreseen to reach about 18.29 billion USD in 2024.
It is estimated that machine downtime costs UK manufacturers £180bn every year. The research by Oneserve, found that ”... Each time the machine breaks down, it takes on average, 9 hours to fix. But some report having to wait 72 hours for a resolution, taking an enormous hit on the production schedule and decreasing productivity.... “. In another words; broken machinery and faulty parts are hampering productivity, equivalent to almost 3% of all working days. Therefore; by reducing maintenance costs in manufacturing environment will result to both efficiency increases in steel pipe production, decrease in maintenance cost and save energy more.

2 Saarstahl AG – Pilot

Introduction to Saarstahl & Process description

The Saarstahl AG - with its locations in Völklingen, Burbach and Neunkirchen along with Roheisengesellschaft Saar in Dillingen (Saarstahl and Dillinger Hütte each with 50%) - is a German steel manufacturing company with a global presence on the steel production market. Saarstahl AG specializes in the production of wire rod, hot rolled bars and semi-finished products of various sophisticated grades. These products are important preliminary products for the automotive industry and its suppliers, general mechanical engineering, oil and gas industry, the mining industry and other steel processing branches. The primary goal of the SAG use-case is to track individual billets in the Nauweiler rolling mill train, thus providing a linkage between various sensor data as well as other relational data on individual billets collected before and after the non-continuous part of the mill train. Figure 1 depicts a schematic overview of the Nauweiler rolling mill.

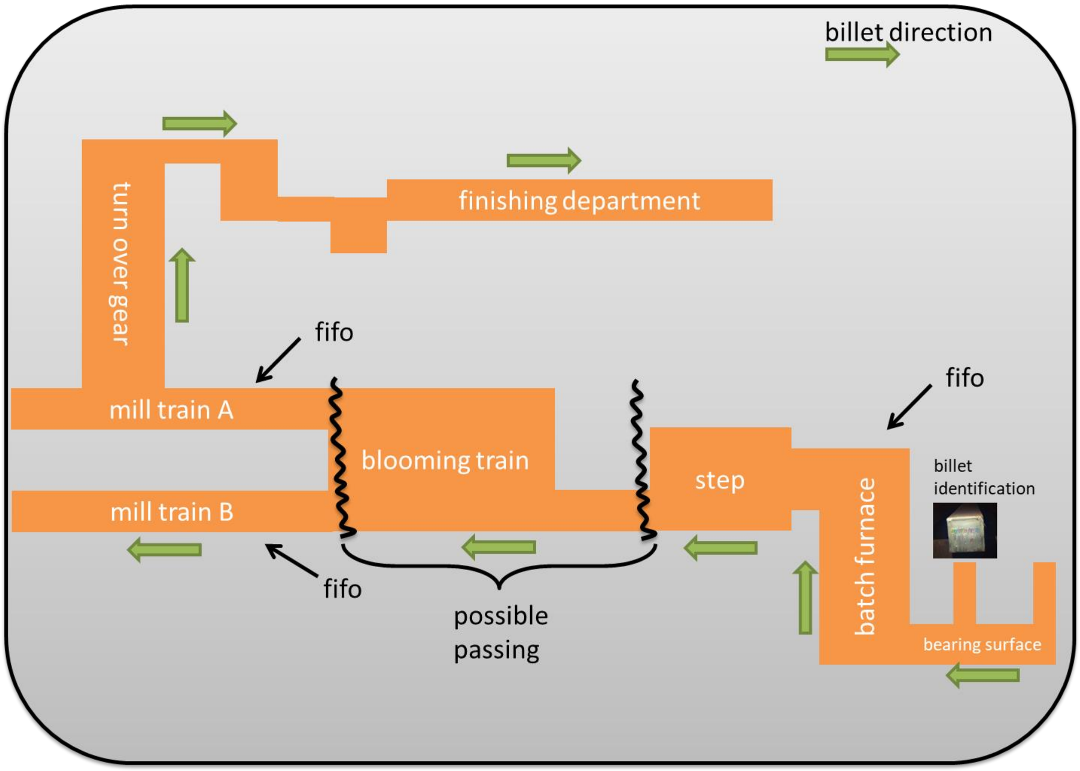


Figure 1: Schematic overview of the Nauweiler rolling mill

The cold steel billets coming from the steel mill have an ID stamped on to them. When a steel billet enters the Nauweiler rolling mill train, this ID is automatically read before the billet enters the oven to be heated for rolling. Upon leaving the oven, the heated billet enters the first sequential rolling stands in the step. After that, the roll strand enters the blooming train, a non-continuous part of the mill train, where it moves back and forth, repeatedly passing through several rolling stands. Here rolled bars can overtake one another, or a bar can receive a too severe bend to be rolled further and needs to be put aside and removed from the mill train area after cooling down. This non-continuous part of the mill train is where a computer vision tracking system is to be installed. Operators will occasionally enter the area while the mill train is active, introducing the need for an automatic anonymization of employees/people in the video stream. The pace with which the roll strand moves back and forth through this section is moderate; not more than 20 km/h. After this non-continuous part, the roll strand enters another continuous part of the mill train where rolling is completed.

Pilot challenges

Table 1, Table 2, and Table 3 collect the challenges for data handling, (cognitive) digital twins and visualization & control in the Saarstahl pilot.

Table 1: Pilot challenges for Digital Twin Data Acquisition/Collection for SAARSTAHL pilot

Sensors	• **Challenge**: Need for Full HD video camera for Blooming and Mill Train • **Requirement**: Full HD camera for industrial environment • **Solution**: Introduction of Machine vision Full HD Cameras
Communication	• **Challenge**: To handle the video stream data in a suitable way • **Requirement**: RTSP stream for life video data, JSON via messaging queue/rest service for communication of tabular data. • **Solution**: For training purposes, video data exchange with technical partner via hard drive.
Cloud platform	• **Challenge**: Use the existing platforms of Saarstahl as the basis • **Requirement**: Utilise the existing Saarstahl infrastructure • **Solution:** Embed the solution into the Saarstahl infrastructure
Data Lake, storage	• **Challenge**: Resolve the management of large volumes of video data • **Requirement**: Ensure management of both training and production data • **Solution**: For productive setting, video data will not have to be persisted. Training data is saved in storage (x TB, type of storage…). Tracked IDs saved in DWH.
Digital Twin – Data driven representation	• **Challenge**: Neural network detecting billets for digital representation of status quo in blooming train. • **Requirement**: Vast amounts of training data required • **Solution**: Set up routine for mass generation of synthetic training data for DL network.
Real time event handling, CEP	• **Challenge**: Video images need to be analyzed in real time to understand the billets movement • **Requirement**: Neural Network inference speed needs to be optimized for real time event handling. • **Solution**: Use local processing – consider FPGA
Cybersecurity	• **Challenge**: Due to the risk of cyber attack the process system is isolated from external access. • **Requirement**: Ensure closed loop system without connection to the exterior • **Solution**: Ensure that the systems in the plant are not externally connected.

Table 2: Pilot challenges for Digital Hybrid and Cognitive Digital Twins for SAARSTAHL pilot

Analytics Models	• **Challenge**: How to match detections among frames • **Requirement**: Analytical modelling for matching DL detections over consecutive frames and different camera viewpoints.

	• **Solution**: Analytical models
Physical Models	• **Challenge**: Billet location needs to be known accurately for ML algorithms to be able to optimize the process • **Requirement**: Seamless billet tracking movement, constrained by physical environment • **Solution**: Instance segmentation technology is used to track the billets
Machine Learning	• **Challenge**: Need to analyse video imagery in order to understand the movement of billets • **Requirement**: Need to have effective training and use of Image analytics including aerial photogrammetry including use Deep Learning Neural Network. Analytical modelling for matching DL detections over consecutive frames and different camera viewpoints. • **Solution**: A visual debugger for neural networks Neuroscope with use of aerial photogrammetry. • **Challenge**: only Real-life training data for DL networks not a feasible solution; • **Requirement**: Provide suitable training data • **Solution**: synthetic training data is needed as addition/supplement
Cognitive Digital Twins	• **Challenge**: Support self-learning of the system also after initial machine learning. • **Requirement**: The system should provide alerts and recommendations for operators and be able to learn continuously • **Solution**: Provide interactive operator guidance

Table 3: Pilot challenges for Digital Twin Visualisation and Control for SAARSTAHL pilot

2D/3D visualisation	• **Challenge**: Suitable visualization for operators. • **Requirement**: tracking of billets must be accurate and in real-time • **Solution**: User preferred visual presentations in Neuroscope tool
Control	• **Challenge:** interfere in real time if critical situation is detected to prevent damage to billet or the roll stand • **Requirement**: sufficiently short inference time of model and suitable visualization for operator • **Possible Solution**: alert operator with sufficient lead time and provide suggestion for action

Pilot specific aim

Figure 2: A billet passing through a rolling stand in the Nauweiler rolling mill

The objective of the SAG use case was to track individual billets in the Nauweiler rolling mill train and thus to be able to associate sensor and other data collected throughout the rolling process to the corresponding billet. Figure 2 depicts a billet in the Nauweiler rolling mill. Combining the data from the rolling mill associated to the billet with data collected beforehand at the steel mill would allow SAG to extend the digital twin of the billet to span the entire production process and enable the twin to acquire cognitive elements. The digital resp. cognitive twin in return could then be used e.g., to optimize production processes, recognize causes for deviations and, depending on the specific situation, react in real time to prevent deviations from occurring. Another benefit of the envisaged computer vision tracking system would have been to detect deviations and erroneous billets.

The tracking system developed in the curse of the COGNITWIN project consists of three parts:

- CV-based blooming train billet tracking, developed by DFKI
- Billet ID identification system, developed by SAG
- Anonymizer system, developed by SAG

Two main points for action were identified at the start of the project

- old billet identification system -> closed proprietary system, performance good but not perfect, and improvement only possible by manufacturer, hardware in use was reaching end of estimated lifetime. When automatic identification fails at present, the ID has to be entered into the system manually.
- blooming train -> initially, billets could not be tracked in blooming train. Due to harshness of environment, rolled bars possibly lying in very close proximity, erroneous bars cooling down in the area and manual operation of some appliances, conventional means such as RFID or other sensors or thermal cameras are not suitable for tracking.

Innovation

Improved billet identification upon entrance into mill train: in-house development. A prototype has been running in parallel to the old system since December 2019. Figure 3 depicts a billet with identified billet ID digits. The prototype is now fully integrated into the Material Flow Tracking (MFT) system.

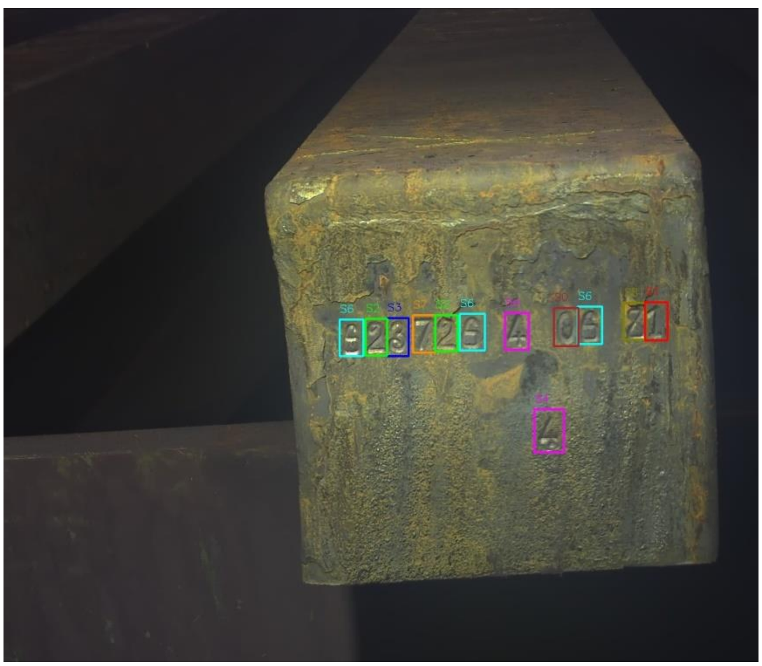


Figure 3: The billet identification system

Blooming train: the situation in the blooming train was analyzed together with local engineers to determine optimal placement of cameras and specify camera requirements, also taking network installation requirements into account. The camera data is to be used as input for a Computer Vision tracking system based on deep learning. Figure 4 depicts a schematic overview of the blooming train.

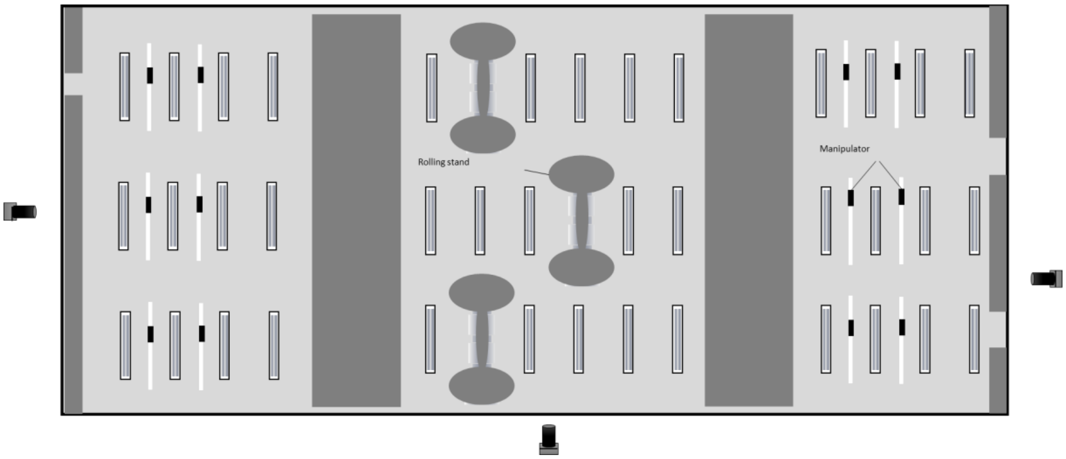


Figure 4: Schematic overview of blooming train with cameras

When running in production, the computer vision tracking system will need to process input from 3 Full HD cameras in real time; the number of frames per second required will need to be determined in the process of model evolution.
Operators will occasionally enter the area while the mill train is active, introducing the need for an automatic anonymization of people in the live and recorded video stream. The anonymizer should obscure employees with workwear and helmets.

Description of Data available

The data provided for the tracking system is a video stream stemming from 3 Full HD Cameras and possibly some additional video or image data. Figure 5 depicts captured images from the three blooming train cameras.

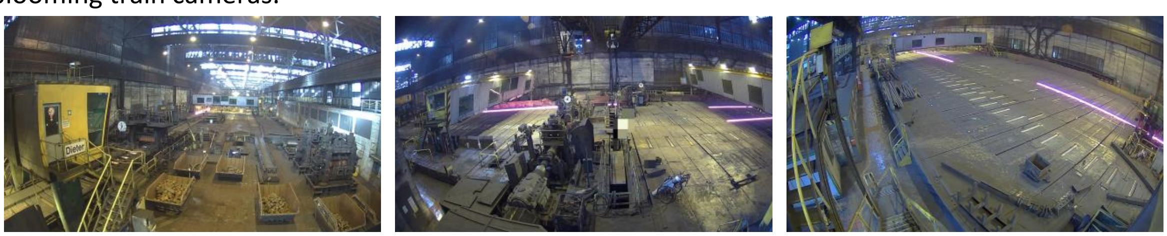


Figure 5: Captured images from blooming train cameras

For training purposes, recorded video files were provided in addition to the synthetic training data. The billet identification system uses sensor data as trigger and images obtained from a Full HD camera to identify billets. Training data was obtained via annotating captured images and generating synthesised data from captured images.
For the anonymization system, training data was obtained from open source data provided online and by generating synthetic data from captured images.

IoT platform and architecture in use

The rolling mill in Nauweiler is controlled by SAG's Manufacturing Execution System and SAG's Material flow tracking system. These applications flexibly exchange data (sensor and controlling data) to due interoperable data models between the assets and the high-level software systems.
Standards in use: OPC Unified Architecture, Enterprise Service Bus (Kafka or RabbitMQ), REST-Service
Components/services within the platform: MES (Manufacturing Execution System; in-house development), MFT (Material Flow Tracking; in-house development), Interfaces (REST or Message queue).

Digital Platform - Overall architecture for SAG Pilot

The IoT architecture at SAG is designed to combine a high flexibility, integrating multiple heterogeneous data sources and catering to a wide variety of users, and meeting performance and stability requirements dictated by production needs and user demands. Figure 6 gives a schematic overview of SAG's IoT infrastructure and Figure 7 depicts tools in use in the pilot IoT infrastructure for the operational phase.

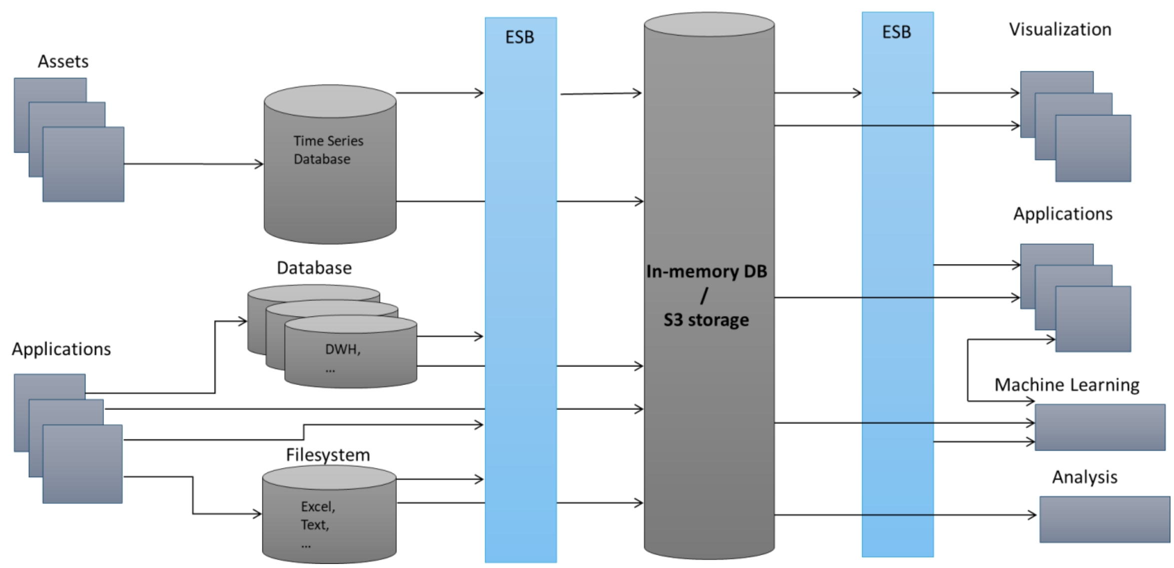


Figure 6: SAG IoT Infrastructure - Schematic Overview (ESB – Enterprise Service Bus, DWH – Data Warehouse, S3 – Simple Storage Service, DB – Database)

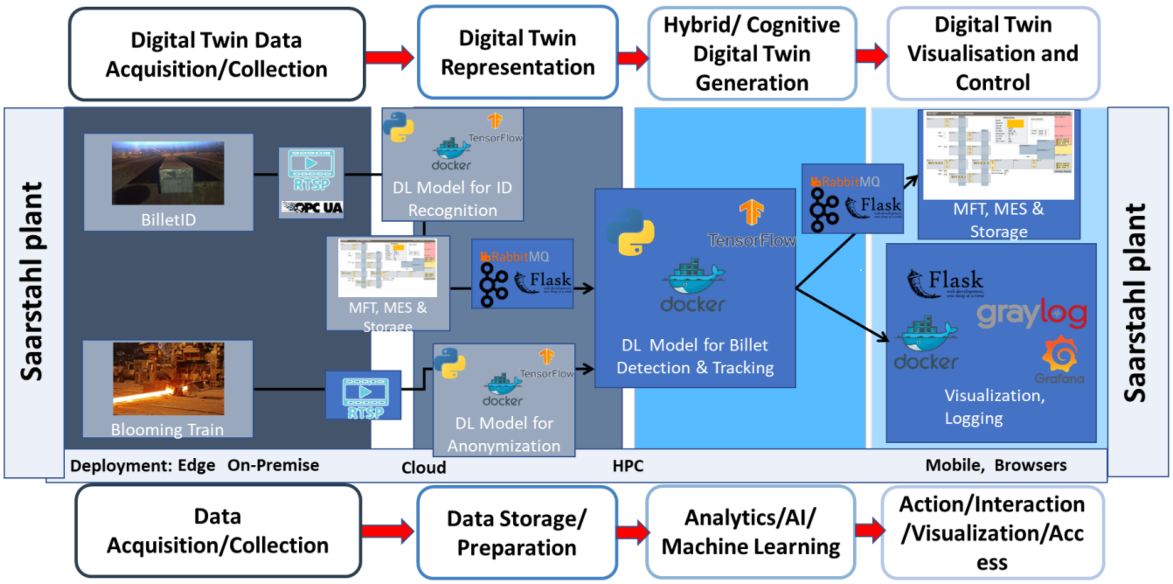


Figure 7: SAG IoT Infrastructure - Operational Phase

The pilot components are integrated into the In-house developed Material Flow Tracking system (MFT), as depicted in Figure 8.

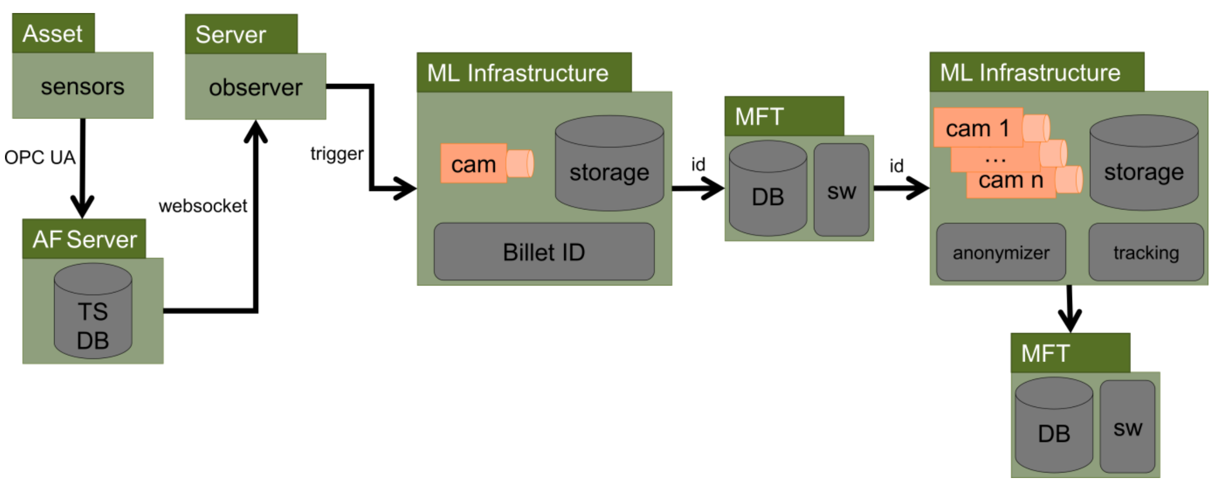


Figure 8: Pilot Integration

Sensors and Data Acquisition

The sensors used in this pilot are cameras, one for the billet ID recognition module, and 3 at the blooming train providing visual input to the tracking system. The video stream is ingested via RTSP. Trigger data for the billet ID recognition module and the blooming train tracking module are obtained via OPC UA through a timeseries database from assets in the corresponding parts of the mill train.

Database and Digital Twin Data representation

Production parameters and sensor data from throughout the steel making process are saved in a dedicated timeseries database for timeseries features and in a DWH for other features. Where applicable, batch or billet IDs corresponding to the data are saved alongside. In other cases, e.g., for weather data, keys such as timestamps allow for a matching of data to a billet or batch. Thus, a digital twin data representation is given at batch or billet level where applicable and where a seamless matching of sensor data to the corresponding ID is possible. At the blooming train section of the Nauweiler rolling mill, the initial situation is that a digital twin representation was only possible at batch level due to the lack of tracking in the blooming train. The pilot aims at allowing a billet level digital twin representation of sensor data in the blooming train section.

Cognitive Digital Twin – Analytics and AI

Billet ID identification

An in-house developed Deep Learning based system recognizes the digits of the billet ID. Combining this output with some 1st principles logic, it provides the corresponding billet ID to other systems. Figure 9 depicts a billet with identified billet ID digits.

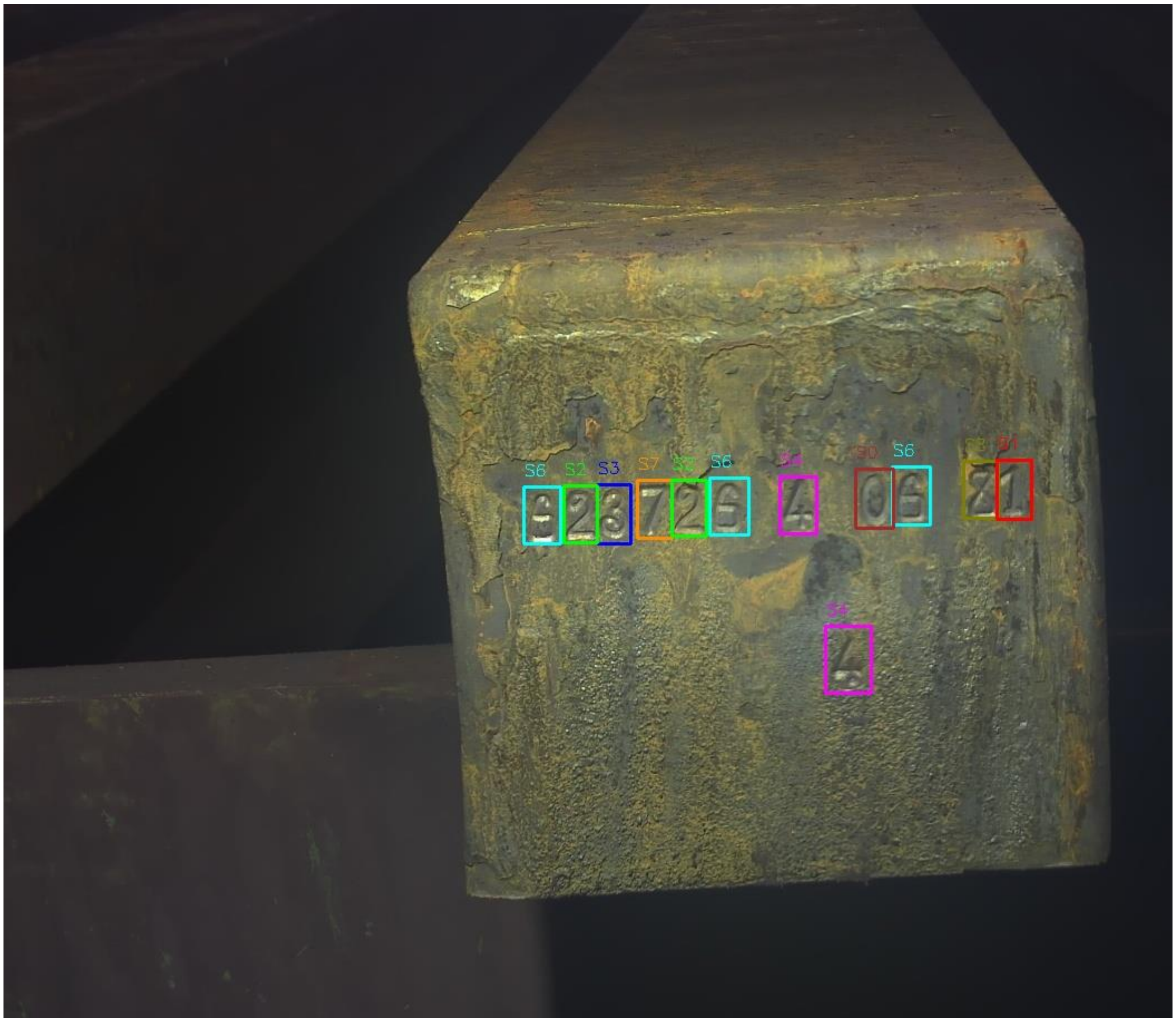


Figure 9: Identified digits of billet ID

The final project phase was used to roll out the punch stamp reader to more locations and to continue work on the integration into existing systems. Figure 10 depicts the integration of the punch stamp reader model and Figure 14 depicts the different locations at which the system is being implemented. A special emphasis was laid on a suitable workflow for machine-human-interaction and the development & deployment of the corresponding User Interface. Cognition is realized by the integration of human knowledge and experience into the Hybrid Twin workflow, as propagated in the Industry 5.0 approach [1]. Figure 11, Figure 12, and Figure 13 depict the user interface and the options for manual override by the operator.

[1] Industry 5.0 (europa.eu)

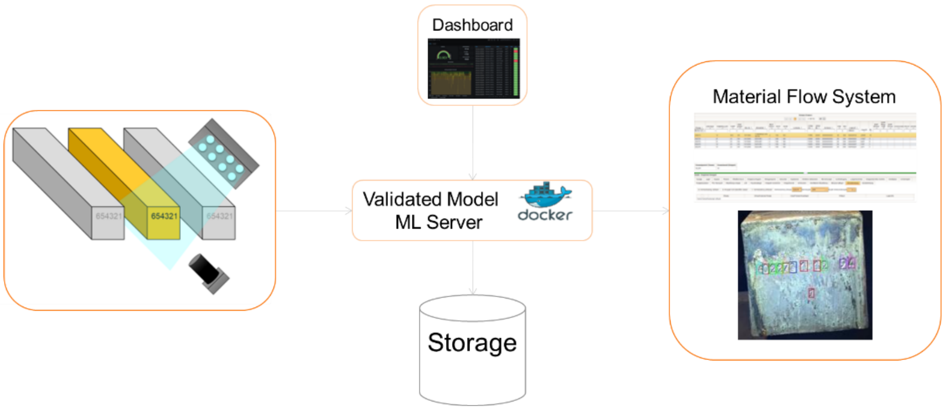


Figure 10: Integration of punch stamp reader

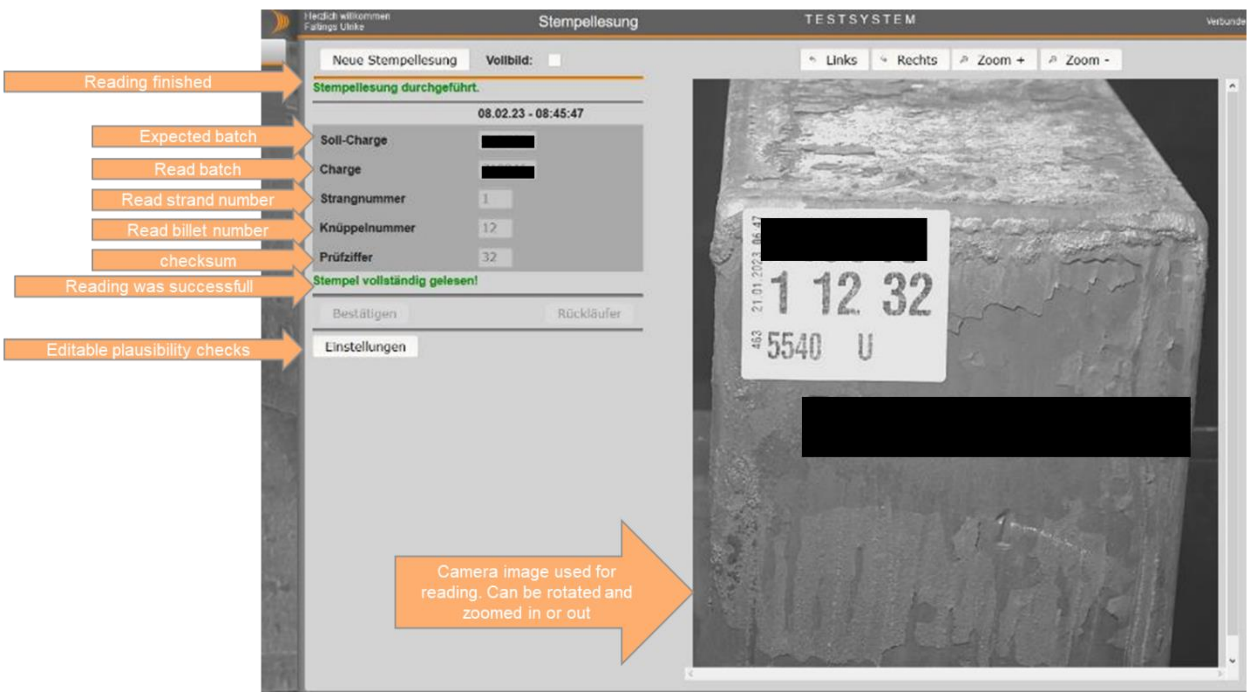


Figure 11: User interface for billet ID reader

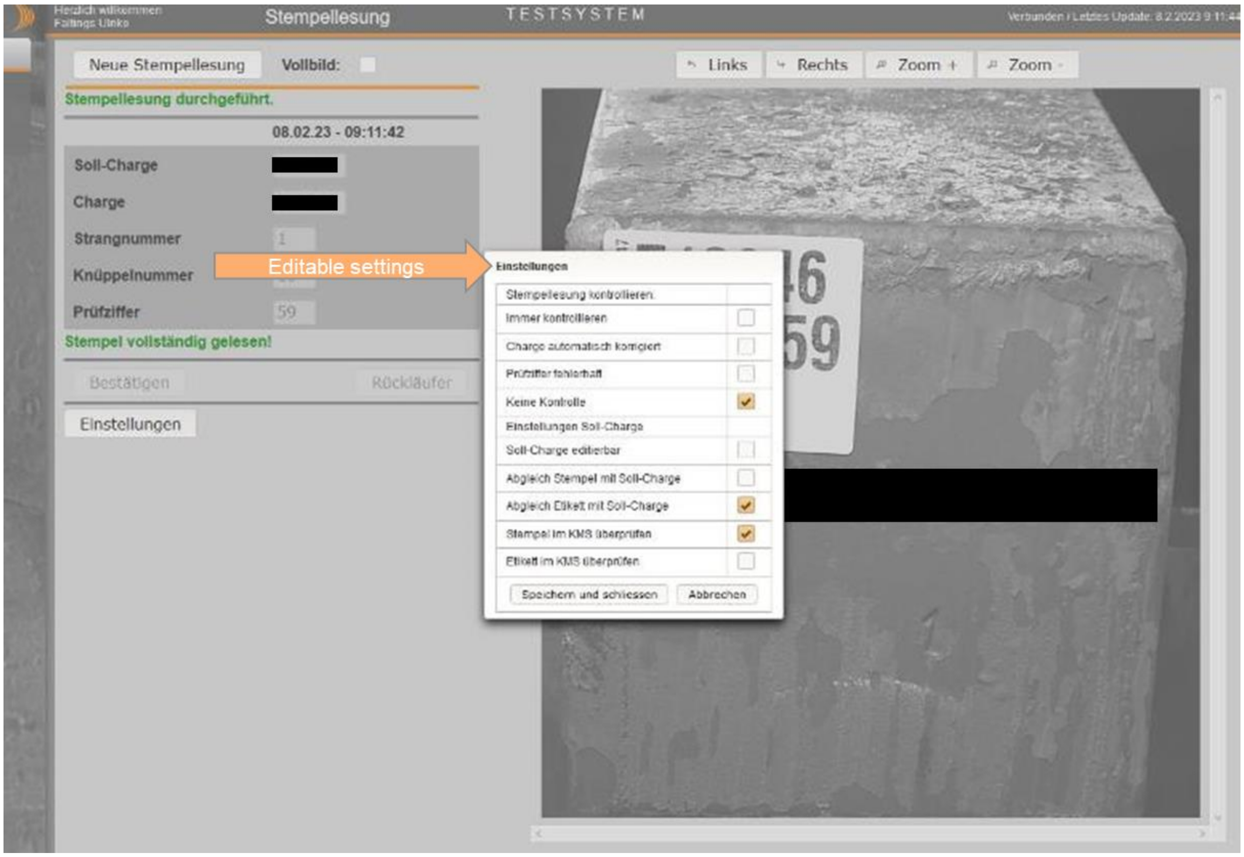


Figure 12: Editable plausibility checks and settings in user interface

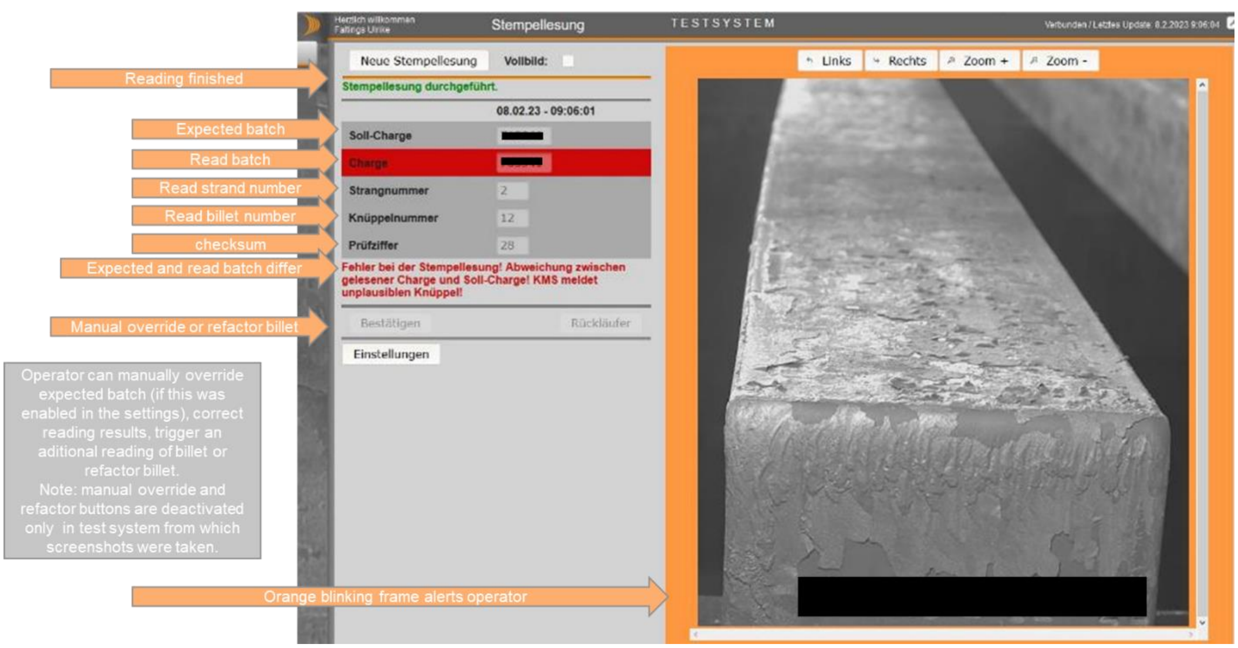


Figure 13: System alerts operator if reading was faulty

Figure 14:Novel billet/slab ID identification system instances at SAG and other SHS-entities

Thus the billet ID identification system reached a full cognitive twin state as envisaged at the beginning of the project, completely fulfilling its contribution to task 2.4.

Anonymizer

An in-house developed Deep Learning based system automatically detects people in the video stream of the blooming train surveillance cameras and pixelates them to prevent unnecessary surveillance of employees. Figure 15 gives a schematic overview of the anonymizer module and Figure 16 shows the output of the model.

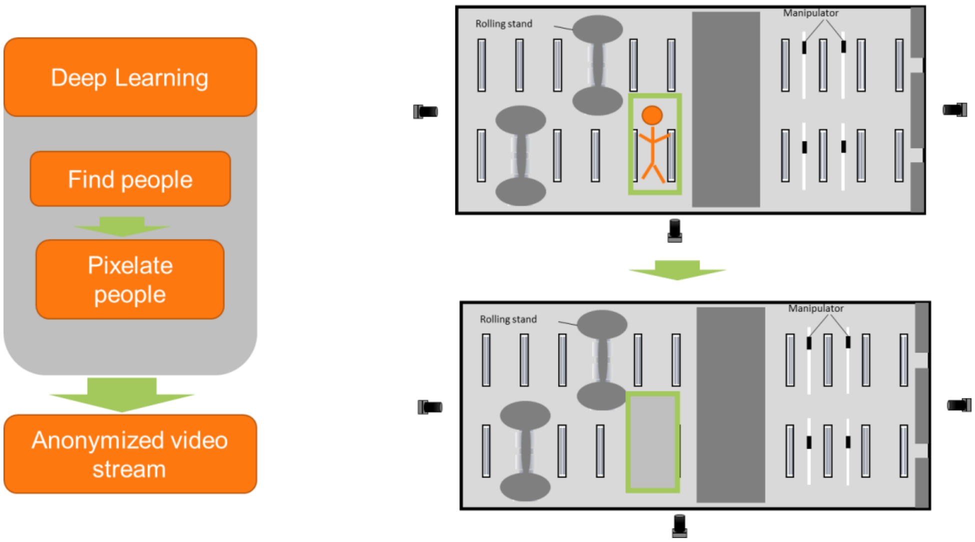


Figure 15: Anonymizer – Schematic overview

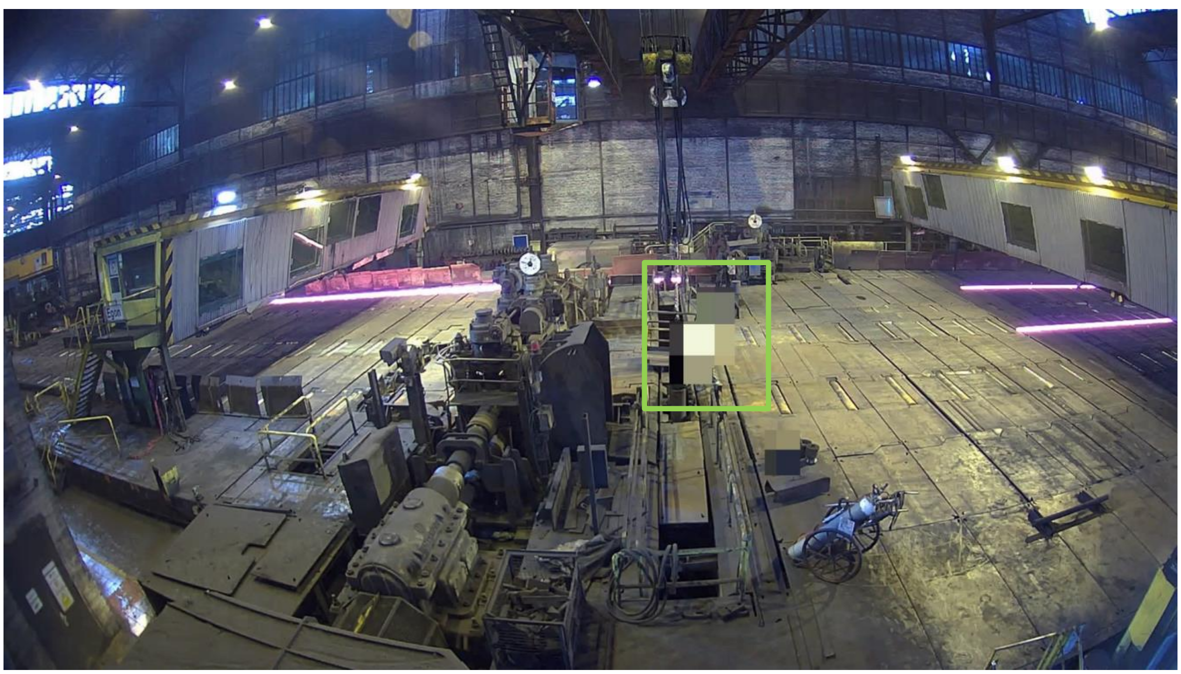


Figure 16: Anonymized employees (highlighted by green rectangle) in recorded video stream

The green rectangle in the Figure 16 is only added in this document for explanatory reasons. The productive system does not highlight employees.
The anonymizer system purposefully does not include human interaction as this would contradict the overall goal of protecting privacy rights and thus is constrained to a hybrid model combining 1st principles and data driven components. The final project phase was used to deploy the anonymizer system and make it available as a service to internal customers. Videos or images can

be sent to the anonymizer system via queue, which returns the anonymized item via queue back to the sender. Figure 17 depicts the deployment of the system.

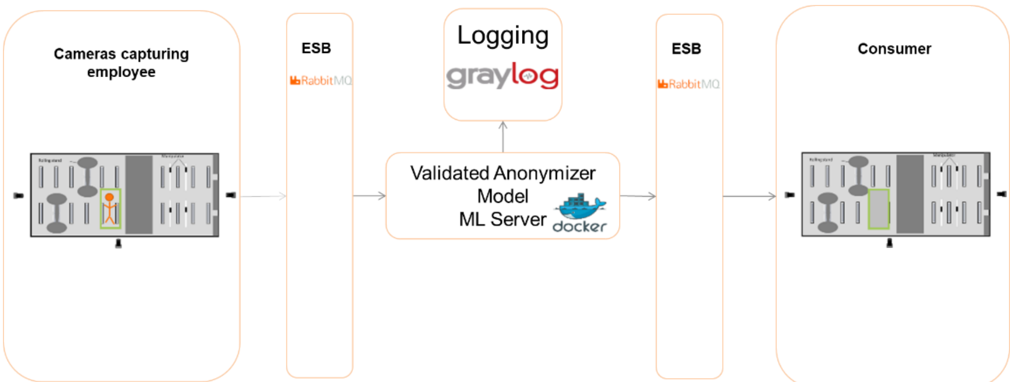


Figure 17: Deployment of anonymizer model

Thus the anonymizer system achieved all goals initially aimed for, completely fulfilling it's contribution in task 2.4.

Blooming train billet tracking

In the previous stages of the project, DFKI synthesized datasets for training the state-of-the-art object detection networks like Mask-RCNN and YOLO. Unfortunately, because of very challenging conditions for billet tracking (high level of occlusions, low level of individual billet characteristic, thin and elongates billet shape), tracking of individual billets via bounding box matching from one frame to the next frame is very unstable. Instead of using the "tracking-by-detection" approach, we propose to track billets with a temporal segmentation network with an architecture illustrated in Figure 18 . The instance segmentation model proceeds on a per-frame basis, guided by the output of the previous frame, which gives additional data source for instance linking about several frames in an image sequence. To be more precise, the network receives as an input the current frame **n** in the video sequence together with the billet instance masks **n-1** for the previous frame **n-1**. The network output is then predicted billet instance masks **n** for the current frame **n**. Since the billet positions in the current and previous frames differ only slightly, the billet linking task becomes much easier. The segmentation results for all three camera views in Figure 19 shows that the network can propagate individual billet masks through the image sequence.

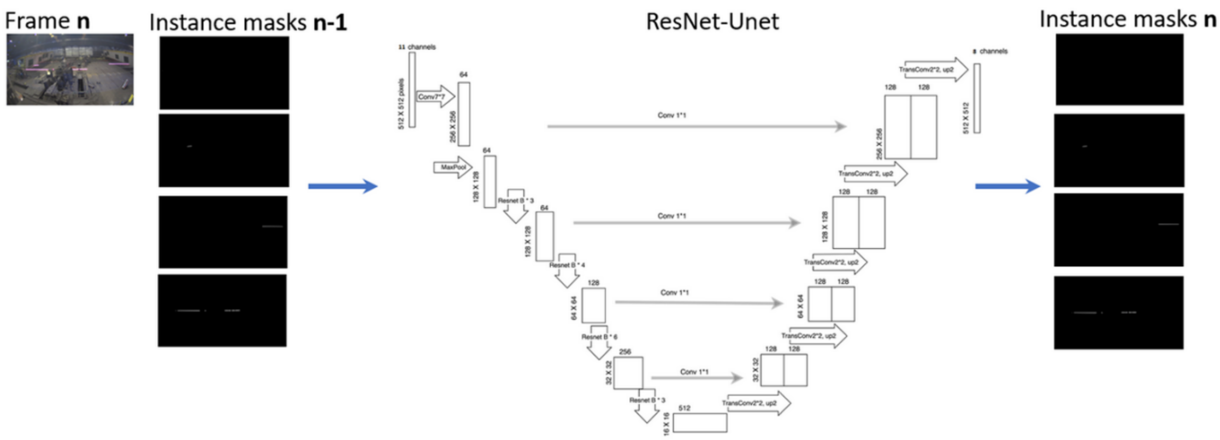


Figure 18: Spatiotemporal deep learning network approach. ResNet-Unet illustration adjusted from Charng et al. Deep learning segmentation of hyperautofluorescent fleck lesions in Stargardt disease. Scientific Reports 10 (16491), 2020.

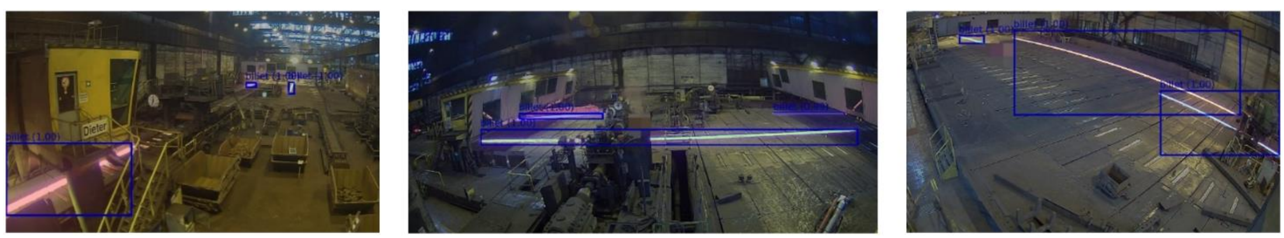


Figure 19: Snapshots from three blooming train cameras in Saarstahl's Nauweiler rolling mill. Billet tracking networks trained solely with simulated data track billet instances in real videos.

One of the most prominent issues of the proposed network is occlusion handling and billet re-identification. If an occluded area has large size, the network has difficulties to predict a correct billet ID after the billet re-appears behind the obstacle. We have tried to solve this problem in two different ways: by modifying the temporal segmentation network with amodal outputs and by post-processing the predicted mask using detection algorithms such as Mask-RCNN.

<u>Amodal temporal segmentation network</u>

Amodal completion is the ability to see an entire object despite parts of it being covered by another object in front of it. We modified the temporal segmentation network in Figure 18 by replacing the network output with amodal billet masks, Figure 20 . Amodal masks show not only visible but also hidden billet parts. The billet masks predicted by the amodal network show fewer billet ID switches, but the network is still not robust enough in the presence of large obstacles.

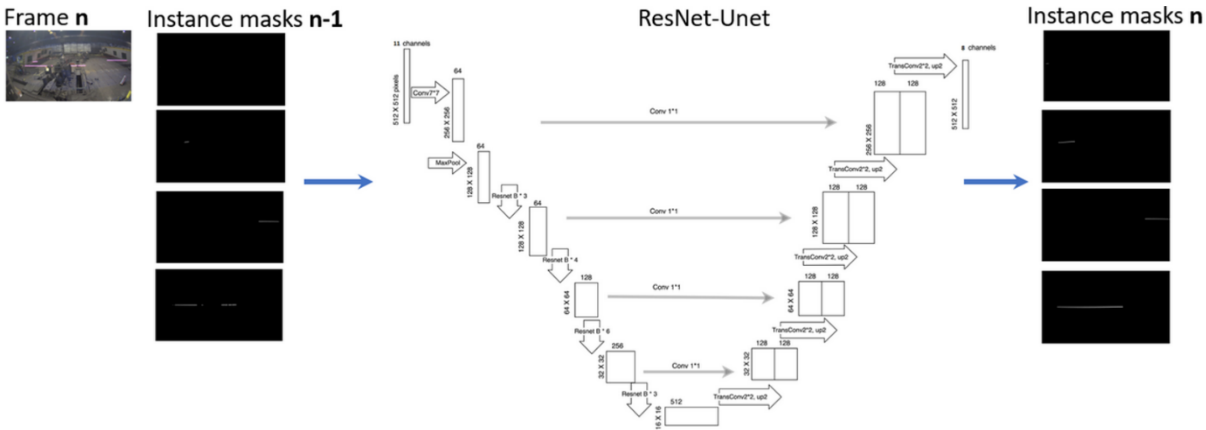


Figure 20: Amodal spatiotemporal DL network approach. Note that output masks are amodal - they show also hidden billet parts.

<u>Mask post-processing</u>

In billet detection experiments, we observed that Mask-RCNN performs well in handling occlusions. This advantage of Mask-RCNN can be used to reduce ID switches for covered billets. If a billet reappears after a large gap and its pixels have wrong IDs (see Figure 21 , left image), the pixel IDs can be corrected if Mask-RCNN detects that the pixels belong to the same billet (Figure 21, right image).

Figure 21: Left image: billet instance temporal segmentation without mask post-processing. Note that pixels of the same billet get wrong ID after the long gap. Right image: Billet instance segmentation with Mask-RCNN post-processing.

The mask post-processing works best for the second camera view since all billets in the view move horizontally. This horizontal orientation gives all the billets tight aligned Mask-RCNN bounding boxes, which minimizes the likelihood of multiple billets appearing in the same bounding box. However, the mask-postprocessing with help of Mask-RCNN bounding boxes works very poorly for the third camera, where billets move diagonally. Due to the elongated shape and diagonal orientation, billets become very large bounding boxes, which increases the chance of detecting multiple billets within the same bounding box. However, this can be overcome by rotating the images prior to the inference such that billets move approximately horizontally.

Challenges for billet tracking

We were not able to reach the deployable status of the tracking prototype, which seamlessly tracks the billets from the furnace to the end of the rolling street via all three cameras. Even though the models show promising tracking results in individual cameras, we could not setup a robust linkage between three cameras because of camera disjoint views due to a lack of time after availability of the necessary technologies. The continuation for this is now progressing afther the end of the project. Choosing other camera positions as depicted in Figure 22 and other camera optics can improve billet tracking. In such camera settings, billets in all three cameras would have approximately the same size and move horizontally, which simplifies usage of standard AI solutions for billet detection and billet tracking between the cameras, by needing less pre-processing on the images, but at the costs of more specialized and expensive camera hardware and a less generalizable system. However, this would not have affected the challenge of handling occlusions, which was the main reason for the delay in the subtask.

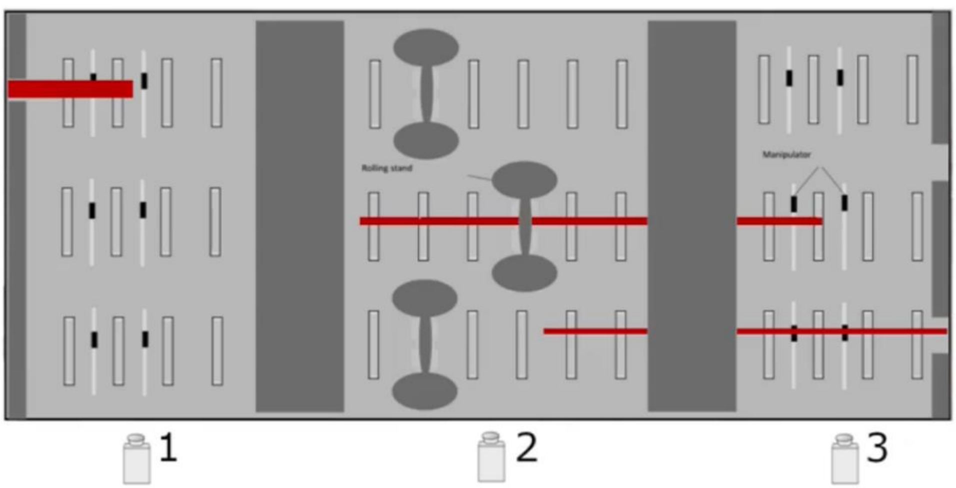


Figure 22: Modified Camera 1 and Camera 2 positions.

The infrastructure for deploying the cv-based blooming train billet tracking system was all prepared. Once the technical development of the system reaches a deployable status, the prototype can be tested in production and the KPIs can be validated.

Demonstrator – Cognitive Digital Twin

The final stages of the different components are demonstrated in the final SAG Demonstrator video [5]:

Demonstration script:

1. Punch Reader

2. Blooming Train Billet Tracking
3. Anonymizer

Conclusion and Summary/Challenges addressed and remaining

Initial challenges have been addressed as follows:
Obtaining sufficient training data for data-driven Deep Learning Computer Vision approaches can be challenging. For the blooming train billet tracking, this has been approached by generating a 3D-Model of the blooming train via photogrammetry and from this, rendering synthetic training data images by inserting synthetic billet instances into the 3D-Model. Challenges to counter in this approach were among others the size of the section of the mill train to be modelled, resulting in a large volume of high-resolution image data required for the photogrammetry approach, taking the lens distortion of the surveillance cameras into account in rendering synthetic images for training and modelling realistic synthetic billets into the images.
Moreover, a new billet ID identification system has been developed in-house to replace the old unsatisfactory one and a Deep Learning based network has been trained to automatically anonymize people in image data, addressing legal and works council issues preemptively.
The billet ID identification system is now fully integrated into the ambient productive systems and provides a graphical interface for the operator to enable the integration of the operator's human cognition into the hybrid twin model, thus constituting a full cognitive twin.
Similarly, the anonymization system is now deployed as a productive system realizing a full hybrid twin. Integrating a human operator into the loop would counteract the intent of the system to protect employees from unnecessary surveillance and was thus omitted.
Aside from the successes in the two aforementioned subsystems of the SAG Use Case, also the more general problem when applying Deep Learning based CV-systems of obtaining sufficient suitable training data was approached by testing and establishing different workflows and routines for the generation of synthetic data, depending on the particular requirements for the task at hand. With these, the SAG Use Case also showcases best practices for implementing DL-based CV solutions in industrial settings at an affordable workload and timespan. This is especially valuable for SMEs that lack the resources for extensive image collection and annotation.

Challenges and requirements remaining are as follows:
The third subsystem of the tracking system/SAG Use Case, the blooming train billet tracking system did not reach a prototype level suitable for testing at the production facility before the end of the project, but the foundation for the further development was provided.
SAG planes to finish the technical development of the CV-based blooming train billet tracking system in-house after the end of the COGNITWIN project. SAG's linked third party SHS has and is

hiring additional personnel for the IT department's AI group in 2023. These additional resources will make it possible to overcome the remaining challenges and obtain a fully functional tracking system integrated into the production environment.

Measurable KPIs and Final impact

- Improve rolling line efficiency by 15%

  By identifying and reacting to situations likely to cause erroneous bars, the rolling line efficiency is improved already. Additionally, providing a linkage in data associated to individual billets throughout the production process will allow SAG to use advanced analytics to identify other causes for deviations in the production process and react to these by e.g., adapting rolling parameters for individual billets. Moreover, the final version of the tracking system should allow for further automatization of the rolling process in the blooming train altogether [although this will most likely exceed the scope of COGNITWIN project].
- Reduce energy consumption (15%) and process emissions (15%)

  Each occurrence of an erroneous bar means a new billet needs to be cast and rolled, leading to additional energy consumption and process emissions. Moreover, the return transport to the steel mill for remelting has further impact on process emissions. Thus, by identifying and reacting to situations likely to cause erroneous bars, this additional impact can be reduced. Moreover, providing a linkage between data associated to individual billets over the entire process will allow SAG to use advanced analytics to identify other causes for deviations in the production process and react to these to reduce the level of scrap from the production goods even further [i.e., the level of billets/rolled bars with too severe deviations to be sold to the costumer that are remelted].
- Automatic error detection

  At present, there is no automatic detection of situations in process that will likely lead to e.g., bent bars or of erroneous bars in assigned section. Erroneous bars are identified manually. The target is to identify over 95% of erroneous bars automatically and to identify over 90% of situations likely leading to erroneous bars automatically. As to the tracking: At present, around 95% percent of all billets enter and leave the mill train sequentially, however, since there is no tracking system installed so far, it is not possible to safely link the sensor data obtained in and after the mill train to data associated to a particular billet ID obtained earlier in the production process. The goal is to track at least 98% of all billets successfully throughout the mill train and recognize when tracking failed such that at most for the two billets in the non-continuous mill train section at that time data from later in the process cannot safely be linked to one of the two respective billet IDs, only to the two IDs together/two-ID-tuple.

The full KPIs evaluation is in progress as the CV-based blooming train billet tracking system becomes deployable in production for further use. However, we still expect that a tracking system as planned for this use case would meet those KPIs.
Although two out of three subsystems of the tracking system reached their final stage and are fully deployed in production as initially envisaged, the full KPIs will be further validated as all three systems next will be deployed together.

The final Saarstahl pilot demonstrator is described in the COGNITWIN Toolbox [1] with the SAG pilot digital twin pipeline description [4] and the final SAG pilot demonstrator videos [5] . This is also further described in the final public deliverable D6.4 Best "Digital Twins" practices report [2].

3 SIDENOR – Pilot

Introduction to SIDENOR & Process description

Sidenor is one of Europe's leading steel manufacturers. The core business of SIDENOR is to recycle steel scrap and transform it into bars of more than 200 different steel grades the production flow chart depends on the customer specification and requirements. All the heats produced in Sidenor has 3 steps in common, the ones belonging to the melting shop (Figure 23), which are Electric Arc Furnace (EAF), Secondary Metallurgy (SM) and Continuous Casting (CC).

1. Electric furnace (EAF): the scrap is melted and once the liquid steel has the defined characteristics it is tapped into one ladle.
2. Secondary metallurgy (SM): Ferro additions are added to the liquid steel for obtaining the desired chemical composition. Moreover degassing, deoxidizing and inclusions conditioning control is done in this step. A tight control of the liquid steel temperature is very important.
3. Continuous casting (CC): once the liquid steel fits the customer specifications, it is necessary to solidify it. During this process the ladle is taken to the CC machine and the liquid steel is poured into the tundish and casted to billet or bloom format.

The liquid steel is contained during the whole production period in the ladle. Therefore, the ladle is very important. The ladle is covered internally with refractory bricks which are the ones in contact with the liquid steel. The ladle refractory wears heat by heat and it is very important to measure the erosion to avoid safety problems.

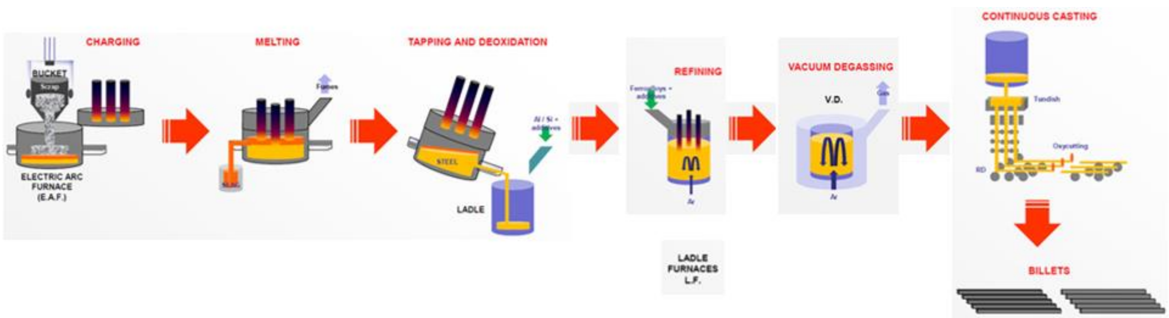


Figure 23: Steelmaking process in SIDENOR's production plant at Basauri, Spain.

Pilot challenges

Sidenor's main goal is to extend the life of the ladle refractory. The ladle consists on several refractory layers which wear in different way depending on their nature.

Refractory bricks located on the ladle walls must withstand the chemical attack produced by the slag. Moreover, the refractory must be stable when talking about avoiding reoxidation of the liquid steel and/or generating inclusions.

The ladle refractory lining (Figure 24) consists of:

- Wear line bricks, made of MgO-C. They are in contact with the liquid steel.
- Permanent line bricks, made of high Alumina or Magnesia. They work as safety lining.

The bricks in contact with the liquid steel are replaced after 1 life cycle, but on the contrary the safety lining is replaced only once per year.

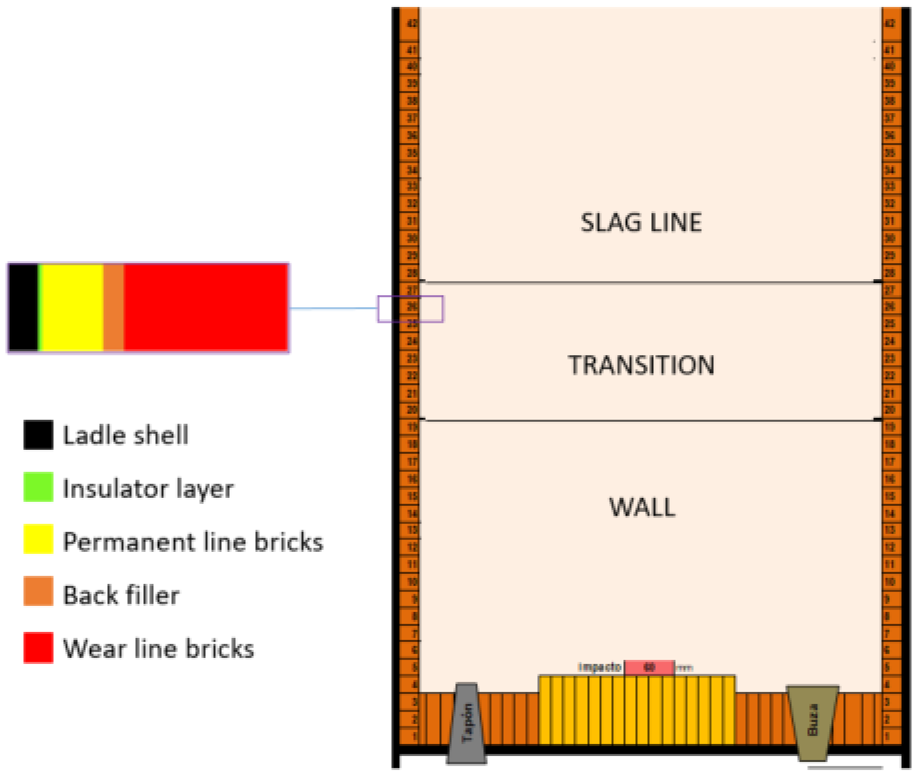


Figure 24: Ladle lining profile

The ladle is divided in different areas depending on the refractory configuration and/or process requirements. These are barrel/wall, transition and slag line. The refractory bricks are different depending on the position due to the differences on the chemical attack or mechanical requirements. These both mechanisms are the ones who erode the ladle. The principal mechanisms are:

- Chemical attack: Corrosion due to the contact of the bricks with the slag of the liquid steel
- Mechanical erosion: the liquid steel is stirred to homogenize the chemical composition and the temperature. This upward flow contributes to the erosion of the bricks

The wear of the refractory bricks is not equal around all over the ladle. Ladle wall nearest to the porous plug wears faster due to the liquid steel flow caused by the gas stirring. The bricks located on the slag line wears faster than the ones located bellow, due to slag chemical attack.

<u>Refractory erosion cycle</u>

The ladle life cycle or campaign is the number of heats produced with the same refractory lining. The operators must check visually the refractory to evaluate if the ladle condition is good enough for one more heat or not. This inspection is made after each heat. The result of the visual inspection added to the checking of several production parameters determine the end of the refractory life.

The final decision is taken based on the experience and the background knowledge. More often than desire, when a ladle is removed for repair or disposal the worker realized that the remaining thickness of the bricks is still useful for operation, implying impact on both productivity and costs.

The working cycle of ladle in the melting shop is cyclic. It rotates from preheating to tapping to SM and finally to CC, and preheating again. During the periods when the ladle is empty and up to the tapping, thermo-mechanical stresses are the main erosion cause, due to the very intensive temperature variations and to the mechanical impact of liquid steel when filling the ladle. On the contrary, during in-ladle metallurgical operations when liquid steel and slag are in contact with the refractory, thermo-chemical stresses are the main causes of erosion.

Another aspect that made difficult to development a better understanding of wear phenomena was the difficulty to get reliable data. SIDENOR measured manually the remaining thicknesses of the bricks and checked the refractory lining visually in hot condition. The measurements was done on several rows of the lining during the demolition or repair of the ladle after each campaign (Figure 25).

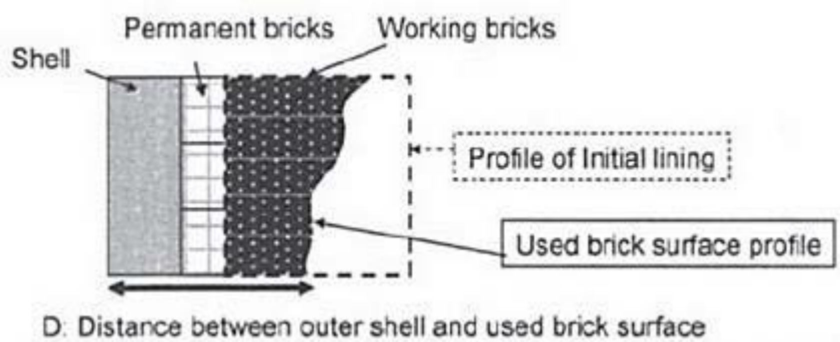


Figure 25: Section of the ladle displaying the remaining thickness

<u>Refractory life control</u>

SIDENOR works with several ladles at the same time. As mentioned above the refractory bricks are made of MgO-C, which are eroded in every produced heat. After a certain number of heats (1 campaign), if refractory lining is below a defined minimum thickness value, risk of liquid steel leakage appears. Therefore, the possibility of causing severe health hazards and production shutdown is too high.

This is the reason why, to assure the correct conditioning of the refractory lining a worker checked the refractory after each produced heat. The technician decided whether to repair or to completely demolish the existing refractory lining. The final decision was made based not only on the by visual inspection but also on checking several production parameters.

As the decision was taken by a person, the criteria were not the same for all the technicians, and sometimes they could decide to repair or replace the refractory bricks when this was not necessary.

Some figures related to SIDENOR's production are:

- 10 ladles in operation are used at the same time.
- The average estimated erosion of the refractory per heat is 3mm/heat.
- Average heats produced until reparation [33-46] and until demolition [69-82].
- The bricks on the bottom ladle are replaced at the end of each cycle.
- If during the visual inspection the thickness of the bricks is ≤50mm, the refractory is changed.

Pilot specific aim

The aim was to develop ML/AI models with cognitive capabilities that predict when the refractory lining in the ladle needs to be replaced or repaired.
The developed models can predict with high accuracy (P<1e-6) when a ladle must be taken out of production to repair.

IoT platform and architecture in use

Architecture of the data systems

The architecture of the data collection and storage systems is shown in Figure 26. The collected set of data, which have different source and nature, are described in the next paragraphs.

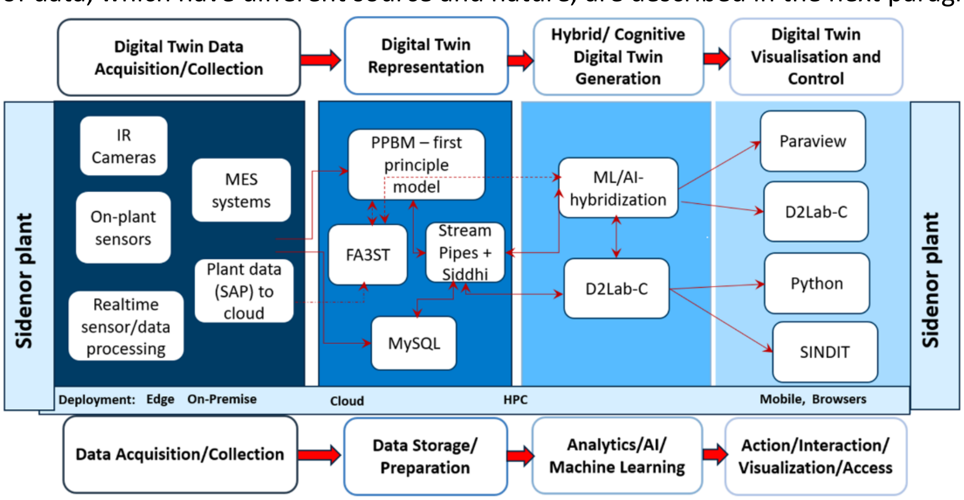
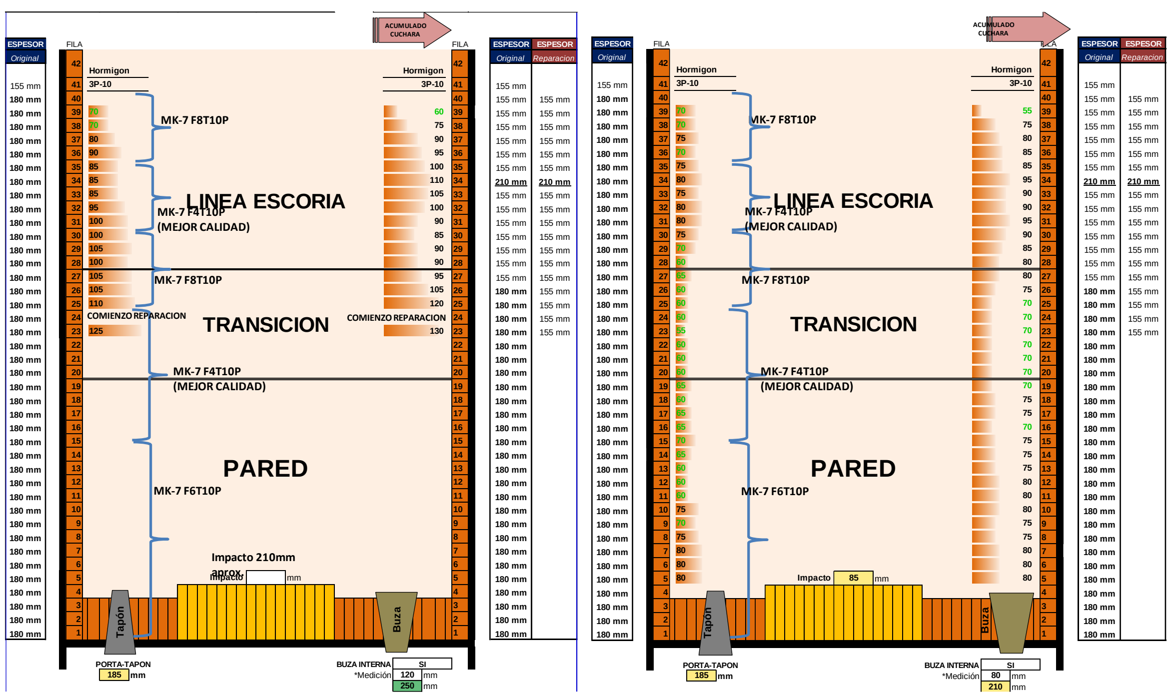


Figure 26 Architecture of the Sidenor data system

The steelmaking MES system Figure 26 centralizes all the information regarding the production process as heats, scheduling, requirements and programming. The main inputs for the MES systems regarding the programming activity are the ERP from the company (a SAP system) and the manual inputs described below. The SAP system stores customer's specification and requirements. The production sensors (PLCs and SCADAs) feed with different types of production data and reports the MES systems. MES system stores data collected manually too. As an example, the results obtained in the chemical laboratory from steel and slag composition, or the measurements done of the ladle refractory thickness.

Data about ladle refractory

As explained above, the ladles used for steelmaking production are repaired or demolished after several heats to assure no breakout. The repair consists of changing the bricks of the slag line and transition zone. These bricks are replaced by new ones, but they are measured row by row before removing, just to know the real thickness and wear. The measurement results are included in file (Figure 27.a) where the history of the ladle is registered. The most affected areas are highlighted, as well as the zones where the refractory has been optimised (in green).



(a) (b)

Figure 27: Remaining thickness of (a) ladle reparation; (b) Ladle demolition

The same measurement is done at the end of the ladle life (Figure 27.b). In this case, each row is measured and the remaining thickness is included in the database.

Another data source are the steelmaking parameters, as for example: time of the ladle with liquid steel, steel grade, vacuum time, etc.... These data are heterogeneous and were treated to obtain the information needed for the developed models. Two types of production data were studied:

- Acyclic data: process parameters taken heat by heat. See Table 4
- Cyclic data: parameters measured and recorded with a time-series data for each heat. The frequency of data collection is 1sec.

heat number	melting date	steel grade	tapping T (ºC)
218298	22-nov-21	40CrMo4FD	1638
218299	22-nov-21	40CrMo4E	1644
218300	22-nov-21	T35CrMo4FDJ	1644
218301	22-nov-21	18CrNiMo7EF	1645
218302	22-nov-21	18CrNiMo7EF	1640
218303	22-nov-21	95Cr6FD	1659
218304	22-nov-21	95Cr6FD	1667
218305	22-nov-21	95Cr6FD	1659
218306	22-nov-21	17CrNi6EF	1673
218307	22-nov-21	17CrNi6EF	1650
218308	22-nov-21	16CrNiPb4E	1672
218309	22-nov-21	16CrNiPb4E	1645
218310	22-nov-21	32CrNiMoBi6EF	1655
218311	22-nov-21	32CrNiMoBi6EF	1625
218312	22-nov-21	50CrV4	1643
218313	22-nov-21	56Cr3	1646

Table 4 Acyclic data: heat number, production date, steel grade and temperature at tapping

The data collected (cyclic and acyclic) were coupled to set up the database. The data model seeks effects from changes in the process data over the wear measured in the bricks. The short-term prediction works as an advisory system recommending ladle repair when residual thickness gets below a safety level. The combination of the data stored from ladle history can explain the rich data obtained with the manual measurement of the ladle profiles. As summary, Table 5 lists the data used in COGNITWIN and how to manage it.

Data source	Overview	COGNITWIN Integration	Granularity of data
MySQL	Level 1 data from PLCs	Yes, TBD	1 second
Informix	Level 2 data – MES	Yes, TBD	Heat (Production batch)
ERP	SAP	No	Heat (Production batch)
Excel	Refractory brick measurements	Yes, files	When ladle is removed for repairing (upper part) or demolition (complete)

Table 5 Overview of data source in Sidenor variables

Data transfer

The production parameters were sent to an FTP server so that all the partners involved in SIDENOR's pilot have access to them and can work with them. The acyclic data were sent to the server once a ladle was emptied (heat casting process is finished), and the cyclic data were sent once per week. The initial idea was to send the cyclic data at the end of each heat, but due to SIDENOR's cybersecurity politics this was not possible. This is the reason why the data are sent to the FTP server twice per week. The ladle wear is calculated at this moment and the result obtained are new inputs for calculating till which heat could be used the refractory.

Physics based (PB) model development

A physics-based model was developed for SIDENOR's pilot. Its objective was to provide information about refractory erosion, for a given ladle, from heat to heat. The model has prediction power as standalone Digital Twin (DT), but more important as part of a Hybrid Digital Twin (HT) and, finally, as part of a Cognitive Digital Twin (CT).

The overall goal was to predict if a ladle can be used at least one more time before relining. The model was built applying the methodology of Pragmatism in Physics-Based Modelling (PPBM).

The ladle operations involve many complex phenomena, where the most important are: transient thermal conduction, convection in liquid steel and liquid slag due to inert gas purging, waves set up by the bubble plume, natural convection in the steel, high power slag heater, slag, metal, refractory, thermal radiation, refractory dissolution, phase enthalpies of slag and metal (melting/dissolution), and handling of the composition and temperature dependent solubility of refractory.

Although the implemented Python model can be classified as a physical model, it requires process data to give accurate predictions. The model is simulating the actual process and the starting values and to a certain extent, the boundary conditions are taken from the data. A detailed description follows:
The modelling approach described above can, in principle, be used to simulate all operations (thermal operation, refining processes, and lining erosion) of a lifetime of a ladle, except for the re-lining process itself. Even though the focus has been on lining erosion during the refining stage, the whole cycle is important to get a realistic temperature in the ladle wall at the start of the refining. This is a function of the history of the ladle, and we have to note that no experimental data is available for the wall temperatures.

To run a simulation of a given heat, it is a relatively limited amount of data that is needed.
1) The total amount of steel that was in the ladle during the refining
2) The total amount of slag that was in the ladle during the refining. This was first done in a simple way by using the acyclic data giving the totals. Later this was extended to use data where the time and amount of each individual addition was done.
3) The added electrical power, purge gas flow rates, vacuum pressure, all as a function of time.
4) The time of each heat.

In addition, we are using the measured temperature in the steel phase. This is reported at 1 second intervals, but the number of actual measurements is limited to a handful for each heat. The orange line in Figure 29 is a result of drawing straight lines between some very few experimental points.

In addition, the model requires data on density, heat capacity and thermal conductivity of all materials. This is partly taken from literature and is partly provided by SIDENOR.

We see a simulation result demonstrating the predictions of erosion profiles and refractory temperatures Figure 28. The inner refractory lining is showing severe erosion at level of the slag line. As a result, a hot spot appears at the outer steel shell, indicating the wear of the refractory. This opens for the possibility to combine the model with thermal images and machine learning to get a very firm understanding of the erosion state inside the refractory. This possibility will not be further exploited in the COGNITWIN project but may be pursued in new projects. During the development physical property data had to be tuned to the industrial observations. Good agreement with measured steel temperatures could only be obtained by significant increase in the thermal conductivity of the refractory and insulation materials. This indicates that steel may penetrate in between the refractory bricks (depends on wetting, liquid hydrostatic pressure, gap sizes). This phenomenon can be included into the model in the future and may explain thermal images that show hotter spots in the lower part of the refractory.

The thermal predictions are important input to the refractory erosion model, as the solubility of refractory into the slag is temperature dependent. In addition, the predicted refractory temperatures at the time of filling fresh steel into the ladle, is the most important parameter for assessing thermomechanical stress erosion. There is no intension to predict this due to the complexity of the problem. However, these predicted temperatures may become a weighing factor for a machine learning algorithm to assess the thermomechanical stress contributions. We expect that the thermomechanical erosion contributions are more evident at the lower part of the ladle as this part is more heavily exposes to the hot metal during tapping into the ladle.

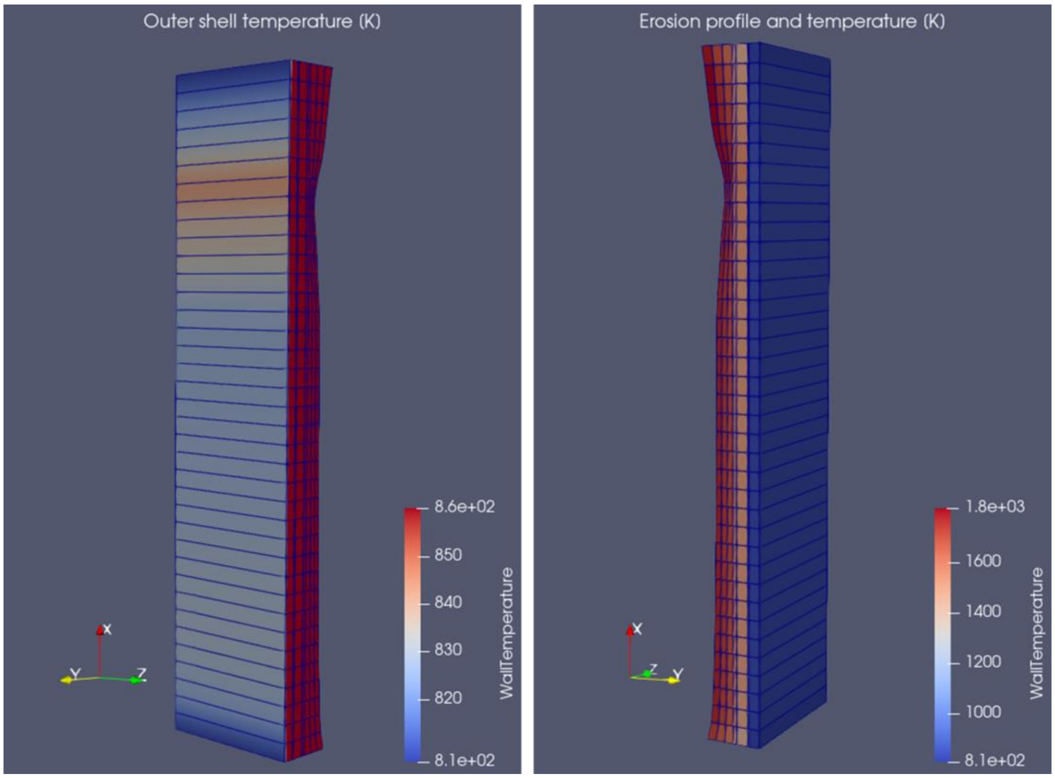


Figure 28 The figure shows a vertical section of a specific ladle refractory (Ladle 4, campaign 51,, use number 68, 2019).. Left figure shows the outer shell surface temperature (range 810 – 8600 K) and right figure shows erosion profile and refractory temperatures, all at 100 minutes after filling steel into the ladle. The erosion profile is the predicted profile at the time of demolition.

We see an example where the model predicts the temperature well in *Figure 29*. The prediction has several undulations which are due to combinations of additions (cooling the ladle) and heat addition. The measured data consists of 5 points (the first point is not real) and gives a snap-shot of the temperature at the time of sampling. As the model assumes perfect internal mixing in the metal and slag, and the sampling is taken at a specific point in the ladle, this accounts for some uncertainty in the comparison.

The model could simulate a complete lifespan of one ladle. All the ladles use in Sidenor during 2019 were simulated by the model and compared to operational data. The simulations handled the transient evolutions of approximately 5000 ladle uses and took around 8 hours on a single CPU.

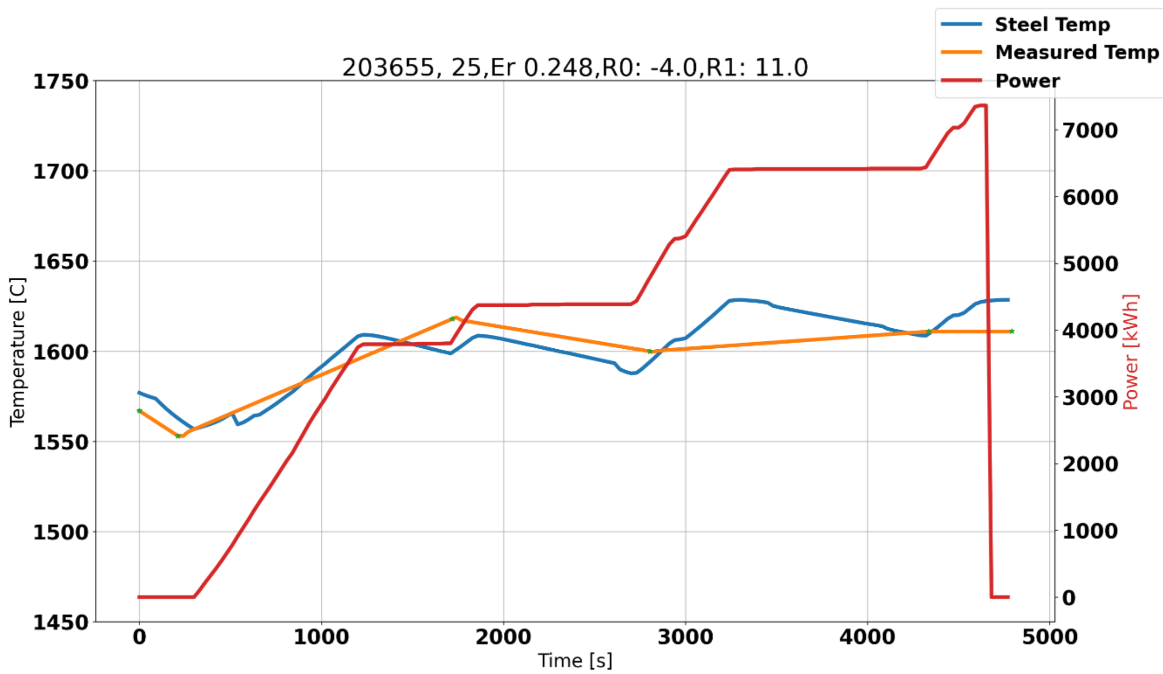


Figure 29 The experimental (points) steel temperature, the predicted steel temperatures and added energy by the slag heater (heat 203655, Ladle 5, campaign number 69, use number 25, 2019).

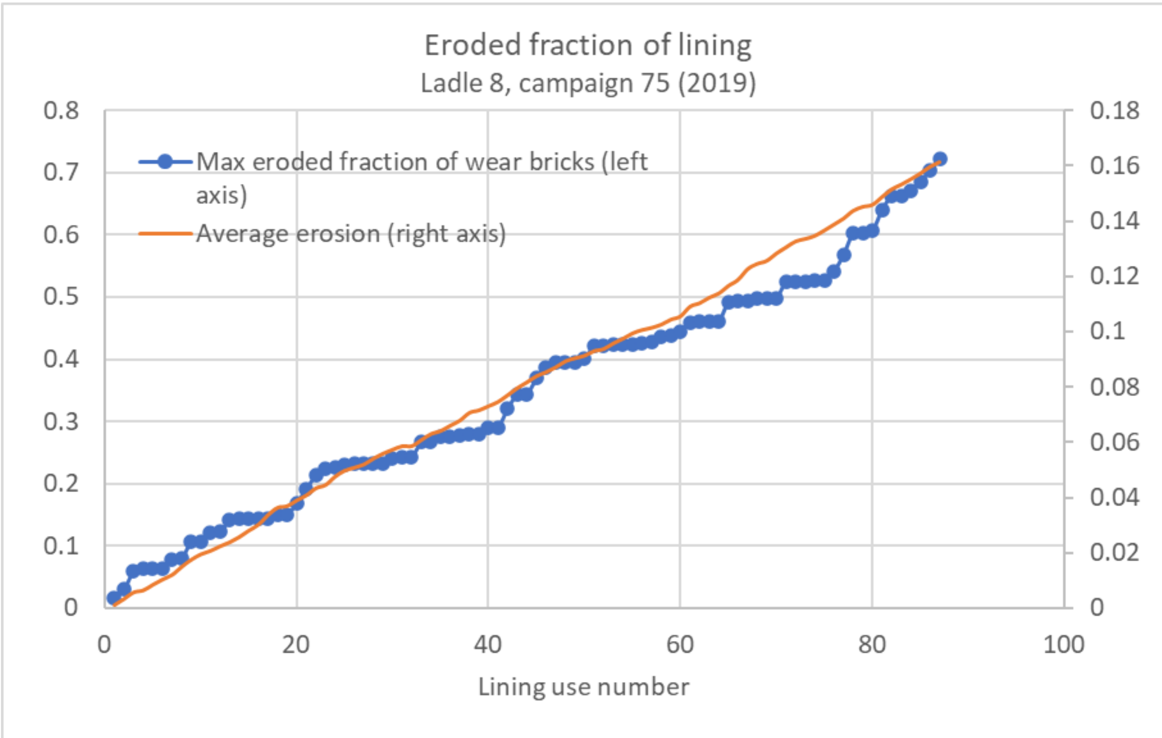


Figure 30 Prediction of the evolution of the refractory lining as it is eroded from use to use. The average is for the entire lining, while the maximum value is dominated by the position of the slag layer.

By running the model over complete campaigns we can compare the predicted erosion profiles with production data. In this case, we begin from the first heat after building the ladle, and simulate each heat. Both thermal and erosion history are accumulated until the operational data tells when the ladle must be repaired. Then, the refractory is repaired in the model, as done in production. The ladle is the restarted, using the repaired erosion profile, and run again until the last heat before demolition. Example data from the demolition is compared with predictions in Figure 31 and Figure 32. The erosion data is obtained by dividing the ladle into two halves and recording the maximum erosion for each of the halves. Quite high erosion is observed in the upper part of the ladle, where the refractory is not covered by the steel nor the slag. This is a special type of erosion that is not included in the model and therefore we use only brick numbers from 5 to 35 for the overall comparison, coming in next paragraph.

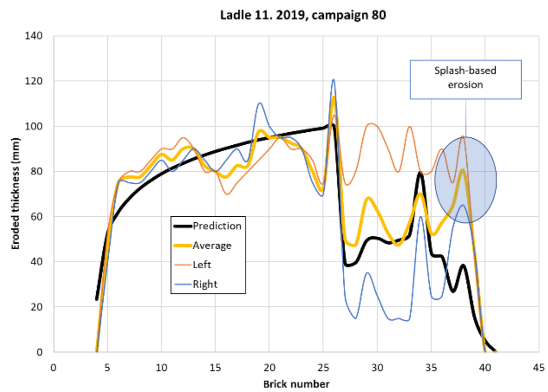


Figure 31 Comparison of predicted versus measured eroded thickness for Ladle 11, campaign 80, 2019.

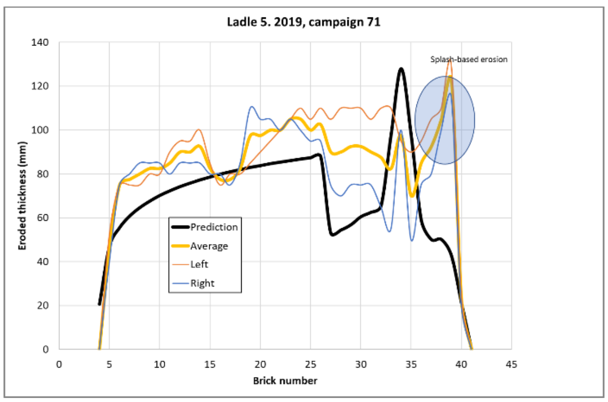


Figure 32 Comparison of predicted versus measured eroded thickness for Ladle 5, campaign 71, 2019.

The overall comparison between the predicted erosion and measured erosion is seen in Figure 33 and Figure 34. Two outliers A and B can be observed. Each point is the result of multiple uses of a ladle, until demolition has happened (typically 80 – 100 uses). As the observations are based on the most eroded positions at each level and we compare with a predicted average erosion, the model is designed to underpredict the data by 10 %.The outlier points may be understood by further analytics, including anomaly detections. However, changes in the quality of refractory bricks are hard to monitor and may influence refractory degradation without knowledge.

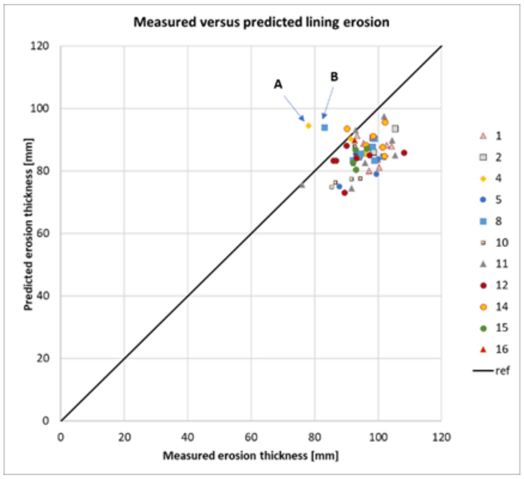


Figure 33 Comparison between measured and predicted erosion thickness at time of demolition of wear lining. Symbols represent different ladle numbers.

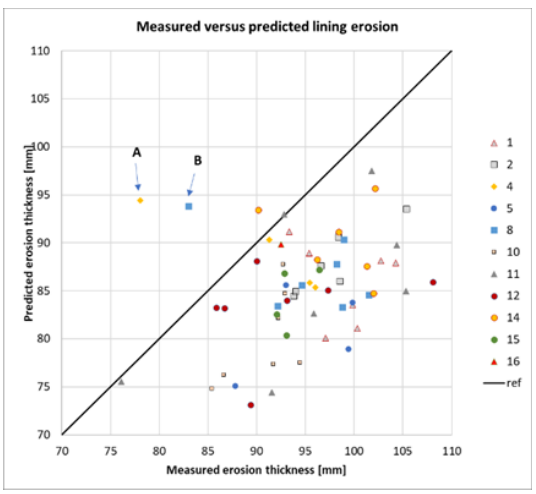


Figure 34 Close-up of comparison between per heat averaged measured and predicted erosion thickness at time of demolition of wear lining. Symbols represent different ladle numbers.

The results shown above show that we have a model that can reproduce data in an acceptable manner and may contribute significantly to supporting the operators in deciding whether one new use of the refractory is recommended or not.

Video demonstration of physics based model

A video demonstrating the model may be found here.

StreamPipes

This section provides high-level overview of industrial pipelines created for the Sidenor pilot using the StreamPipes platform.

First, it shortly describes StreamPipes – its purpose and features. Afterwards, it provides overview of created pipelines with elements that were used, both built-in and custom-built, alongside their brief description.

StreamPipes

Apache StreamPipes is a self-service (Industrial) IoT toolbox that enables non-technical users to connect, analyze and explore IoT data streams. It provides a set of built-in elements and an editor that enable users to easily create and manage pipelines. Additionally, it provides means of creating custom pipeline elements, further increasing its functionalities and applicability.

StreamPipes' core features:

- Connecting with IoT data – **Data Sets** and **Data Streams** types of pipeline elements enable users to connect data with the following pipeline elements, via the built-in StreamPipes Connect library with support for generic protocols such as HTTP, Kafka, MQTT, Files or specific adapters for open data sources.
- Analyzing data – **Data Processors** type of pipeline elements enables users to process and analyze data using a real-time algorithm toolbox, ranging from simple filters up to pre-trained Neural Networks. Some of the included Data Processors are: Trend Detection, Peak Detection, Trigonometry Functions and Frequency Calculation.
- Exploiting data – **Data Sinks** type of pipeline elements enables users to trigger notifications, send data to third-party systems or visualize data. Some of the included Data Sinks are: Apache Kafka, Apache CouchDB, RabbitMQ and Email, PostgreSQL.

Pipeline #1 (acyclic data)

The main purpose of this pipeline is to create a path from the SIDENOR ***acyclic*** sensory data, to the Neural Network output, depicting the state of the tool through several pipeline elements. In addition, this pipeline triggers notifications when the value of a particular property goes above a certain threshold, displaying time between consecutive heats and information about each heat.

These pipeline elements include elements that simulate the connection between acyclic sensory data and the rest of the pipeline (Data Stream – marked in yellow), elements for processing, including one with the Neural Network (Data Processors – marked in green) and elements that provide output of the pipeline – visualization, notifications, etc. (Data Sinks – marked in blue) (Figure 35).

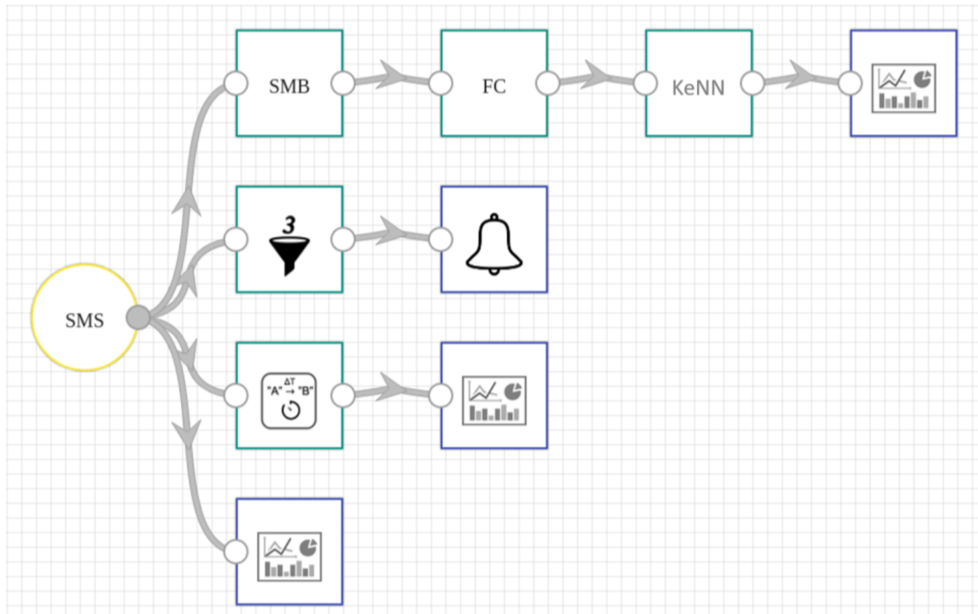


Figure 35 Developed pipeline (arrows represent data flow)

The first pipeline element, named SMS for *Sidenor Measurements Simulation*, simulates connection between acyclic sensory data and the rest of the pipeline, by reading row-by-row of .csv file. Each row represents one heat and contains values of parameters for said heat. Simulates real-world situation in which, after each heat, data measured for said heat would be sent to the pipeline.

The first element in the top row, named SMB for Sidenor Measurements Buffer, represents a buffer that orders heats according to ladle, cycle and phase they come from. Our Neural Network requires input that is calculated based on all heats from the same ladle, cycle and phase. For each new heat, this element adds it to the list that holds heats that came from same ladle, cycle and phase and then forwards the entire list to the next element.

The following element in the top row, named FC for *Factored Contributions*, serves as a pre-processing element that prepares forwarded data for inference done by next pipeline element with integrated Neural Network model. This element calculates the "contribution" of each parameter to the wear of the ladle and outputs a vector where each value represents the "contribution" of the corresponding parameter.

Said vector represents input to the KeNN (*Keras Neural Network*) element. It represents input of both the mentioned pipeline element and the Keras Neural Network model loaded within it. After inference, NN outputs class which states the condition of the ladle (*lower/higher degradation than predefined threshold*). This output value is forwarded to the visualization element.

The first element in the second row, named *Numerical Filter*, filters events (heats, in this context) based on the value of the selected numerical property. In this case, it filters heats based on the consumed electricity (*Kwh_rr*) - if a heat has the value of the consumed electricity greater than the specified threshold, it gets forwarded to the next element (*notification element*), otherwise, it gets ignored.

The first element in the third row, named *Task Duration*, computes the time difference between two events (heats, in this context). It forwards the measured difference to the visualization element.

The last element in the second row, named *Notification* and marked with *bell*, displays a notification in the UI panel of StreamPipes. In this pipeline, it displays a notification regarding heats that have consumed an electricity value above a certain threshold (Figure 36).

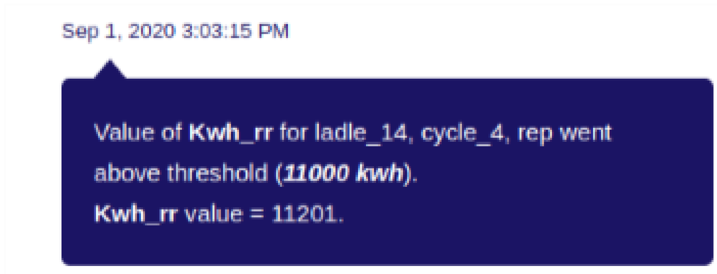
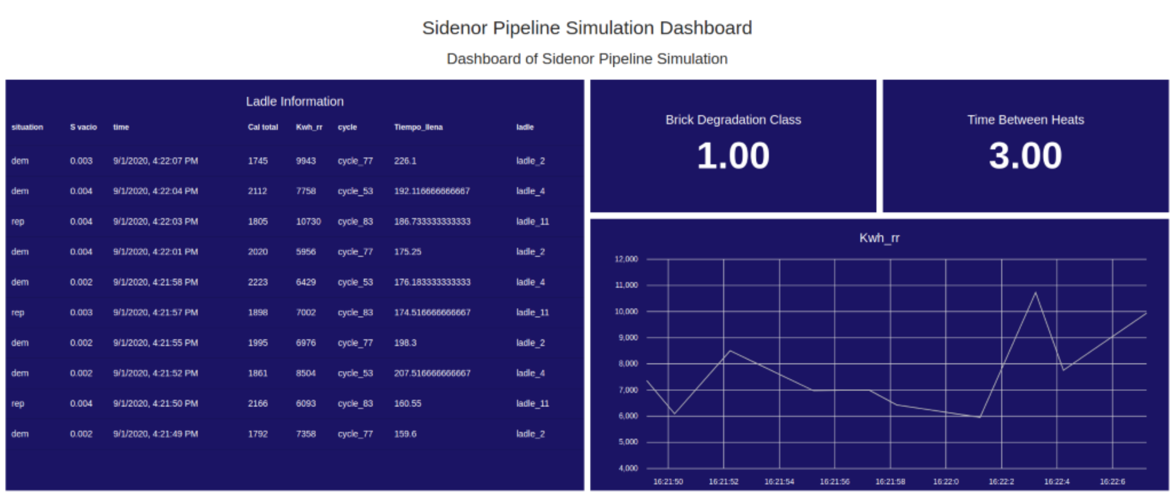


Figure 36 Displayed notification

The rest of the elements colored in blue, named Dashboard Sink, serve as a visualization tool. They visualize data streams in the StreamPipes dashboard (Figure 37).



situation	S vacio	time	Cal total	Kwh_rr	cycle	Tiempo_llena	ladle
dem	0.003	9/1/2020, 4:22:07 PM	1745	9943	cycle_77	226.1	ladle_2
dem	0.004	9/1/2020, 4:22:04 PM	2112	7758	cycle_53	192.116666666667	ladle_4
rep	0.004	9/1/2020, 4:22:03 PM	1805	10730	cycle_83	186.733333333333	ladle_11
dem	0.004	9/1/2020, 4:22:01 PM	2020	5956	cycle_77	175.25	ladle_2
dem	0.002	9/1/2020, 4:21:58 PM	2223	6429	cycle_53	176.183333333333	ladle_4
rep	0.003	9/1/2020, 4:21:57 PM	1898	7002	cycle_83	174.516666666667	ladle_11
dem	0.002	9/1/2020, 4:21:55 PM	1995	6976	cycle_77	198.3	ladle_2
dem	0.002	9/1/2020, 4:21:52 PM	1861	8504	cycle_53	207.516666666667	ladle_4
rep	0.004	9/1/2020, 4:21:50 PM	2166	6093	cycle_83	160.55	ladle_11
dem	0.002	9/1/2020, 4:21:49 PM	1792	7358	cycle_77	159.6	ladle_2

Figure 37 Visualization of outputs of Keras Neural Network (Brick Degradation Class), Task Duration (Time Between Heats) and Sidenor Measurements Simulation (Ladle Information, Kwh_rr)

Pipeline #2 (cyclic data)

This pipeline is meant to process ***cyclic*** *(transient)* sensory data using various analytical methods. Currently, we have implemented an element that calculates MEWMA (Multivariate Exponentially Weighted Moving Average) over input data and forwards its output.

This pipeline's elements include: an element that simulates the connection between the cyclic sensory data and the rest of the pipeline (Data Stream – marked in yellow), element(s) for analytical processing (Data Processors – marked in green) and element(s) that provide output of the pipeline – visualization (Data Sinks – marked in blue) (Figure 38).

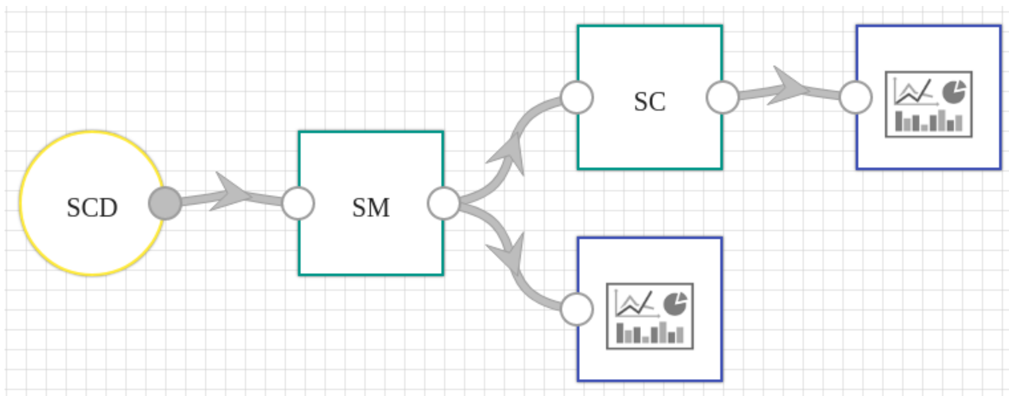


Figure 38 Developed pipeline (arrows represent data flow)

The first pipeline element, named SUCM for *Sidenor Unbatched Cyclic Measurements Simulation*, simulates the connection between cyclic sensory data and the rest of the pipeline. Each event represents cyclic data for an entire heat, hence *unbatched – not divided into batches*. This first element simulates real-world situation in which, after each heat, the data measured for said heat would be sent to the pipeline.

The following element, named SM for *Sidenor MEWMA*, receives cyclic data for an entire heat and calculates MEWMA for it, outputting result of its execution, including detected anomalies along with the information which parameters caused it.

The third element, named SC for the *Sidenor CEP*, performs Complex Event Processing (CEP) using the Siddhi engine. It performs a CEP query on the received data. It applies complex logic to the "main" outputs of this pipeline (*results of various analytical methods*). In addition, this element provides the point of connection for this and the previous pipeline. Complex Event Processing can be applied on outputs of multiple pipelines connecting them into one complex pipeline.

The pipeline elements at the end are used for visualization of the performed analysis. At the moment, we are presenting raw data, as it is received (*JSON format*) (

Figure 39).

Heat Information

{"timestamps":["2019-03-05 20:37:40","2019-03-05 20:37:41","2019-03-05 20:25:19","2019-03-05 20:25:18","2019-03-05 20:32:30","2019-03-05 20:37:48","2019-03-05
20:32:31","2019-03-05 20:37:49","2019-03-05 20:37:47","2019-03-05 20:37:42","2019-03-05 20:37:43","2019-03-05 20:17:37","2019-03-05 20:17:38","2019-03-05 20:27:21","2019-
03-05 20:37:51","2019-03-05 20:27:22","2019-03-05 20:39:30","2019-03-05 20:37:50","2019-03-05 20:27:20","2019-03-05 20:17:25","2019-03-05 21:38:13","2019-03-05
20:19:08","2019-03-05 20:17:26","2019-03-05 20:19:07","2019-03-05 20:17:27","2019-03-05 20:19:06","2019-03-05 20:27:18","2019-03-05 20:27:19","2019-03-05 20:46:14","2019-
03-05 21:22:12","2019-03-05 21:22:13","2019-03-05 20:17:51","2019-03-05 20:46:19","2019-03-05 21:22:10","2019-03-05 21:22:11","2019-03-05 20:27:14","2019-03-05
20:46:15","2019-03-05 20:27:15","2019-03-05 20:46:16","2019-03-05 20:27:16","2019-03-05 20:46:17","2019-03-05 20:27:17","2019-03-05 20:46:18","2019-03-05 20:13:17","2019-
03-05 20:35:03","2019-03-05 20:13:18","2019-03-05 20:35:02","2019-03-05 20:13:19","2019-03-05 20:17:14","2019-03-05 20:35:05","2019-03-05 21:22:09","2019-03-05
20:35:04","2019-03-05 20:35:06","2019-03-05 20:18:50","2019-03-05 20:46:22","2019-03-05 20:18:51","2019-03-05 20:46:23","2019-03-05 20:18:10","2019-03-05 20:18:54","2019-
03-05 20:19:23","2019-03-05 20:18:11","2019-03-05 20:19:22","2019-03-05 20:17:40","2019-03-05 20:18:52","2019-03-05 20:19:21","2019-03-05 20:25:20","2019-03-05
20:46:20","2019-03-05 20:18:53","2019-03-05 20:46:21","2019-03-05 20:37:37","2019-03-05 20:37:38","2019-03-05 21:03:37","2019-03-05 21:03:38","2019-03-05 21:03:39","2019-
03-05 21:38:32","2019-03-05 20:37:39","2019-03-05 21:38:33"],"ladle":"ladle_2","cycle":"cycle_3","situation":"rep","root-causes":[["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],
["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],
["Caudal_gas"],["Caudal_gas"],["Consumo_electrico"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],
["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],
["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],
["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],
["Caudal_gas"],[],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],
["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"],["Caudal_gas"]]}

Figure 39 Raw JSON representation of MEWMA output

Knowledge Graph Based Solution for Cognition

An advantaged functionality of Cognitive Digital Twin (CT) is "cognition": a digital twin (DT) able to "reason" and "act" taking into consideration of the data as well as the states and results of other hybrid twins (HT). To enable cognition, the CT not only composes the data and information from DTs and HTs but also extract knowledge from the domain expert in order to perform self-learning, situation analysis and decision-making in the same way that the human operators handle the information. Therefore, knowledge representation is an important requirement to support this vision. The knowledge must be integrated and represented formally so that the CT can process, analysis and reasoning automatically. In this pilot, we employ semantic knowledge graph technologies for knowledge representation. Accordingly, SINDIT[2] (SINTEF Digital Twin Framework) - a knowledge graph based Digital Twin Framework – was employed to achieve the "cognition" for SIDENOR pilot.

[2] https://github.com/SINTEF-9012/SINDIT

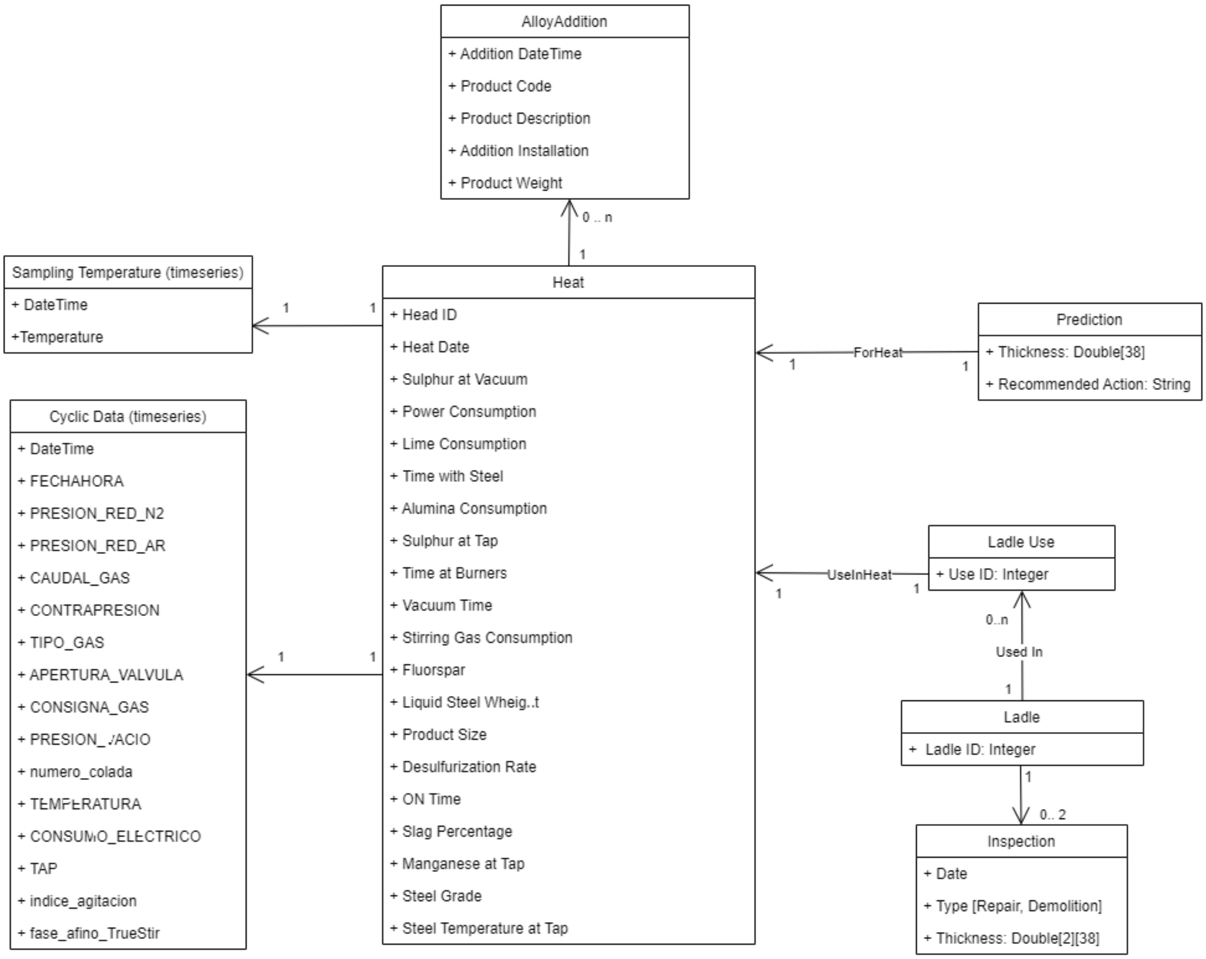


Figure 40 Information Model of SIDENOR Pilot.

Figure 40 illustrates the information model of the pilot which contains both acyclic data (e.g., process parameters) and cyclic data (e.g., timeseries sensor data). The data was used by hybrid digital twins (e.g., physical models, hybrid models) to simulate and make prediction about the conditions of the ladles. The input data as well as output of these hybrid models are integrated into SINDIT in order to support automatic reasoning and decision-making support for the human operator.

Figure *41* illustrates the corresponding SINDIT knowledge graph which contain the Ladles (blue node), Ladle Use (pink), Heats (dark green), Alloy Additions (purple) and prediction models output (light green). To import the knowledge graph to SINDIT, we develop mapping rules using RML[3] to

[3] https://rml.io/specs/rml/

map structural data from relational database to semantic knowledge graph. Particularly, tools such as Mapeathor[4] and Morph-KGC[5] were employed to develop and execute the mappings.

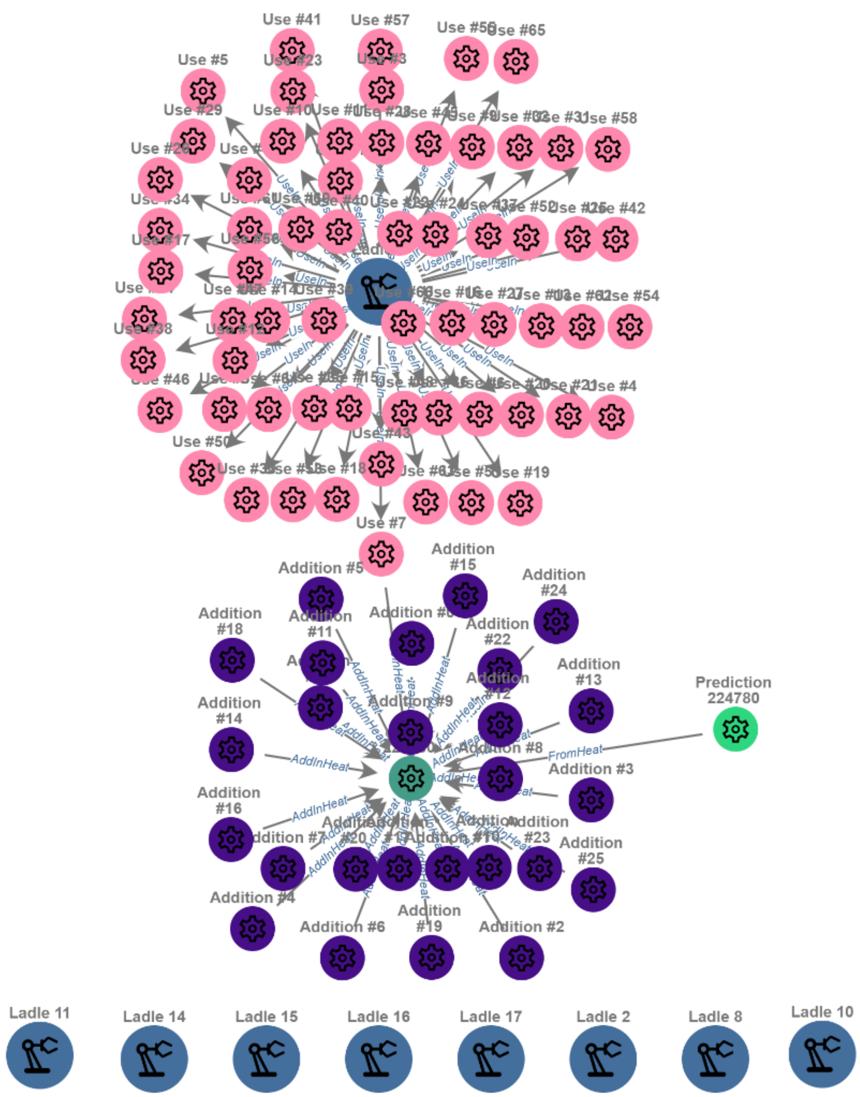


Figure 41 SINDIT Knowledge Graph

Domain expert knowledge for making assessment (e.g., whether to repair or demolish the ladles) is also integrated into SINDIT under the form of reasoning rules (IF-THEN rules). Such domain knowledge was illustrated in Figure 42. Figure 43 is a simplified version of a rule to inform the

[4] https://github.com/oeg-upm/mapeathor

[5] https://github.com/morph-kgc/morph-kgc

operator to carefully analyse the ladle before using it for the next heat as it is reaching the repair or demolish cycle. Figure 44 is another example of the reasoning rule to notify the operator whenever the predicted thickness of the ladle is below operational range (e.g., 50 mm).

Domain Knowledge
+ Maximum Wear Brick Thickness = 180 mm
+ Minimum Wear Brick Thickness = 50 mm
+ Average ware rate = 3 mm/heat
+ Repair Use Range = [33, 46]
+ Demolition Use Range = [69, 85]
+ Operation Thickness Range = (70, 180] mm
+ Analysis Required Range = (50, 70] mm
+ Repair/Demolish Range = [0, 50] mm

Figure 42 Domain Expert Knowledge for making decision on the Ladle.

```
[?PredictionModel, :hasRecommendAction, "Analysis Required" ] :-
[?PredictionModel, :assetType, "PredictionModel"],
[?PredictionModel, :FromHeat, ?Heat],
[?Use, :UseInHeat, ?Heat],
[?Use, :UseNum, ?UseID],
NOT [?PredictionModel, :hasRecommendAction, "Repair or Demolish" ],
FILTER((?UseID >= 33 &&  ?UseID <= 46) || (?UseID >= 69 &&  ?UseID <= 85))
```

Figure 43 Reasoning Rule: Recommend analysing the ladle if the use number is within a specific range.

```
[?PredictionModel, :hasRecommendAction, "Repair or Demolish" ] :-
[?PredictionModel, :assetType, "PredictionModel"],
[?PredictionModel, :hasPredictionResult, ?Result],
[?Result, :hasPredictionValue, ?value],
FILTER(?value <= 50) .
```

Figure 44 Reasoning Rule: Recommend to repair or demolish the ladle if the predicted thickness is below 50mm.

The reasoning result is also integrated into the knowledge graph and visualised in a separate window in SINDIT in order to support the operator to make decision on the ladles after every heat as can be seen in Figure 45.

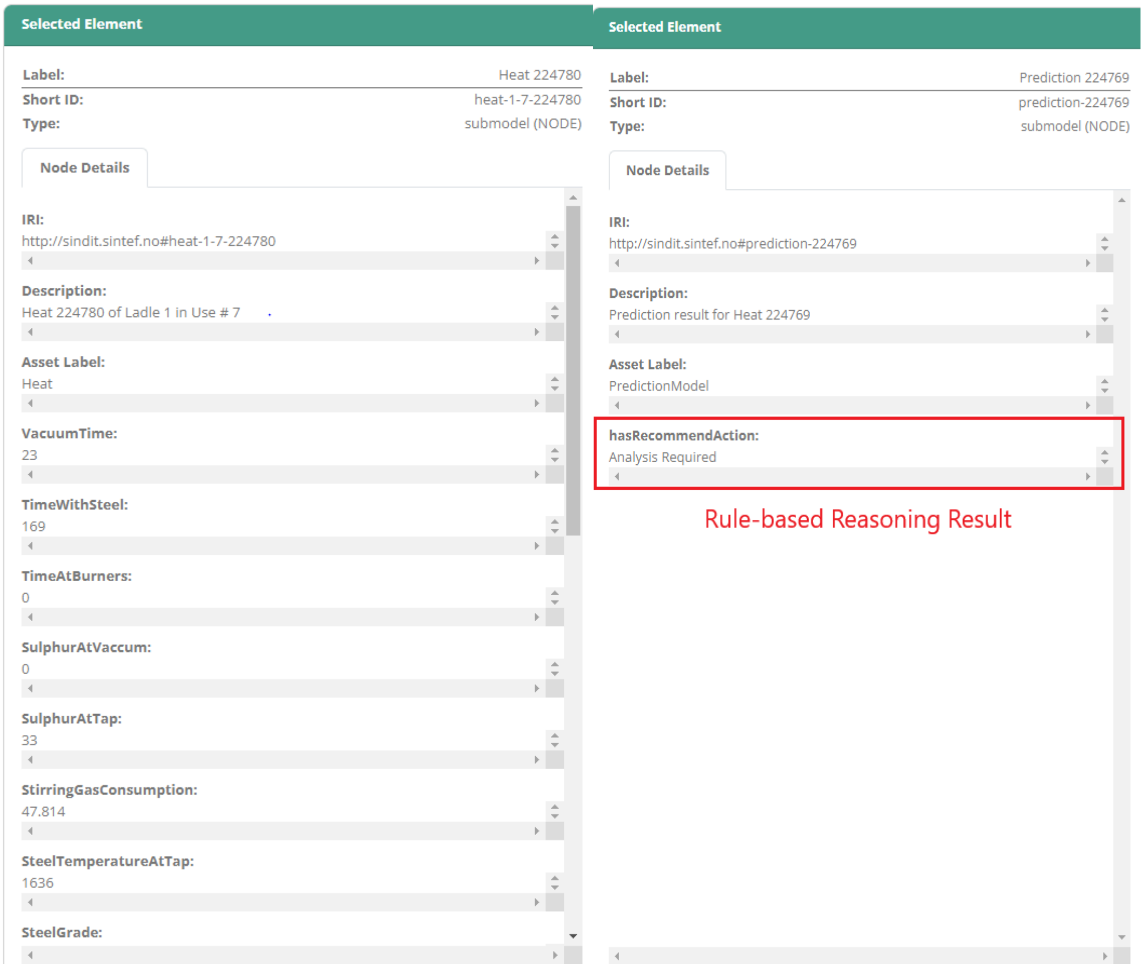


Figure 45 Properties of selected nodes shown in a separate window: (a) Heat node and (b) Prediction node.

FA^{3}ST (Fraunhofer AAS Tools for Digital Twins) Service

Fraunhofer IOSB has developed a digital twin for a ladle based on Asset Administration Shell (AAS) specifications ("Details of the asset administration shell - part 1 version 3.0rc01," [Online]. Available: https://industrialdigitaltwin.org/wp-content/ uploads/2021/09/07 details of the asset administration shell part1 v3 en 2020.pdf; "Details of the asset administration shell - part 2 version 1.0rc02," [Online]. Available: https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/Details of the Asset Administration Shell Part2 V1.html). The digital twin was developed by applying the following steps:

1. The AAS-compliant digital twin type 1 was modeled, i.e. the AAS model was created as a file. The AAS is a static file serialized according to Part 1 of the AAS specification. We used the AASX Package Explorer (https://github.com/admin-shell-io/aasx-package-explorer) to model the digital twin manually and stored it as a JSON file.

To develop a model, we analyzed all the data types provided by Sidenor. This included not only the acyclic data, cycle data, and lining data, but also the created models (e.g., physics-based). Additionally, we analyzed the existing AAS submodel templates (https://industrialdigitaltwin.org/en/content-hub/submodels) to be as standard compliant as possible.

The AAS model contains (i) general information about the ladle, such as the number of heat runs, the original thickness of the bricks, or information from inspections; (ii) operating parameters of the current and historical heat runs of the ladle; (iii) prediction results of the simulations, (iv) meta-information about the simulations used, etc. We used the AAS properties to represent single measurement per parameter. To represent the time-series measurements per parameter, collected every second during each ladle usage, we used the template “SubmodelTemplate Time Series Data” (SMT). This is the AAS submodel template currently in development that aims to provide a standardized metamodel and API for modelling time series data. In Figure 46 the AAS model is partially shown within the AASX Package Explorer.

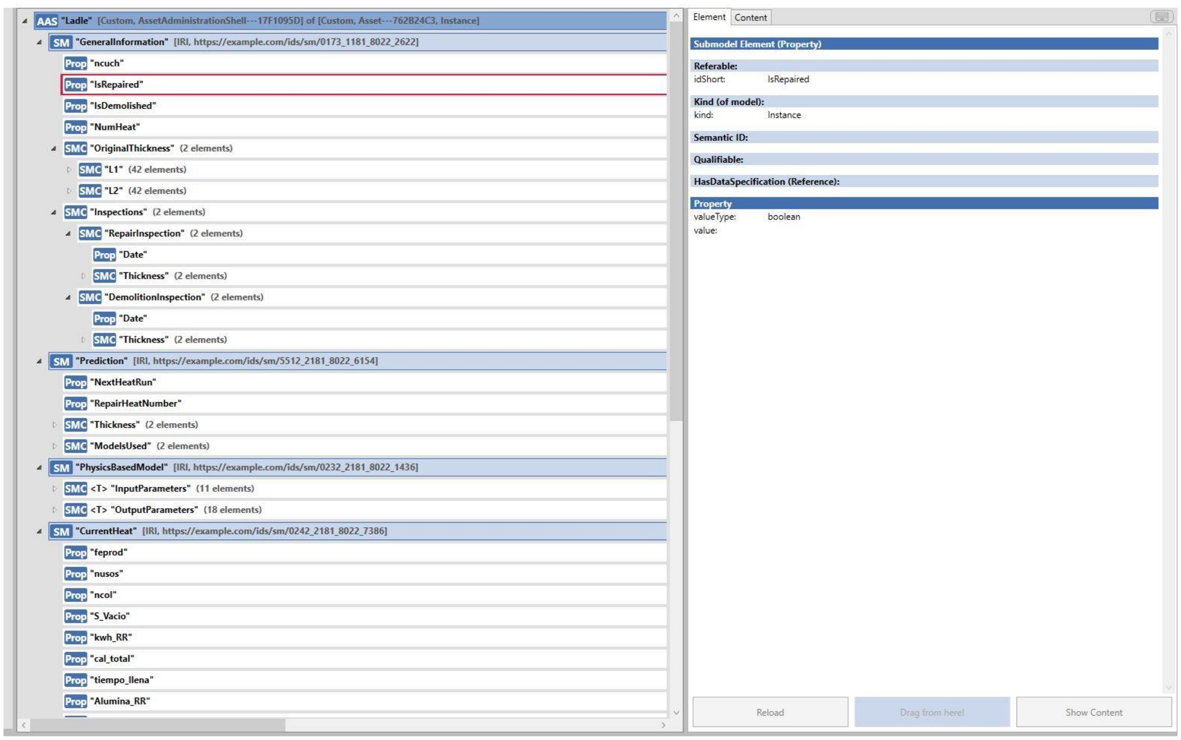


Figure 46:AAS model opened in AASX Package Explorer

2. We used the FA³ST Package Explorer Converter (https://github.com/FraunhoferIOSB/FAAAST-Package-Explorer-Converter) to convert the Sidenor AAS JSON file created with AASX Package Explorer in step 1 to a FA³ST-compatible

version. We note that the AASX Package Explorer uses AAS meta model v2.x while FA³ST uses v3.x.

3. We used the FA³ST service (https://github.com/FraunhoferIOSB/FAAAST-Service) to create the AAS compliant digital twin type 2. The AAS is a software component that provides a standardized API according to Part 2 of the AAS specification. Information is exchanged automatically via external software that communicates with the AAS via the API.

This means that we created the software representation of the ladle. This was done by starting the FA³ST service with the AAS model created in step 2. The result is the digital twin for Sidenor with an AAS-compliant digital twin API interaction of the digital twin with the outside world. More information on FA³ST could be found in D4.4 deliverable.

4. The final step was to ensure that the digital twin for Sidenor is always synchronized with the underlying physical asset, i.e. the ladle. To do this, we connected the digital twin of Sidenor developed in step 3 to the data source. In this case, this was the InfluxDB provided by Nissatech, which contains real-time data. The connectivity was achieved by (i) creating a configuration file that defines how to access the data stored in the external database and by (ii) starting the FA³ST Service with both the Sidenor AAS model and the configuration file. More information on that could be found in the D5.4 deliverable.

The key advantage of the AAS-compliant digital twin for Sidenor is the improved interoperability and thus the reduced effort required to integrate the digital twin into the potential applications.

Measurable KPIs and Final impact

The model gives an objective prediction which takes into account the history of the ladle. Up to the beginning of the project the refractory was rebuilt after 69 to 81 heats, but we detected that in some occasions the ladles could have worked longer. In other words, using the model developed within the COGNITWIN, we are now able to predict if the ladle's lining will last one more heat taking into account the safety limits defined by Sidenor.

Due to cybersecurity restrictions, the possibility of giving Nissatech access to Sidenor's production data was forbidden so the partners working in this WP agreed to send the data to a FPT server. Moreover, the possibility of sending the production data and how to manage the files took longer than expected. Nevertheless, the problem was solved and the model was able to run online with the cyclic and acyclic production data sent.

After a deployment delay at the end of the project, the cognitive digital twin pipeline has been installed with Sidenor to support the future refractory wear decision making

The estimation was calculated taking into account the results obtained with the predictive model. The model run the heats already produced and these results were compared with the real measurements. The improvement on the refractory life could reach 14% what means saving around 54.000€ per year, taking into account the refractory cost of 2021.

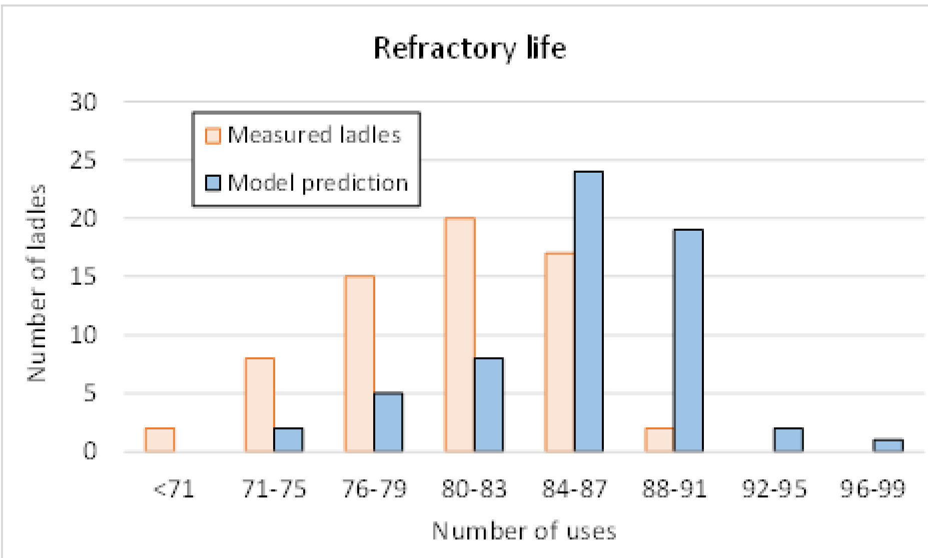


Figure 47:Comparison between measured and predicted refractory thickness

Conclusion and Summary

The Hybrid Digital Twin with enhancement to a Cognitive Digital Twin has in particular taken advantage of the use of a physical model combined with a data driven model. As part of this development based on PPBM (Pragmatism in physics-based modelling), a physics-based model for ladle lining lifetime has been developed. The model is using physical conservation principles and has been tuned to operational data. In addition to its prediction capability, the model has been proven useful to filter or clean data which does not comply with the basic conservation principles. ML and regression models have been explored, but these model does not have the capability to explain observations without excessive amounts of data.

A specific challenge with ladle lining erosion was that no quantitative state of the refractory exists except at the time of lining repair (only part information) and the time of ladle demolition ("complete" information). The "complete" information consists of a measurement of the most worn brick at each brick row in the ladle. This is done for each of two half sectors of the ladle. The average erosion and the variation of erosion at each height has recently been mapped, partly because of question resulting from the PPBM-based model. This information will be very useful for further development of the model. The model predicts erosion profiles which are ensemble averages. Accordingly, a safety margin has to be applied. Some uncertainty in the data must be acknowledged and the model prediction is therefore taken as a recommendation for the ladle operator.

The model has been integrated into a StreamPipes-based application, where both historic and online process data is available. Prediction is here made available to operators. Historic and current prediction data is saved in the database and is available for other tools. One specific tool which has been explored is a knowledge-graph based solution for adding new cognition elements.

The work has led to new understanding about phenomena that impact outer surface temperatures of a ladle, and quantitative information about how much different phenomena contribute to ladle lining erosion has been obtained. This will help operators the better understand what their decision windows are.

The developed model have several possibilities for further improvement and will have value for many other steel companies. The generic version of the PPBM-based model is available at github.com[6]. This model has been included in the final Sidenor digital twin pipeline.

The final Sidenor pilot demonstrator is described in the COGNITWIN Toolbox [1] with the Sidenor digital twin pipeline description [6] and the final Sidenor demonstrator video [7] . This is also further described in the final public deliverable D6.4 Best "Digital Twins" practices report [2].

4 NOKSEL – Pilot

Introduction to NOKSEL & Process description

NOKSEL is one of the leading steel pipe manufacturers of Turkey with its plants located in Iskenderun at the region Hatay and in Hendek at the region Sakarya since 1987. Noksel serves domestic and international markets by manufacturing spiral welded steel pipes for petroleum, gas, water and piling industries.

Turkey is the biggest producer of spirally welded steel pipe in Europe with production capacity of 5.2 million tons of steel pipe per year. In Turkey NOKSEL continuously is in the First Biggest Industrial

[6] https://github.com/SINTEF/refractorywear

Companies List issued research by İstanbul Chamber of Industry since 1996 b and the second largest in company steel pipe industry. Besides NOKSEL Turkey, Noksel España S.A. was established in Spain in 2008. All manufacturing plants of NOKSEL are planned mainly for the production of the pipes in API standard. NOKSEL produces pipes in accordance with AWWA, DIN, BS, ASTM, ISO, EN, UNI and AFNOR standards to serve the petroleum, gas, waterline and construction industries. With a full commitment to superior performance, the Company constantly strives to ensure that its quality policies and principles are in full compliance with all national and international regulations and standards. To optimize information management, NOKSEL has been using the SAP system for its own business operations and MIS systems since 2005 and the company continues to invest in digitalization of its premises.

Noksel's pilot case aims at the development of Digital Twin for the production process of Spiral Welded steel Pipes (SWP). The digital twin will collect, integrate and analyse multiple sensors' data streams in real-time, and enable predictive maintenance by a smart condition monitoring system. Real-time data acquisition, communication networks for monitoring, and automated recommendation generation are among the key innovative features of this pilot.

Smart components that use sensors to collect in real-time condition of the equipment's and their position are connected by a cloud-based system that processes the collected data streams. Theses inputs are analysed against business and other contextual data through smart visualization systems. The digital twin models allow joining the physical and the virtual worlds to create a new networked layer in which intelligent objects interact with each other to virtualize the steel pipe manufacturing process on the Spirally Welded Pipe (SWP) machinery shown in below Figure 48: Schematic lay out of SWP Machinery and Figure 49: Photo of SWP Machinery.

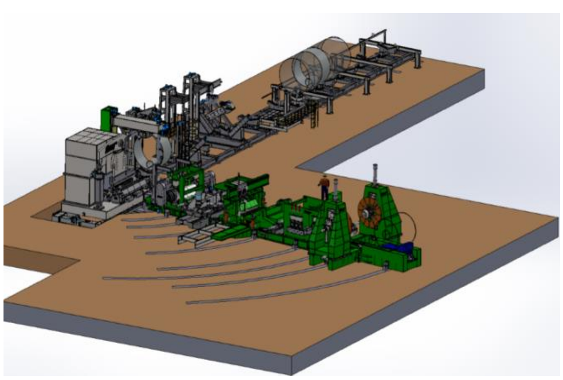


Figure 48:Schematic lay out of the SWP Machinery

Figure 49:Photo of the SWP machinery.

SWP machinery is used in the production of spirally welded steel pipe. In these machines the hot rolled sheets are combined by turning at a certain angle or flatly using the submerged arc welding

(SAW) method. The general name of this process is spirally welded steel pipe production process, in Figure 50.

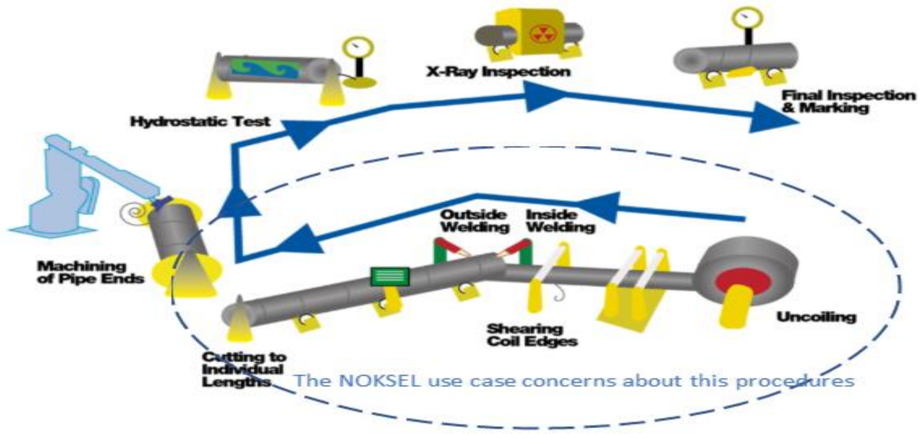


Figure 50: NOKSEL's use case processes

SWP is a multi-step, manufacturing process that consists of the following steps:

1) Preparation of hot rolled coils,
2) Coil ends welding, (skelp end welding)
3) Edge preparation of the coils,
4) Transforming the coils to pipe, (pipe forming)
5) Welding operations,
6) Pipe production,
7) Pipe cutting,
8) Repair welding,
9) NDT(Non-Destructing) testing,
10) Acceptance of the pipes,
11) Coating and lining,
12) Final testing and acceptance of the pipes.

The objectives of the NOKSEL case on the SWP machinery are threefold: 1) real-time condition monitoring, 2) predictive maintenance, and 3) digital twin generation.

Due to the very large size of SWP machinery, and the high number of its components, development of an user-friendly software for SWP's real-time condition monitoring is a challenge. The condition monitoring has been realized by means of a platform. The platform was connected with PLC, SCADA and ERP interfaces of the production plant. In the aimed platform, data gathered from different sources, such as sensors, operational data input, automation system input were analyzed for understanding current situation of the SWP machinery. Using the digital twin technology, SWP's outputs were shown to the operator. Smart components that use sensors to gather data about real-time status, working condition, or position have been connected to a cloud-based system.

By applying AI/ML/DL based analytics and computation, and using the data acquired by the platform, smart predictive maintenance is aimed via cognitive digital twin. The Digital Twin, has l visualized maintenance predictions real-time, and notified personnel with current condition of the machinery, abnormalities and alarms.

Both data driven and model driven twins were generated for the case. Both synthetic and real data collected from the platform have been developed as planned to be used for data driven twin generation. The 1st order, or the process modeling were used to generate the model driven twin. By combining the data and the model driven twins, a hybrid digital twin was created. Adding the cognitive elements to conduct predictive maintenance by enhancing the digital twin via AI/ML/DL components, a cognitive digital twin has been generated as aimed (Albayrak and Unal, 2021).

User stories

The users of the system are Maintenance Operators and Maintenance Managers. The following sentences brief the requirements of the users in forms of user stories.

- As a user, I want to monitor the condition of the system so that I follow the current status of the machine.
- As a user, I want to learn the remaining useful life of machine so that I conduct predictive maintenance and take actions before the machine's breakdown.
- As a user, I want collect data from sensors and PLCs so that the acquired data set is used for condition monitoring.
- As a user, I want conduct data pre-processing so that the data set becomes ready to be used for ML/DL model training.
- As a user, I want to execute selected ML/DL models so that I detect anomalies in real time.
- As a user, I want visualize past sensor data so that I see the data trend of sensors.
- As a user, I want visualize operational data so that monitor the real-time condition of the SWP machinery.
- As a user, I want to generate synthetic data for electromechanical components so that I create machine break down and/or faulty cases.

The pilot requirements written above as user stories are detailed in forms of user stories/use cases by Table 6 - Table 15.

Table 6: Use Case NOKSEL-UC-00

Use Case Template	Description

Use Case Name	Condition Monitoring
Use Case ID	NOKSEL-UC-00
User story expression of use case	As a user, I want to monitor condition of the system so that I follow the current status of the machine.
Goal	To monitor the condition of the machine and environment
Measurable KPIs for the goal (if any)	The latency between data acquisition and data visualization < 200ms.
Actors and stakeholders involved	Maintenance operator, Maintenance Manager
Input data	Sensor data, and data from the PLC (alarm and status data) used for condition monitoring
Output data / actions	Quality data that is cleaned, scaled, filled in, analyzed, normalized, and preprocessed and displayed to the user
Summary description – Main success scenario	1. User selects the machine component for which the condition is to be monitored. 2. The system displays the 3D view of the selected machine component and presents the stream data measured by the sensors implemented on the machine.
Extensions, exceptions, variations	If the user selects a specific place (hotpoint) on the machine component's 3D view, the system displays the sensor values that are implemented only on the selected place.
Possible generalisation of use case	The use case can be generalized for condition monitoring of other assets and systems on IoT platforms.
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps: Digital Twin Data Acquisition, Digital Twin Representation, Digital Twin Visualization and Control

Table 7: Use Case NOKSEL-UC-01

Use Case Template	**Description**
Use Case Name	Predictive Maintenance
Use Case ID	NOKSEL-UC-01
User story expression of use case	As a user, I want to learn the remaining useful life of machine so that I conduct predictive maintenance and take actions before the machine's breakdown.
Goal	To estimate RUL and to be able to conduct predictive maintenance
Measurable KPIs for the goal (if any)	The models should provide TP = 100%. The models should provide FP = 100%.

	(Both true negatives and false negatives will be 0, thus Type I and Type II errors will be 0)
Actors and stakeholders involved	Maintenance operator, Maintenance Manager
Input data	Sensor data, and data from the PLC (alarm and status data) used for condition monitoring and data-driven AI models
Output data / actions	Estimates on the RUL is provided.
Summary description – Main success scenario	1. The system runs ML/DL algorithms on the stream data 2. The system displays the anomalies in data 3. The system calculates RUL and informs the user.
Extensions, exceptions, variations	None
Possible generalisation of use case	The use case can be generalized for predictive maintenance of other assets and systems flow
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps: Digital Twin Data Acquisition, Digital Twin Representation, Hybrid/Cognitive Digital Twin Generation, Digital Twin Visualization and Control

Table 8: Use Case NOKSEL-UC-1

Use Case Template	**Description**
Use Case Name	Industrial Big Data Processing
Use Case ID	NOKSEL-UC-1
User story expression of use case	As a user, I want conduct data pre-processing so that the data set becomes ready to be used for ML/DL model training.
Goal	To prepare data for ML/DL training
Measurable KPIs for the goal (if any)	There will be no empty cell in the data set All data fields in the data set will be labelled
Actors and stakeholders involved	Maintenance operator or maintenance manager
Input data	Sensor data, and data from the PLC (alarm and status data) used for remaining useful life estimation and type of ML/DL algorithm to be trained
Output data / actions	Quality data that is cleaned, scaled, filled in, analyzed, normalized, and preprocessed and get ready to be used for ML/DL model training
Summary description – Main success scenario	1. Data set is selected and uploaded. 2. Rules to fill in empty data is selected by the user.

	3. Data is normalized. 4. Mean, median and variation of the data set is calculated. 5. Processed data is stored.
Extensions, exceptions, variations	None
Possible generalisation of use case	The objective function of the ML/DL models to be trained may be functions other than predictive maintenance.
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps: Digital Twin Data Acquisition, Digital Twin Representation

Table 9: Use Case NOKSEL-UC-2

Use Case Template	Description
Use Case Name	Real time anomaly detection using pretrained ML/DL models
Use Case ID	NOKSEL-UC-2
User story expression of use case	As a user, I want to execute selected ML/DL models so that I detect anomalies in real time.
Goal	To estimate anomalies on stream data.
Measurable KPIs for the goal (if any)	None
Actors and stakeholders involved	Maintenance operator and/or maintenance manager
Input data	Sensor data, PLC data, alarm and status data
Output data / actions	Anomalies are detecting by the selected ML/DL models and the results are graphically displayed
Summary description – Main success scenario	1. The users select the pre-trained ML/DL model set to be executed on the stream data. 2. The selected models are executed on the stream data. 3. The results of the selected ML/DL models are calculated and displayed to the user
Extensions, exceptions, variations	If the user does not select any ML/DL model a warning message is displayed to the user in order to state that at least one model must be selected.
Possible generalisation of use case	The pretrained ML/DL models can be generated for purposes other than predictive maintenance and anomaly detection

Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps: Hybrid/Cognitive Digital Twin Generation, Digital Twin Visualisation and Control

Table 10: Use Case NOKSEL-UC-3

Use Case Template	Description
Use Case Name	Big Data Visualization
Use Case ID	NOKSEL-UC-3
User story expression of use case	As a user, I want visualize past sensor data so that I see the data trend of sensors.
Goal	To display collected past sensor data in graphs.
Measurable KPIs for the goal (if any)	The zoom in and zoom out of displayed data should be <200ms.
Actors and stakeholders involved	Maintenance operator, maintenance manager
Input data	Past sensor data collected
Output data / actions	Sensor data displayed in graphics with respect to time
Summary description – Main success scenario	1. The user wants to visualize past sensor data in graphs 2. The data is displayed to the user
Extensions, exceptions, variations	When the user zooms in and/or zooms out the time scale of the x-axis changes, and the data is displayed in the graphs accordingly
Possible generalisation of use case	The visualized big data does not need to belong to the sensors in the NOKSEL pilot, any time series big data can be graphically visualized.
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps: Digital Twin Representation, Digital Twin Visualisation and Control

Table 11: Use Case NOKSEL-UC-4

Use Case Template	Description
Use Case Name	Operational Data Visualization
Use Case ID	NOKSEL-UC-4
User story expression of use case	As a user, I want visualize operational data so that monitor the real-time condition of the SWP machinery.

Goal	To monitor operational data regarding the SWP components
Measurable KPIs for the goal (if any)	Latency in visualization > 200ms
Actors and stakeholders involved	Maintenance operator, maintenance manager
Input data	Sensors and PLC
Output data / actions	Graphical representation of operational data
Summary description – Main success scenario	1. The user selects the component of the SWP machine for which the condition is to be monitored. 2. The system displays the 3D model of the component and graphics associated to the hotspots on the component. 3. The user selects the hotspot. 4. The system displays the values of the sensors associated with the selected hotspot in suitable graphics.
Extensions, exceptions, variations	None
Possible generalisation of use case	The component visualized may belong to a different machine other than SWP, and the sensor set installed on the component may be different.
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps (Delete if not related): Digital Twin Visualisation and Control

Table 12: Use Case NOKSEL-UC-5

Use Case Template	Description
Use Case Name	New trained ML/DL model inclusion
Use Case ID	NOKSEL-UC-5
User story expression of use case	As a user, I want to add a new pre-trained ML/DL model so that it can be used for anomaly detection in real-time on-stream data
Goal	To include a new ML/DL model in the list of the models to be executed without updating the source code
Measurable KPIs for the goal (if any)	None
Actors and stakeholders involved	System Administrator, maintenance manager
Input data	Pre-trained ML/DL models
Output data / actions	Updated list of ML/DL model

Summary description – Main success scenario	1. The user selects the new model to be added to the list of pretrained ML/DL models 2. The system displays the list of files to be selected. 3. The user selects the model to be included in the model set. 4. The selected model is added to the list of pre-trained ML/DL models.
Extensions, exceptions, variations	If the file selected by the user is not a model, the system displays a warning message to the user.
Possible generalisation of use case	The pre-trained model can be for a different purpose other than anomaly detection. Any process can be called and added to a multi selection list.
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps: Hybrid/Cognitive Digital Twin Generation, Digital Twin Visualisation and Control

Table 13: Use Case NOKSEL-UC-6

Use Case Template	Description
Use Case Name	Synthetic data generation for generic electro mechanical components
Use Case ID	NOKSEL-UC-6
User story expression of use case	As a user, I want to generate synthetic data for electro mechanical components so that I create machine break down and/or faulty cases.
Goal	To generate synthetic data needed for the ML/DL training
Measurable KPIs for the goal (if any)	None
Actors and stakeholders involved	Maintenance Manager
Input data	Matlab model for the electro-mechanical components, parameters, real-data, limits
Output data / actions	.mat file including the faulty data generated
Summary description – Main success scenario	1. The user selects the Matlab model for the electro mechanical component. 2. The system loads the model. 3. The user installs the virtual sensors. 4. The system generates data including faulty cases.

	5. The system saves the data in .mat file.
Extensions, exceptions, variations	None
Possible generalisation of use case	Different model files and sensor types can be used to generate synthetic data
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps: Hybrid/Cognitive Digital Twin Generation

Table 14: Use Case NOKSEL-UC-7

Use Case Template	Description
Use Case Name	Industrial Big Data Acquisition
Use Case ID	NOKSEL-UC-7
User story expression of use case	As a user, I want collect data from sensors and PLCs so that the acquired data set is used for condition monitoring.
Goal	To collect machine and environment data
Measurable KPIs for the goal (if any)	Data are collecting at the determined frequencies
Actors and stakeholders involved	Maintenance operator, IoT Platform
Input data	Sensor data, and data from the PLC (alarm and status data) used for condition monitoring
Output data / actions	Quality data that is cleaned, scaled, filled in, analyzed, normalized, and preprocessed and get ready to be used
Summary description – Main success scenario	1. Data sources are mapped. 2. Data is transferred
Extensions, exceptions, variations	None
Possible generalisation of use case	The use case can be included in condition monitoring and ML/DL training
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps: Digital Twin Data Acquisition

Table 15:Use Case NOKSEL-UC-8

Use Case Template	Description
Use Case Name	Preventive Maintenance

Use Case ID	NOKSEL-UC-8
User story expression of use case	As a Maintenance Manager, I want the system to be able to control welding cell temperature so that the system acts proactively and eliminates human-in-the-loop for climate control operation.
Goal	To conduct preventive maintenance
Measurable KPIs for the goal (if any)	Energy consumption reduced by at least 10%
Actors and stakeholders involved	Actor: TIA CONTROL Stakeholders: Maintenance operators, Maintenance managers, Air Conditioner in the Welding Cell
Input data	Indoor environment, Welding Machine Generators temperatures, and Welding Machine Generators current value
Output data / actions	The command to manipulate air conditioner in the welding cell
Summary description – Main success scenario	1. For each welding machine generators current (voltage) are measured. If any of these values is greater than 0.1 amper, air conditioners are turned on. 2. Temperature in the welding cell and welding machine generator are measured. If for more than one hour welding cell temperature is in between 20-25 Celsius or if environment temperature is less than 20 Celsius then, the air conditioner status is set to close.
Extensions, exceptions, variations	The cause of machine breakdown may stem from a reason that is not associated with welding wire cut
Possible generalisation of use case	The use case is applicable to similar SWP machines that are used in the steel pipe production process industry
Use case analysis – related to which Digital Twin pipeline steps	This use case has been supported through the following Digital Twin pipeline steps (Delete if not related): Digital Twin Acquisition, Digital Twin Representation, Digital Twin Hybrid Analytics Models, Digital Twin Visualisation and Control

Current challenges

Table 16 - Table 19 collect the challenges for data handling, (hybrid) digital twins and visualization & control in the NOKSEL pilot.

Table 16: Pilot challenges for Digital Twin Data Acquisition/Collection for NOKSEL pilot

Sensors	• **Challenge**: The environment in which the sensors to be installed is a harsh environment. • **Requirement**: Multi sensor data will be collected for condition monitoring in near real time. • **Solution**: Special attention was given to thoroughly determine the sensor set and the locations of them to be implemented, by using information collected from the TEKNOPAR's expert engineers, on site domain experts, the machine manufacturers and the academicians. • **Challenge**: Provide necessary sensor data to monitor machinery • **Requirement**: Make use of suitable sensors, including vibration sensor and energy sensors for motors, in addition to pressure and temperature sensors. • **Solution**: Four new sensors have recently been installed, including Analog vibration sensor on DC motor and Energy analyzers.
Communication	• **Challenge**: There are two different PLCs to be used. • **Requirement**: The sensor data from two different PLC's will be collected. • **Solution**: Distributed sensor data is bundled, and uploaded over Profinet to an OPC gateway and MQTT broker. PN/PN coupling was applied to connect the PLC's. • **Challenge**: Low performance in data transfer from Kafka to PostgreSQL. • **Requirement**: Data shall be transferred at near real time. • **Solution**: The bridge written in Java was written in Phyton.

Table 17:Pilot challenges for Digital Twin Representation for NOKSEL pilot

Cloud platform	• **Challenge**: Data losses due to OPC Server's (KepServerEX) being turned off by the operators. • **Requirement**: OPC server shall be up and running with high availability. • **Solution:** OPC server is moved to virtual servers that are on the same network of the PLC.
Data Lake, storage	• **Challenge**: Bottlenecks in Apache Kafka • **Requirement**: Data shall be stored permanently, and there should not be bottlenecks in Apache Kafka**.** • **Solution**: Kafka's topics, partitions and consumers' configuration were rearranged. PostgreSQL and Cassandra have been used together as Kafka's consumers.

Digital Twin – Data driven representation	• **Challenge**: Store data in database in a way to fulfill the storage and analysis requirements. • **Requirement**: Data shall be stored in databases and be used for analytics and real time monitoring. • **Solution**: Cassandra tables were rearranged to store component-based data per-table.
Real time event handling, CEP	• **Challenge**: None • **Requirement**: None • **Solution**: None
Cyber Security	• **Challenge:** Security • **Requirement:** Data should be accessed by authorized and authenticated users, data should be coded. • **Solution:** Authentication and authorization was performed. Data encryption via coding was done.

Table 18:Pilot challenges for Digital Hybrid and Cognitive Digital Twin Generation for NOKSEL pilot

Analytics Models	• **Challenge:** Data cleaning and labelling • **Requirement:** Pre-processing for ML/DL for predictive maintenance • **Solution:** Industrial Big Data Processing implementation • **Challenge:** The number of features is high • **Requirement:** Efficient ML models generation • **Solution:** PCA on data • **Challenge:** Difficulty in analysis due to high volume of data • **Requirement:** Efficiency and performance • **Solution:** Pandas Profiling is used
Physical Models	• **Challenge**: Missing models • **Requirement**: 1st order model generation • **Solution**: Models are generated in Matlab Simulink
Machine Learning	• **Challenge**: Algorithms did not learn very well on the collected data set • **Requirement**: Quality ML models trained • **Solution**: A thorough analysis was conducted and feature selection is applied • **Challenge**: Missing data • **Requirement**: ML/DL model training for predictive maintenance • **Solution**: Synthetic data generation and data balanced sampling • **Challenge**: Low quality data • **Requirement**: ML/DL model training at high correctness • **Solution**: Under sampling applied on to the data
Cognitive Digital Twins	• **Challenge**: Ensure that the predictive maintenance also takes into account operators domain knowledge and experiences

	• **Requirement**: Combine the Digital Twin based recommendations with the inclusion of operators domain knowledge for self learning and being proactive • **Solution**: Extract the tacit knowledge from experts as a basis for developing the cognition. Partially by applying unusuality detection on past data • **Challenge**: Fast query of semantics (ontology) • **Requirement**: Ontology should be queried fast and stored • **Solution**: Ontology model development and storage in a relational database using Protege and On2RDB respectively

Table 19:Pilot challenges for Digital Twin Visualisation and Control for NOKSEL pilot

2D/3D visualisation	• **Challenge**: Slowness in big data visualization • **Requirement**: Latency < 200ms • **Solution**: View creations for zoom in and zoom outs • **Challenge**: Custom 3D visualizations • **Requirement**: none • **Solution**: User preferred visual elements adjustments (i.e. colour, light, etc.)
Control	• **Challenge**: Controlling ambient temperature of the welding room which is a closed environment • **Requirement**: To act proactively and prevent machine failures due to increased temperature in the welding room • **Solution**: A control panel has been designed, developed by TEKNOPAR and installed at NOKSEL site to control air conditioners in the welding cell

Pilot specific aim

The objectives of the NOKSEL case on the SWP machinery are threefold: 1) real-time condition monitoring, 2) predictive maintenance, and 3) digital twin generation. The aim of the pilot case is to improve the predictive maintenance capabilities and thus increase the total equipment usage performance by analysing operational and automation data received from different sensors with digital twin supported condition monitoring platform to be developed in serial production of steel pipes.

Steel pipe sector is a sector where operation run on 24/7 basis. The cost of machines breakdown is very high. Due to the multi-step nature of the process (Figure 52) if a section stops due to a malfunction, the entire production is stopped. An efficient Predictive Maintenance approach has the potential to increase the machine uptime. Hereby, equipment availability, performance and quality have been increased. It also supports reduction in maintenance times, maintenance and operational costs and operator risks. Flexibility, agility, profitability and competitive advantage in production has been provided.

Innovation

The main innovation is the development of a Digital Twin for the SWP in steel pipe production. The digital twin collects and analyses multiple sensors' data in real-time and enable a smart condition monitoring system for predictive maintenance. Real-time data acquisition, communication networks for monitoring, and automated recommendation generation are among the key innovative features of this pilot.

Digital Twins can serve as the basis for advanced analytics and AI applications. Advanced analytics and AI applications can be part of a Digital Twin, making it an intelligent and self-contained entity. Abburu et.al. defines three types of digital twins: 1) digital twins, 2) hybrid digital twins, and 3) cognitive digital twins. A hybrid twin is a digital twin, and a cognitive twin is a hybrid twin (Figure 51). A digital twin is a physical replica of a physical system. A hybrid twin using different set of data (such as sensor data, databases, simulators, etc.), contains a set of interconnected models, where a cognitive twin is proactive and learn by itself. The "cognitive element" in the Digital Twins are introduced by learning from historical process data and events to predict unwanted events in the operation before these events happen (Albayrak and Ünal, 2021).

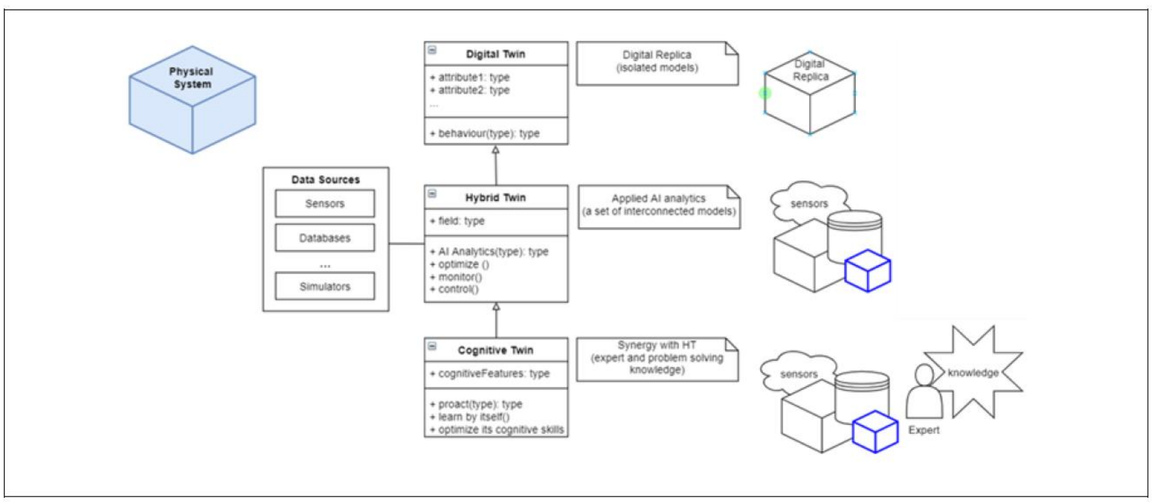


Figure 51:Hybrid and Cognitive Digital Twins are both Digital Twins (Albayrak and Unal, 2021)

Smart components that use sensors to gather data about real-time status, working condition, or position were connected to a cloud-based system that receives and processes all the data the sensors monitor. This input has been analysed against business and other contextual data through smart visualization systems. The digital twin model allows joining physical and virtual worlds to create a new networked layer in which intelligent objects interact with each other to virtualize the steel pipe manufacturing process on the SWP machinery.

IoT platform and architecture in use

Currently the SWP Machinery is being tracked by NOBİS System (a Delphi based special software created by NOKSEL and machinery adjustments are made by terminals with TEKNOPAR's installed software. NOBIS is also integrated with the SAP system.

The digitalisation architecture is closely related to the processes of the SWP machine. At NOKSEL, involved components are; coil feeding station, pinch roll, skelp leveller, skelp and welding station, edge milling machine, main drive unit, forming station, internal welding and external welding. Figure 52 displays the current processes of SWP realized by the SWP machine components at NOKSEL:

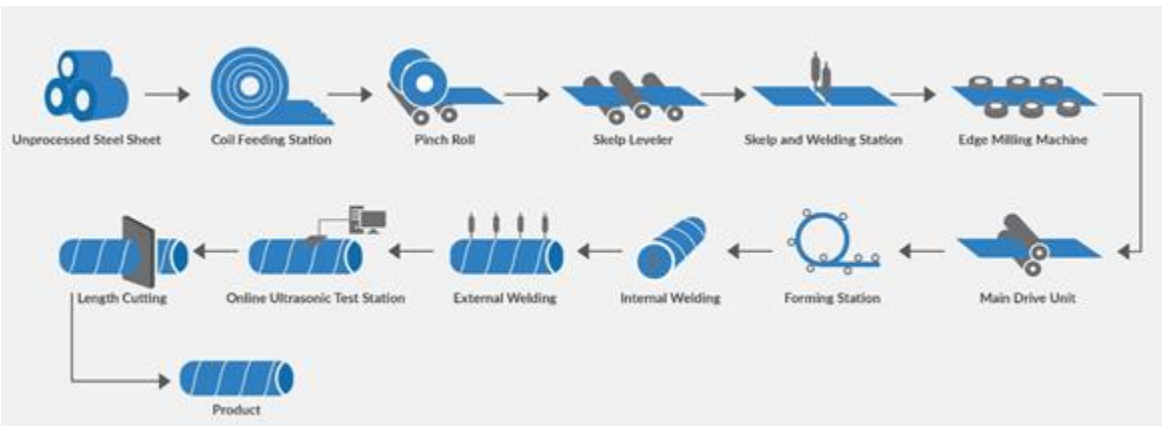


Figure 52: Steps in the SWP process at NOKSEL.

The existing architecture at NOKSEL's pilot facility is referred as AS-IS architecture and the planned future architecture is referred as TO-BE architecture. Figure 53 presents the AS-IS architecture that was valid in the beginning of the project. The TO-BE architecture has been completed in 2022 2022 when the Cognitive Digital Twin of Noksel's use-case had been completed in the COGNITWIN project.

The three-tier architecture pattern comprises the edge, platform and enterprise tier that handles the data and control flows. In the developed platform, the edge tier collects data from the edge nodes within industrial automation system. The platform tier receives, processes and forwards control commands from the enterprise tier to the edge tier. The enterprise tier has domain-specific applications.

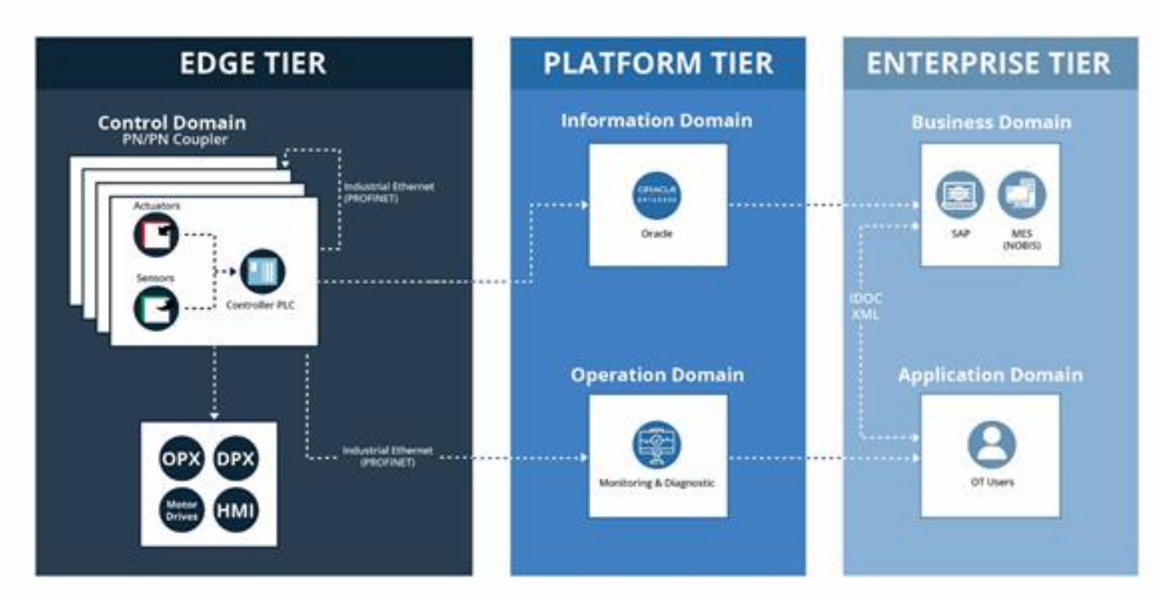


Figure 53 : AS-IS: Existing Architecture at NOKSEL at the beginning of the COGNITWIN project

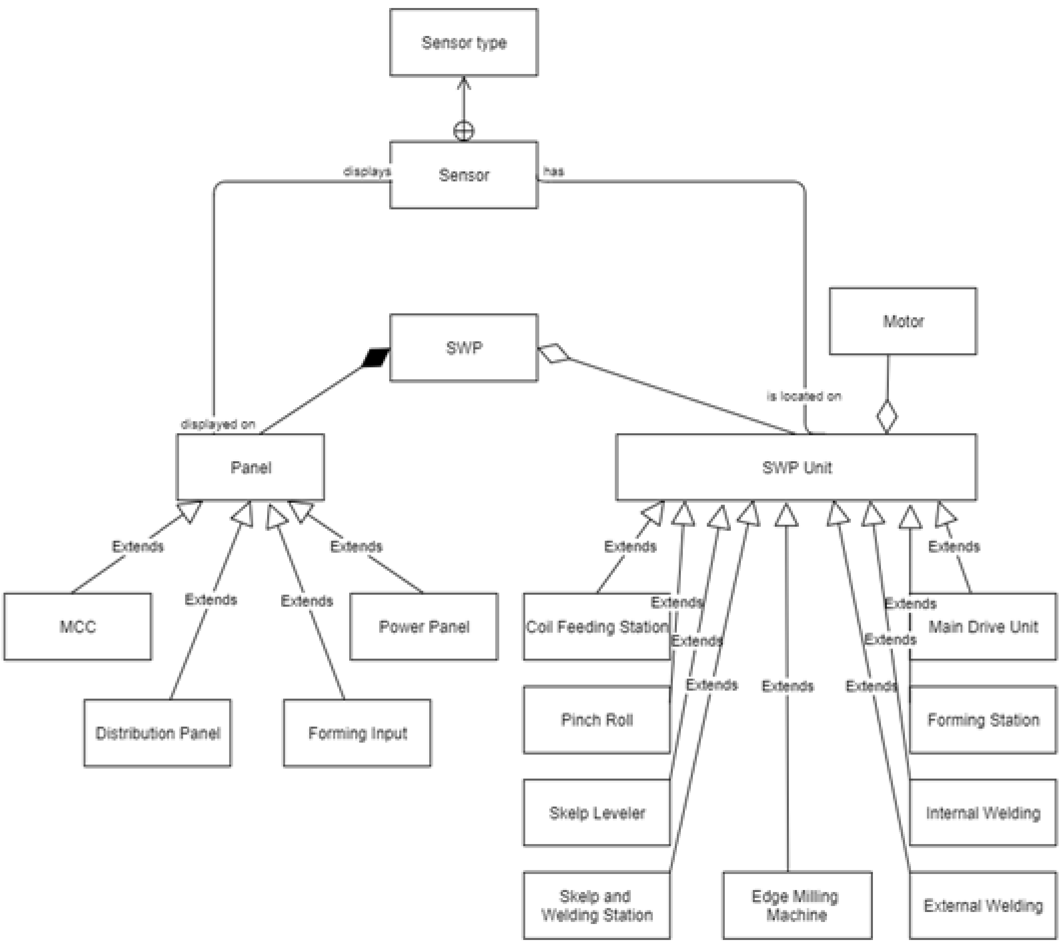


Figure 54 : System's Generic Static View of the Architecture at NOKSEL

SWP is associated panels and is composed of different SWP units. On these units, there are limit sensors located. These sensors are displayed on the existing panels, which are MCC, distribution panel, forming input panel and power panel. In a UML class diagram, Figure 54 presents a top-level generic architecture of the SWP at NOKSEL from a static point of view.

Currently, there is an IoT platform where the machine is monitored with SCADA systems on the shop floor. The initial existing sets of sensors are composed of limit sensors:

- Direction limit sensors of carrier cars (forward and backward limits)
- Right/Left roll handler limit sensors (in out limits)
- Pre delivery up location limit sensors
- DP1-DP5 limit sensor list

Note: We have the complete list of existing sensors and their specifications at NOKSEL. We do NOT present the list and sensor specs in this document for confidentiality reasons.

At NOKSEL, three types of motors (AC, DC and servo motors) exist (Figure 55). The motors are associated with the drivers that are specific to the associated motors.

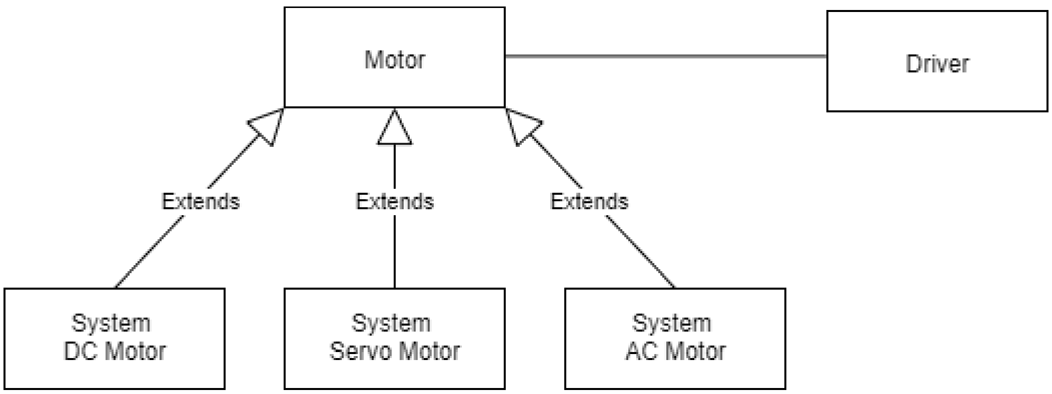


Figure 55 : Engine Types at NOKSEL

Panels at NOKSEL are classified into two main groups; Distribution panels and Operator panels. There are five distribution panels at NOKSEL's existing IIoT platform (Figure 56).

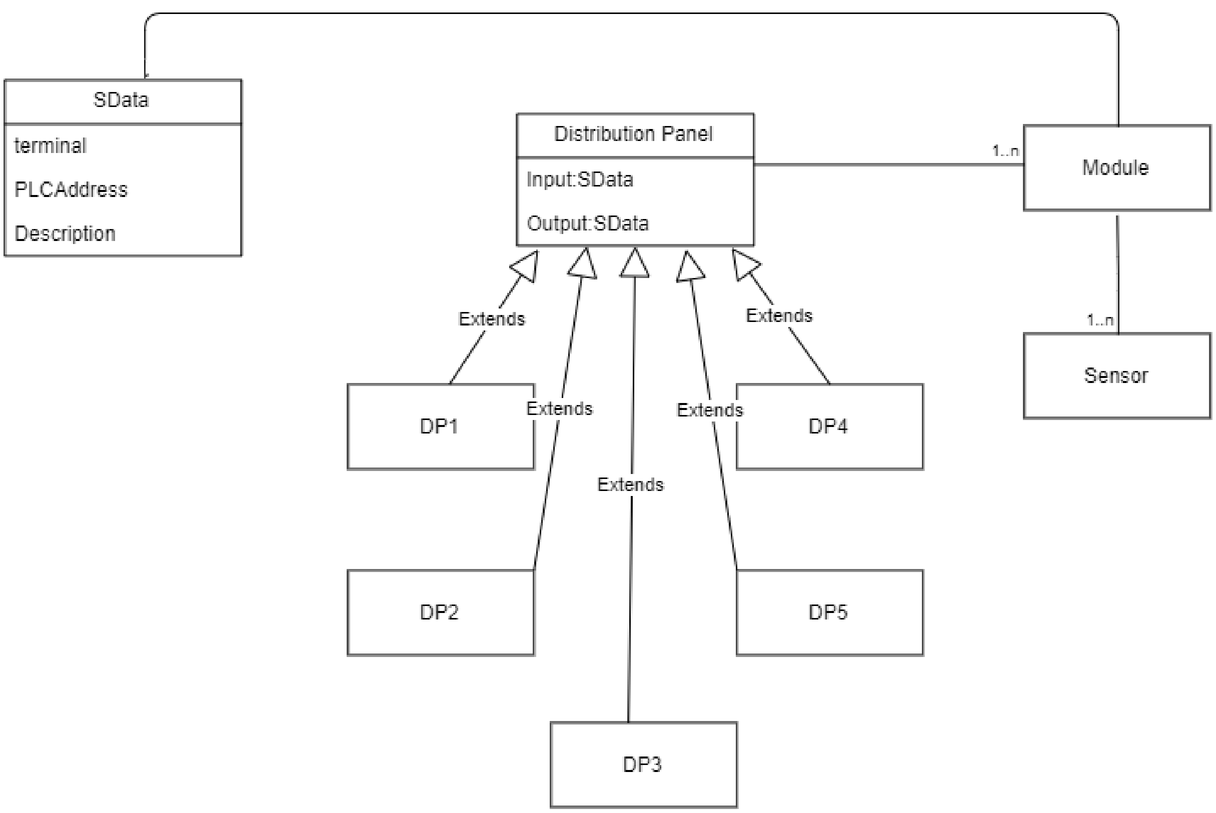


Figure 56 : Existing Digitalization at NOKSEL: Distribution Panels

Each distribution panel (DP) is associated to sensors and switches via modules. Table 20 displays the DP's and their associated types of sensors and switches at NOKSEL's existing IIoT platform.

Panel	Type of System Sensor/Switch	Number of Sensor/Switch
DP1	Proximity End. Sensor	19
DP1	Pressure Switch	4
DP2	Proximity End. Sensor	15
DP2	Pressure Switch	1
DP3	Proximity End. Sensor	15
DP3	Laser Photosensor	6
DP5	Proximity Sensor	9

Table 20: AS-IS: Existing Digitalisation at NOKSEL: System Sensor/Switches and Distribution Panels.

At NOKSEL, there are ten operator panels, each of which is associated to different units of SWP. The units that are associated with the existing sensors are listed in Table 21.

Name	Unit	Number of Modules	Number of Buttons/Keys
OP1	Coil Feeding Station	2	48
OP2	Pinch Roll	3	64
OP3	Skelp Leveller	3	66
OP4	Edge Milling Machine	2	48
OP5-6	Main Drive Unit, Forming, Internal Welding	4	80
OP7	Length Cutting	2	48
OP9K	Forming by Internal Welding and External Welding	4	36
OP10K	Internal Welding, External Welding	3	56
MOP	Main Operator Panel	4	96
KMOP	Welding Main Panel	5	80

Table 21: AS-IS: Existing Digitalisation at NOKSEL: Operator Panels.

Database and Datasets for Digital Twin Pilot - Noksel

Data to be used include sensor data, PLC data. Energy consumption data and ERP (NOBİS) data have also been utilized. The IoT system components and tools used are given in detail in template tables filled for D4.1. Briefly, current IoT system has its components like PLC hardware and software, analogue/ digital modules, communication modules and SCADA system installed in the production plant. Establishing the necessary infrastructure to communicate with the intermediate module software (OPC server) and streaming from PLC to cloud through sensor network are under development.

A fully asynchronous communication structure with the event-bus method is used for the transmission of data collected from the source with OPC. Data transmission is provided in the JSON format. In the architecture managed on the basis of Microservice, Cassandra is used as the NoSQL database with a database presented as a log file to users.

Cassandra is a database that provides continuous availability, high performance, and scalability. PostgreSQL, a relational database (RDBMS), is used by the interface program that provides user interaction to display time series data in real time.

Dataset stored in the databases include sensor data collected from the installed sensors, and alarm and status data that are retrieved from the PLCs via OPC. Id, time and value fields are stored in the database. A total of 120 sensor values are monitored on the SWP machine to capture data on temperature, vibration, pressure, current, oil temperature and contamination. A value is taken every 10 milliseconds from the vibration sensors, once every 100 milliseconds from the current sensors, and every 1000 milliseconds from the temperature and pressure sensors plus alarm and status fields. SWP machinery has a total of 120 sensor values, 122 alarm and 175 status data which create 11 GB of incoming data in one day (in 24 hours).

Description of Data available

Current tracking system provides downtimes periods and types, total working durations, effective working durations, number of produced pipes, meters of produced pipes, weight of produced pipes, which pipe produced of which labelled raw material which shows the quality of raw material.

Only daily electrical consumptions have been started to be recorded since November 2019. Before then, electrical consumptions have been recorded monthly in 2019.

Starting end of 2020, data from the new sensors installed, and from the added PLC were started to be collected. Sensor data, alarm, status and production process data were acquired. A custom ERP system's data, named NOBİS data has been integrated and used together with the sensor data for AI related calculations.

Since 2022 with COGNITWIN project in NOKSEL's use case; The digital twin established in our İskenderun facility collects real-time data from PLCs, more than 125 sensors, and uses the acquired data together with alarm and status data of the Spiral Welding Machine (SWP) used in steel pipe manufacturing. Our maintenance operators are now able to visualize multisensory data on specially designed dashboards to fulfil NOKSEL's requirements. The dashboards complement TEKNOPAR's previously developed SCADA and control systems.

The demonstrator below in Figure 57 shows how stream data collected real time can be used together with the artificial intelligence models to monitor the process, to determine anomalies in advance, and to predict machine breakdowns early enough to take precautions.

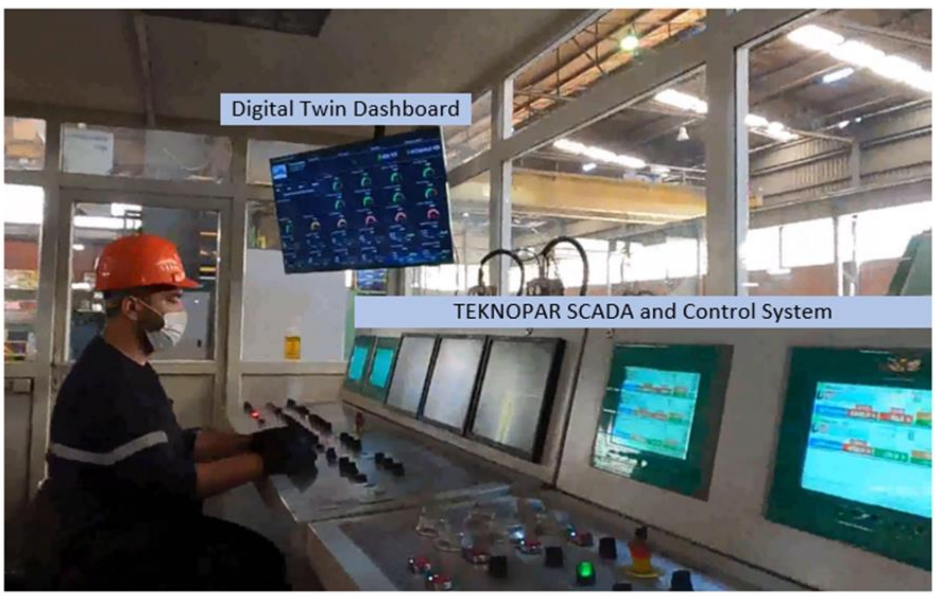


Figure 57:Cognitive Digital Twin System Control for Monitoring

- **Digital Platform for Digital Twin Pilot - Noksel**

Noksel Digital Platform – Overall architecture

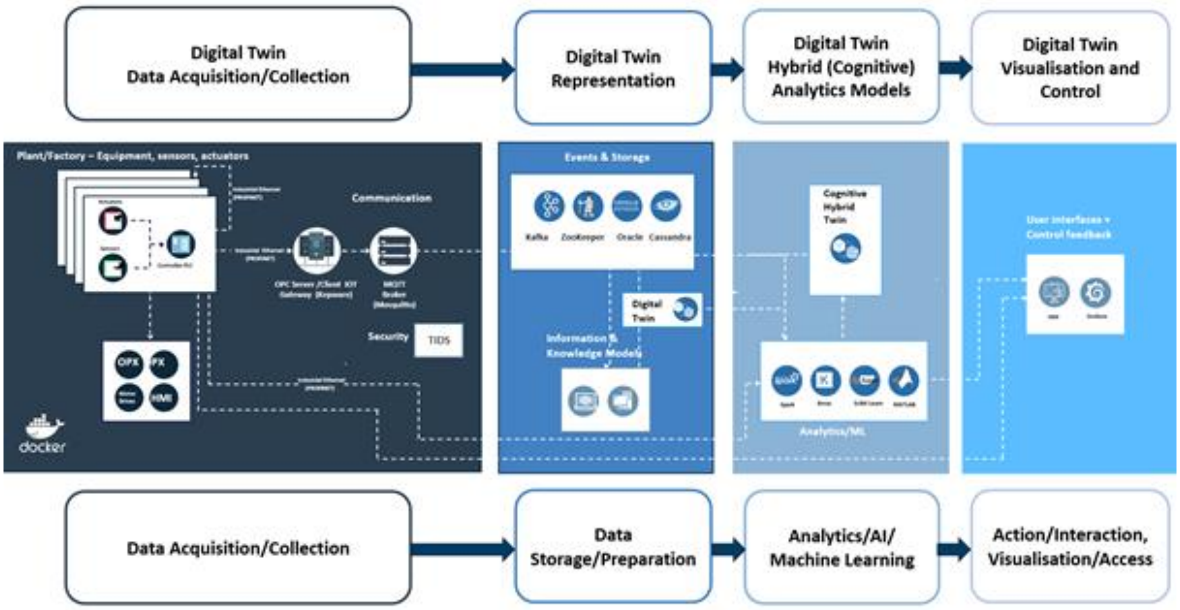


Figure 58: Updated Topology Aligned with Pipeline Architecture for Noksel Pilot ((Unal, Albayrak, Jomaa, & Berre, 2021))

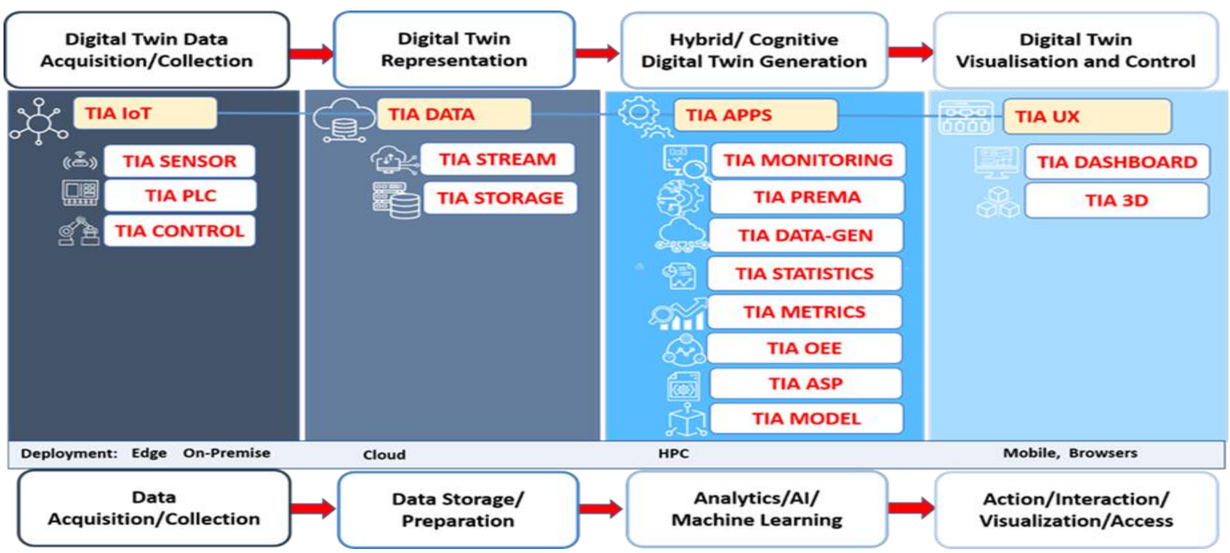


Figure 59:Pipeline Architecture for Noksel Pilot mapped to TIA PLATFORM tools

The four-tier architecture pattern that handles the data and control flows. In the developed platform, the first tier collects data from the edge nodes within industrial automation system. Second tier is performed DT representation. Third tier is related to DT Hybrid (Cognitive) Analytics Models. Fourth tier is associated with DT visualisation and control. The four-tier architecture of the developed platform is given in Figure 58,Figure 59.

The architecture can also be presented in the following Figure 60 as aligned with BDVA reference architecture.

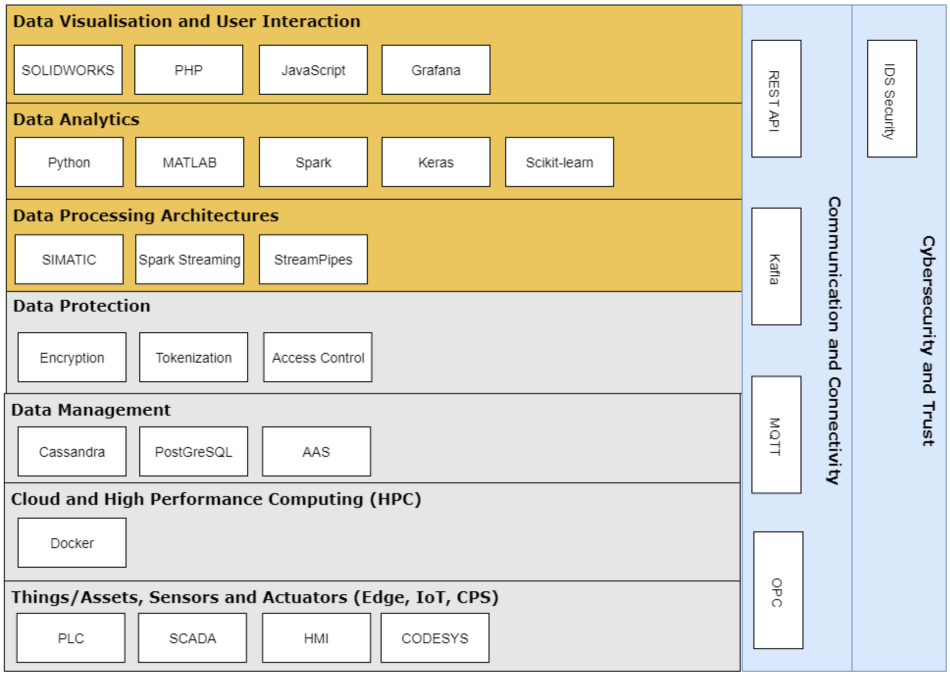


Figure 60:NOKSEL Pilot architecture presented as aligned to BDVA reference architecture

Open Platform Communications (OPC is the interoperability standard for the exchange of data in the industrial automation among devices from multiple suppliers. The OPC Unified Architecture refers to the service-oriented architecture. KEPServerEx is an OPC server that serves as a connectivity platform with device drivers, client drivers, and advanced plug-ins to fit the communication requirements of the industrial control systems. The IoT Gateway is a plug-in to allow data from PLCs and other devices to be delivered to third-party nodes and clients via HTTP and MQTT protocols. MQTT Broker transports data between the Edge gateway and other nodes in the platform.

Eclipse Mosquitto is an open-source broker that implements the MQTT protocol to carry out messages using a publish/subscribe model.

Kafka is an open-source distributed event streaming platform used for building data pipelines, data integration, streaming analytics, and scalable processing applications. Initially developed as a distributed messaging queue, Kafka is based on distributed commit logs which hold the ordered sequence of events. Kafka can read, write and process streams of events in a vast array of programming languages and has interfaces to connect to numerous event sources and sinks. Kafka is highly scalable and supports data persistency. It provides high throughput and low latency. Many commonly used programming languages are supported by Kafka.

Apache ZooKeeper is used for the management of consensus between producers and consumers and for synchronization purposes. Zookeeper coordinates the Kafka system for the cluster integrity, broker status, and coupling of producers and consumers. ZooKeeper also manages the failure and use of replications, and the authorization process and access control lists are stored in this service.

Cassandra is a NoSQL database that provides continuous availability, high performance, and scalability. Due to its high scalability, Cassandra can hold petabytes of data and perform thousands of transactions in a very short time. Apache Spark supports batch processing and stream processing by micro-batching as an in-memory batch data processing platform. Spark offers increased performance with its in-memory processing. Grafana is an open-source, general purpose dashboard and graph composer, which runs as a web application. It retrieves data from multiple different data sources and has ready visualization elements to display data. Grafana can easily be integrated with many services such as to send alarms and notifications as SMS, or as message etc. SOLIDWORKS is used for 3D visualization for the Digital Twin.

Benefits of reuse, interoperability, and flexibility are gained by using micro services.
Micro services are supported by Docker. Docker is a packaging and deployment methodology in order to easily manage the variety of the underlying hardware resources efficiently. Docker Swarm is used for clustering and scheduling in Docker containers by establishing and managing a cluster of Docker nodes as a single virtual system. Docker Compose helps create a stack of multiple simultaneous running Docker containers.

SIMATIC and Kafka are used in the Data Processing Architecture layer. In the Data Analytics layer the components are developed by using Python, MATLAB, and Spark.HTML, JavaScript, Grafana, and SOLIDWORKS are used in the Data Visualisation and User Interaction Layer. Cyber security and Trust are supported by the IDS Security component.

Noksel Digital Platform – Data Acquisition – including sensors

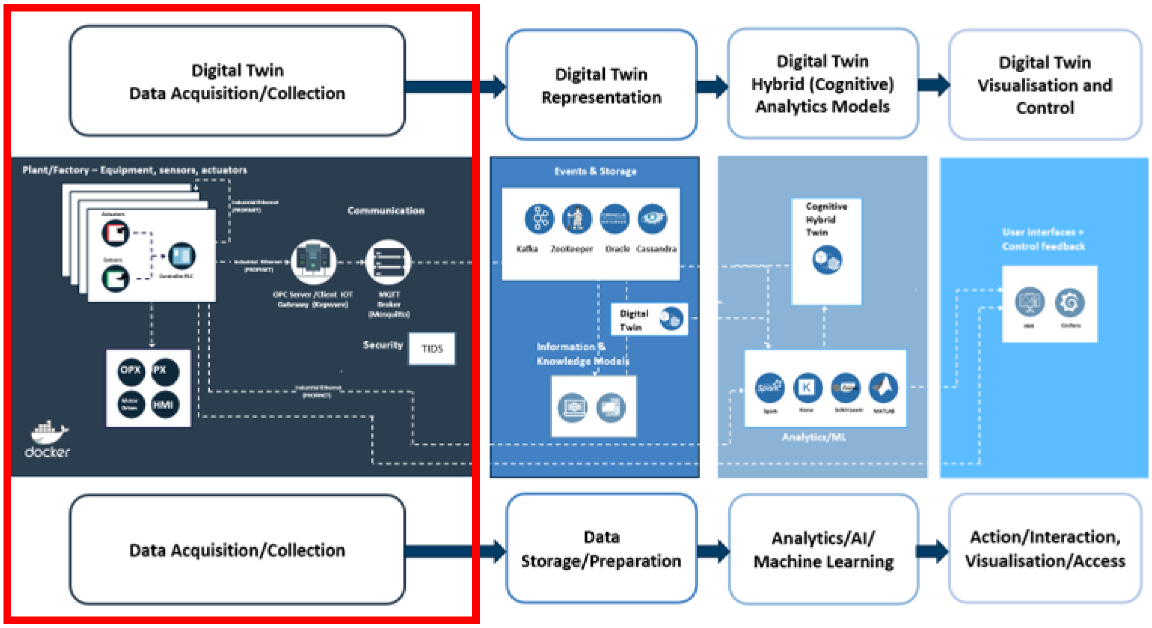


Figure 61: Pipeline Architecture for Noksel Pilot: DT Data Acquisition/Collection ((Unal, Albayrak, Jomaa, & Berre, 2021))

Figure 62 Figure 62 shows the hardware topology of the current system. With the developed digital twin-supported condition monitoring platform d, an infrastructure that aims to analyze the operational and automation data received from sensors and PLC/SCADA used for PM, which helps to increase the overall equipment performance.

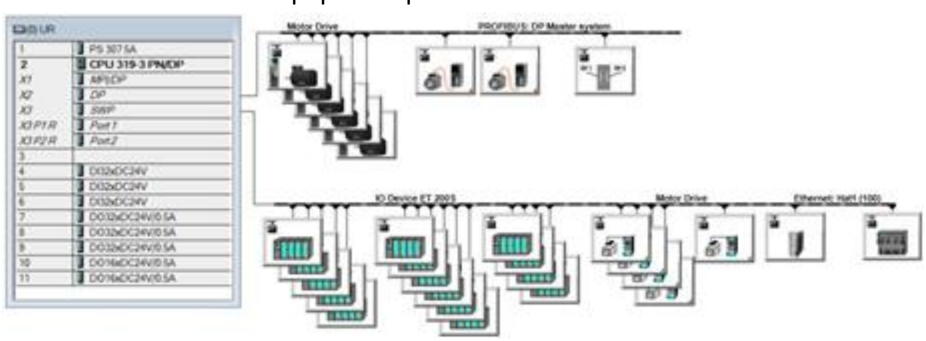


Figure 62: Existing hardware topology

Communication between these two topologies is provided with the industrial communication protocol PROFINET, and the two structures communicate with each other. Data required from the existing structure obtained by using the existing controller. Figure 63 presents the added hardware topology.

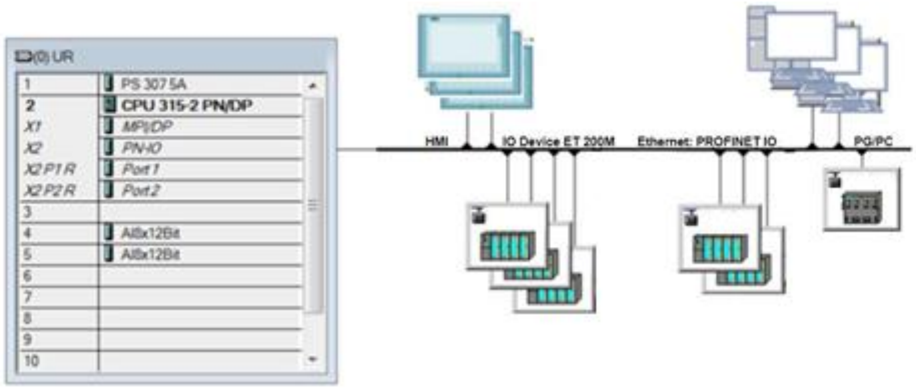


Figure 63: Added hardware topology

The former PLC model used for process control was S7 300. The operation details of the components, status information, process information, such as speed and power, production details, and system alarms are kept on this PLC while the newly added sensor data and alarms are located in the S7 1500 PLC. The existing PLC data was transferred to the S7 1500 PLC through the PN/PN Coupler module, allowing all data tracking to be carried out over the new PLC.

The PN/PN Coupler module provides the simple connection of two separate PROFINET networks. The PN/PN Coupler enables data transmission between two PROFINET controllers. The output data from one network becomes the input data of the other. For data transfer, additional function blocks are not required and the transfer is realized without latency. In order for the new sensors added to the system not to affect the existing process, a new PLC is employed and the controls are implemented over it. The communication structure between the PLCs is designed using the PN/PN Coupler module as shown in Figure 64.

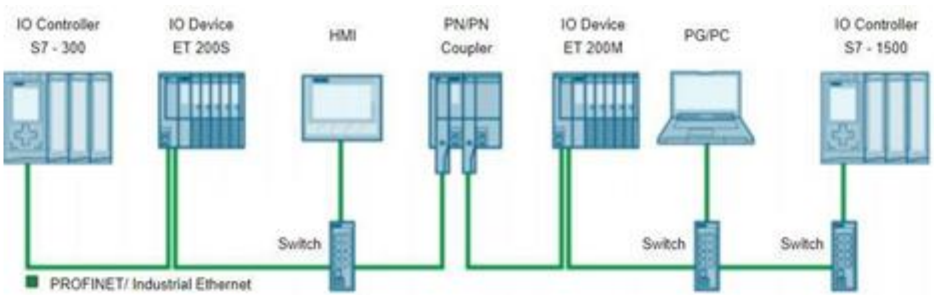


Figure 64: Coupling of the PROFINET subnets with the PN/PN Coupler

PLC transmits the data it receives from the sensors to OPC, which then transfers the data to the platform via MQTT. The received data is transmitted to Kafka, which passes it on to the Cassandra and PostgreSQL databases to be stored for further processing or later access. The data are stored in the Cassandra database with three columns: id, time, and value. The id column shows the component to which the data belong. The JSON format streams of the data transferred to the Cassandra database are presented to users as a log file.

A fully asynchronous communication structure with the event-bus method is used for the transmission of data collected from the source with OPC. Data transmission is provided in the JSON format. In the architecture managed on the basis of Microservice, Cassandra is used as the NoSQL database, and PostgreSQL, a relational database (RDBMS), is used by the interface program that provides user interaction.

SENSORS

Table 22:Sensor List

IFM	HYDAC	PHOENIX Contact	SIEMENS
Temperature and Vibration sensor	Temperature, Pressure and Contamination sensor	Current sensor	Energy Analyser

While selecting sensors for NOKSEL steel pipe production plant physical constraints of the environment and the pilot's requirements were taken into account. All of the sensors were installed at the sensor places determined in the SWP machine (Table 22Table 22:Sensor List). Regardless of sensor type, the selected sensors have been preferred by TEKNOPAR experts because of our long-time experiences in automation, hydraulic and pneumatic systems. Moreover, manufacturers of the motors, used in SWP machine, suggested these sensors.

IFM sensors provide convenience in harsh industrial environments where mounting on the engine is challenging. HYDAC's sensors have been chosen owing to that they provide convenience in spaces where mounting on the engine is difficult on the hydraulic and pneumatic systems that will affect the manufacturing process. HYDAC's was specifically developed for OEM (Original Equipment Manufacturer) applications. PHOENIX's sensors can be measured by placing sensors on the thick cables that feed the motor to measure the current value of the high ampere consumed. Another reason for this sensor selection is that the current value can be calculated thanks to the magnetic field formed in the cable. SIEMENS's products have been generally chosen for our projects because of their standards. That's why SIEMENS energy analyser has been used for measuring energy consumption.

The selected products have product characteristics that can be made precise measurements at the critical points desired to be followed on the machines.

CONTROL PANEL

Figure 65:Control panel developed for Noksel pilot

A control panel, in Figure 65, has been developed for the air conditioner so that commands can be sent according to the decision taken by the system. This panel transforms the update commands for the air conditioner status determined by the system into a structure that the air conditioner can understand and transmits it to the air conditioner. The panel has been used to control temperature at the welding cell at Noksel.

The above control panel designed, implemented and developed by TEKNOPAR for the Noksel pilot has been installed on the closed room for the welding machines, shown in Figure 66:

Figure 66:Welding Machines in a Closed Room at Noksel

Noksel Digital Platform – Data storage/preparation

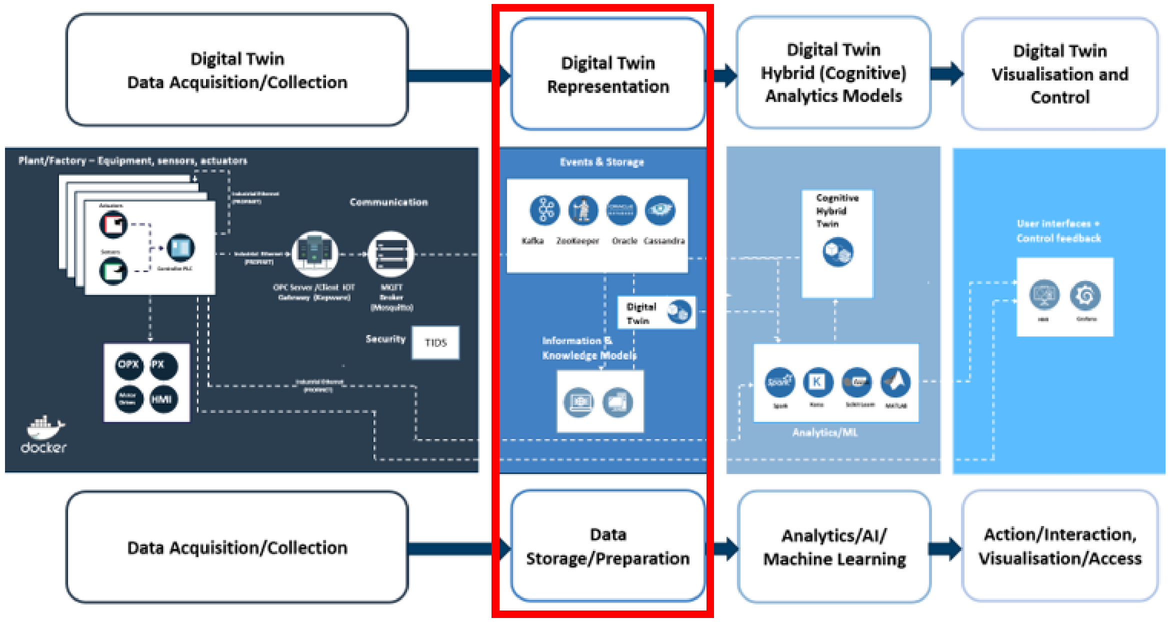


Figure 67: Pipeline Architecture for Noksel Pilot: DT Representation ((Unal, Albayrak, Jomaa, & Berre, 2021))

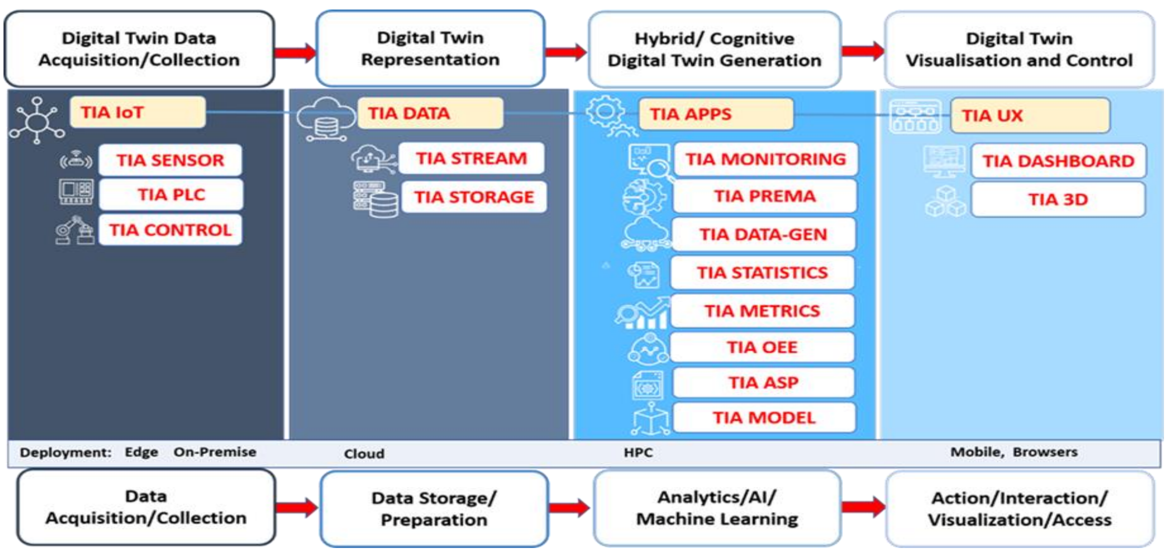


Figure 68:Pipeline Architecture for Noksel Pilot: DT Representation mapped to TIA PLATFORM elements

The data obtained from the sensors, such as temperature, pressure and vibration, voltage, and current are transmitted to MQTT over OPC, and then to Kafka in the JSON format. Apache Kafka is a data streaming platform developed specifically to transmit real-time data with a low error margin and short latency. Kafka achieves superior success in systems with multiple data sources, such as sensor data and reduces the inter-system load. It has an integration that can also process big data coming from sensors operating at high frequencies. Figure 67 displays the pipeline architecture for NOKSEL pilot Data Representation step.

Instant data received by Kafka is transmitted to the Python-based server, where the attribute extraction process begins. Incremental principal component analysis (PCA), which is the most well-known method used in big data flow, applies PCA stages to the instantaneous data using data in a certain window range, and thus large data that cannot fit into the memory can also be processed effectively. PCA basically performs dimensional reduction by making the incoming high-dimensional data low-dimensional, providing more accurate results for machine learning, and therefore it is frequently used for categorization problems.

Noksel Digital Platform – Analytics/AI/Machine Learning

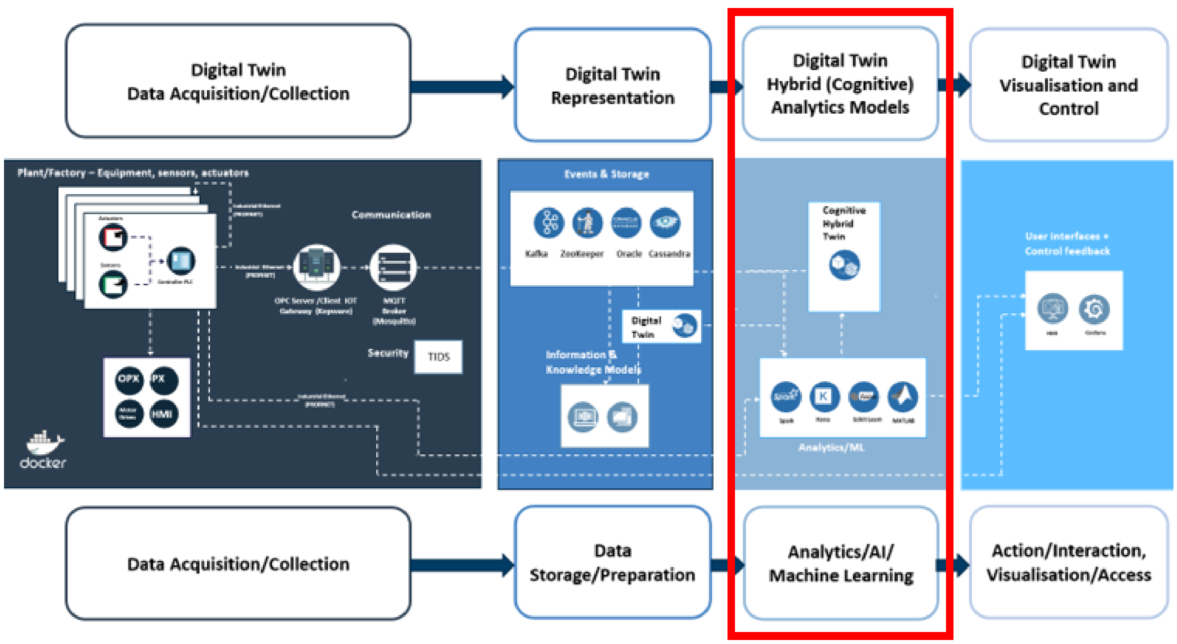


Figure 69: Pipeline Architecture for Noksel Pilot: DT Hybrid (Cognitive) Analytics Models ()

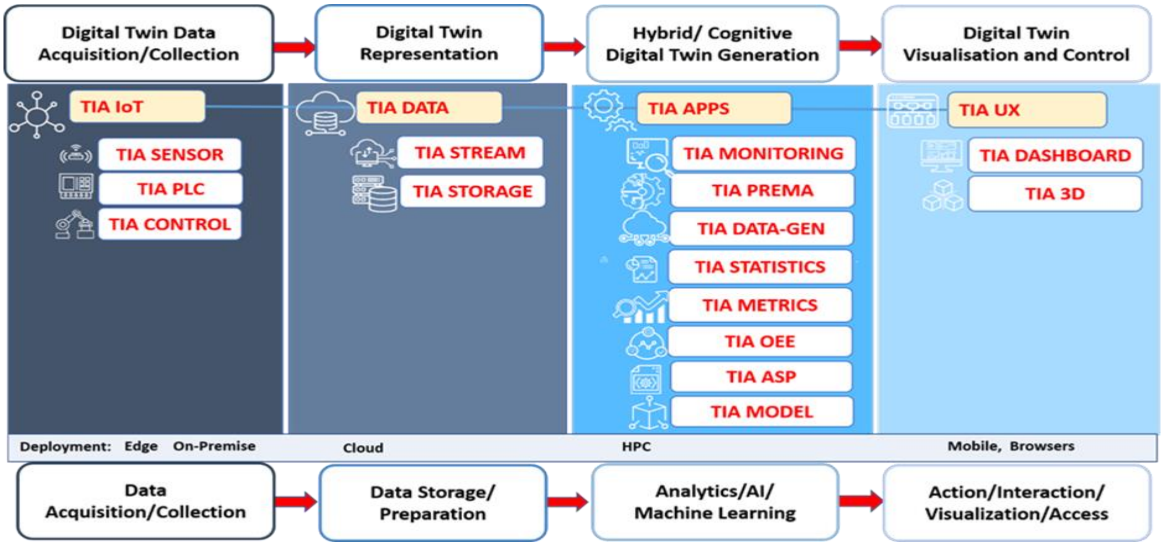


Figure 70:Pipeline Architecture for Noksel Pilot: DT Hybrid (Cognitive) Analytics Models mapped to TIA PLATFORM

In the MLL module, six different machine learning algorithms are applied to the data passing through the incremental PCA stage to detect anomalies. Prediction results are produced using data from three different machine learning libraries. First Spark MLlib is produced entirely by Spark, which uses Spark's engine optimized for large-scale data processing. In the pilot, the remaining useful life (RUL) of the SWP machine predicted. To predict RUL, we have collected sensor data, selected some ML models and trained the algorithms with different data sets. We also compared the performance of various ML and DL algorithms. The RF (Random Forest) and gradient boosted tree machine learning algorithms belong to this library. Keras library utilizes TensorFlow, and is used for deep learning. The LSTM algorithm of this library is utilized. This open- source neural network library makes it simpler to work with artificial neural networks through its user interface facilities and modular structure. The Scikit-Learn library is another open-source machine learning library that contains several algorithms for regression, classification, clustering. We used algorithms like SVM, KNN and multi-layer perceptron (MLP) from Scikit-Learn library. In addition, auto encoder, generative adversarial networks (GANs), deep belief networks, and K-means algorithms are considered being used.

An MLL application is developed for comparing the machine learning models used for predictive analysis. The application enables users to select the machine learning model for a given set of data, and then compares the output using graphical elements.

Noksel Digital Platform – Action/Interaction-Control-Visualisation

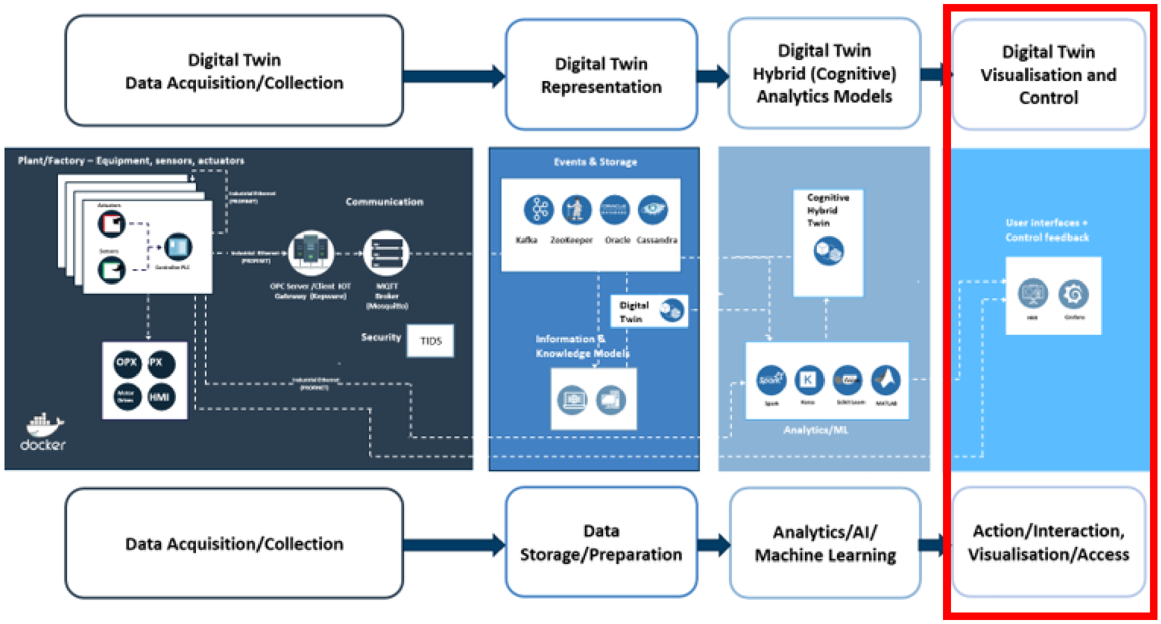


Figure 71: Pipeline Architecture for Noksel Pilot: DT Visualisation and Control ((Unal, Albayrak, Jomaa, & Berre, 2021))

This component contains designing dashboards suitable for sensor data, error detection, and transfer of regular information obtained from data processing to the real-time status monitoring system, and development of end-user (mobile/desktop/web) applications. Selected GUIs for the component is presented in Figure 73. Figure 71 displays the pipeline architecture for NOKSEL pilot Visualization and Control step.

Three.js, an open-source JavaScript library, was used to develop animated or non-animated 3D applications that can be opened in the web browser using WebGL. Three.js is supported by all WebGL supported web browsers. In addition to Three.js for visualization of the Digital Twin elements.

For the web interface, the JSON data received with JavaScript have been parsed and then transferred to PHP pages. In this communication, the post method has been used in the requests sent with JavaScript. With the help of PHP, the information was placed in HTML objects. Grafana technology is used in the process of placing graphics within the card object. Dynamic graphics created on Grafana are placed on cards in iframe tags.

The generated output includes predictive analytics results.

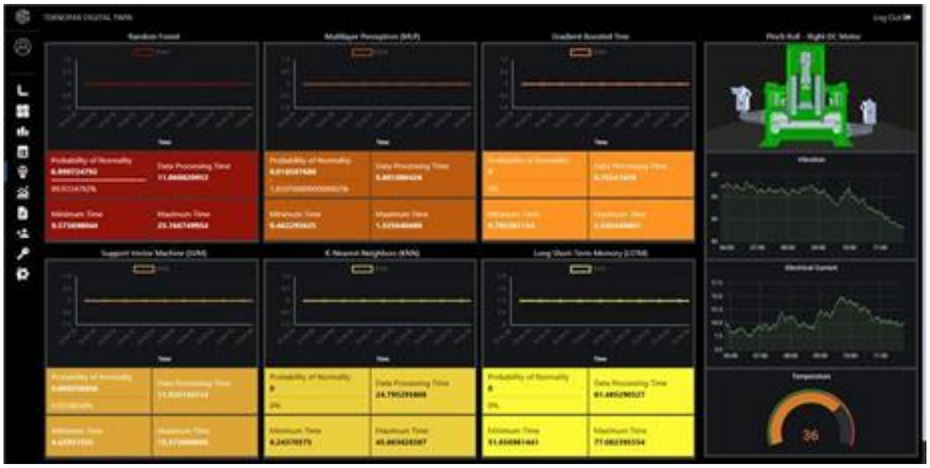


Figure 72: Artificial Intelligence/ Machine Learning/ NN Algorithms Application GUI

Figure 73: TEKNOPAR's platform visualization and digital twin GUIs

New displays have been generated for different applications including but not limited to OEE, energy consumption, and breakdowns.

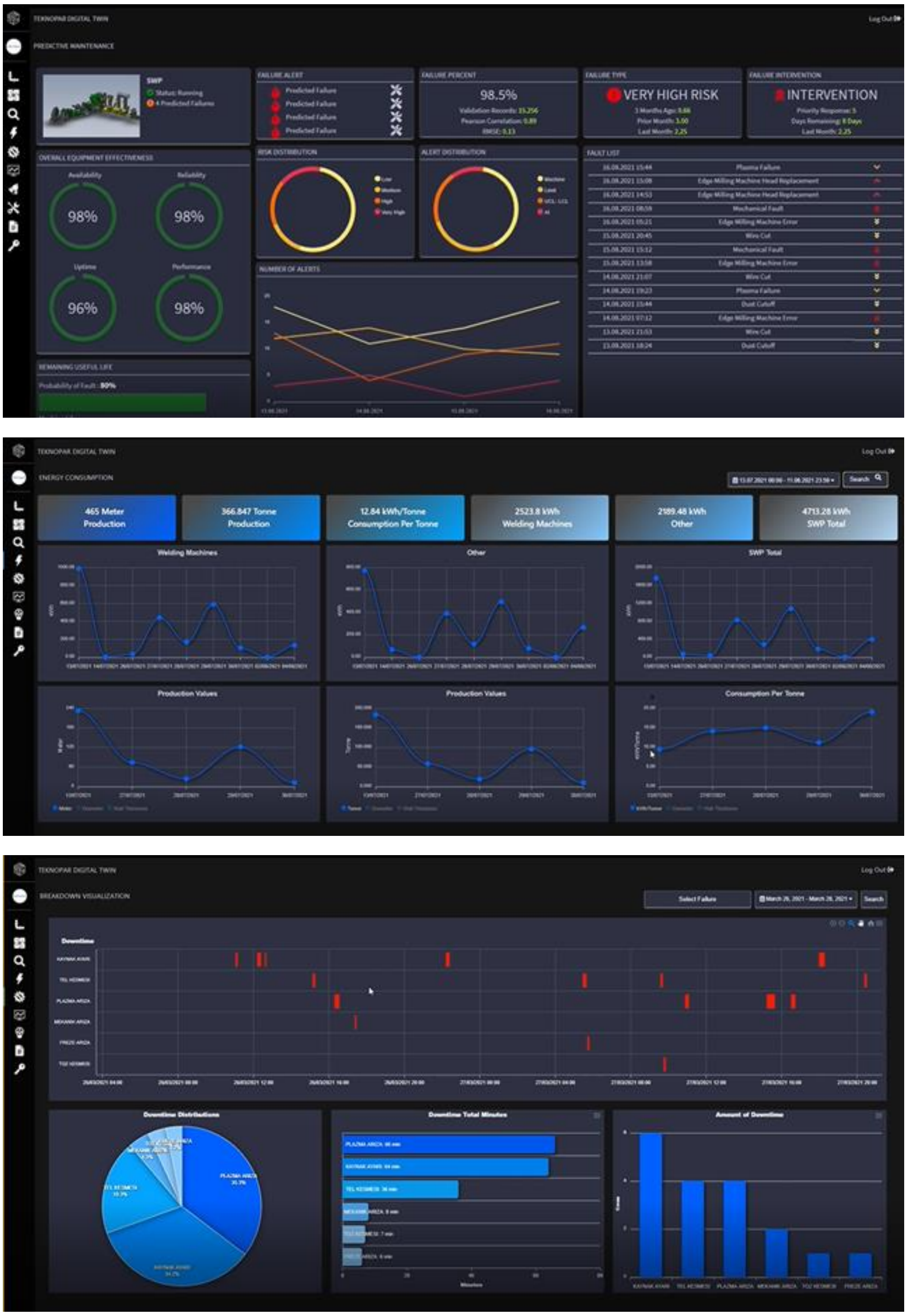


Figure 74:Selected new displays generated for the Noksel pilot

StreamPipes and AAS Studies Validation

In the Noksel pilot, some of the tools developed in the COGNITWIN project have been validated and tested by TEKNOPAR. Regarding StreamPipes, following three pipelines have been created and shown to be working (Figure 75):

1. A Pipeline to retrieve data in CSV format and after processing it to visualize it graphically
2. A Pipeline to enable selection of ML algorithm to be executed on stream data and to be visually presented
3. A streamline to transmit data over MQTT to Kafka broker.

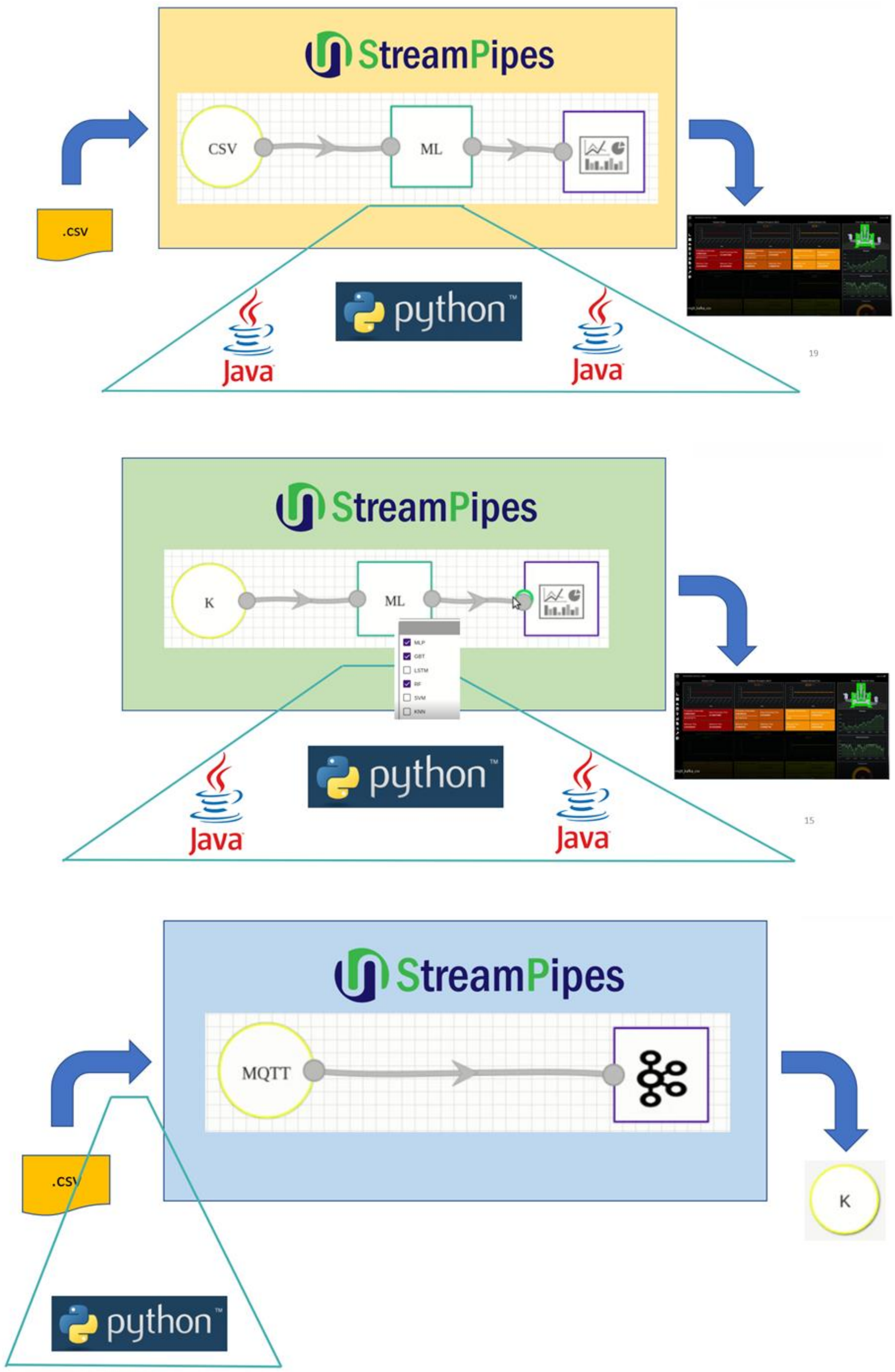


Figure 75:Apache StreamPipes pipelines demonstrated at the NOKSEL pilot

Asset Administration Shell (AAS) appears to be the key concept of Platform Industrie 4.0 in order to enable interoperability. The AAS can directly be adopted to implement Digital Twins. As a result, all industries may benefit an open and standardized metal model, standardized data models with homogenized semantics and standardized APIs and infrastructure services. Regarding AAS studies, Package Manager and AASxServer were used and two different pipelines have been developed: Pipeline 1) A pipeline to stream data display, and Pipeline 2) A pipeline to preprocess stream data. Both of the demonstrations were successfully completed (Figure 76)

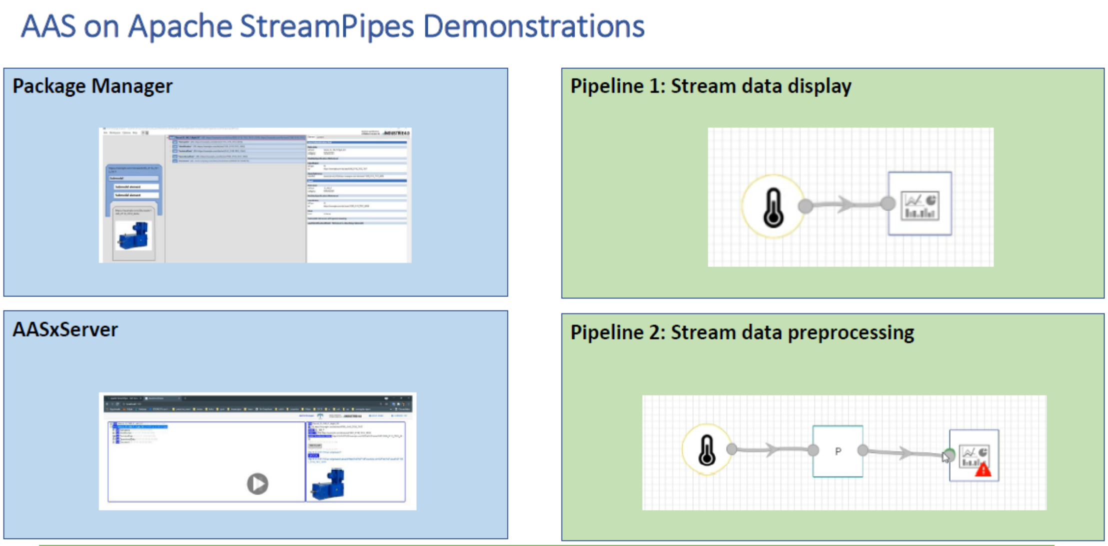


Figure 76:Pipelines used for AAS validation at NOKSEL pilot

TEKNOPAR has designed and implemented different technical solutions using emerging digital twin technologies/tool in order to generate AAS and IDS in a real manufacturing environment. The Admin Shell IO, Eclipse BaSyx, IDS connectors have been developed and used together with Apache StreamPipes and Node-RED. The developments have been verified and validated by testers (Yallıç at al 2022).

Demonstrator of Digital Twin Pilot – Noksel

The demonstration of the Digital Twin pilot of Noksel is presented through the 5 steps as listed below:

- Pilot summary
- Digital Twin
- Real Time predictive maintenance
- MATLAB Models
- Streampipes

Measurable KPIs and Final impact

The SWP consists of various parts for pipe preparation, welding, cutting, etc. all of which are controlled by sensitive servo systems. Due to the process's multiple steps, if one stops working due to a breakdown, the entire production is stopped. For the performance monitoring of the SWP machine, sensors, and energy analyzers have been installed. Production and failure data were obtained by integrating NOBIS, NOKSEL's MES solution. In addition, domain knowledge was obtained from experts through monthly meetings and field visits. It has been determined that the welding machine components are the most important components of the SWP machine, with the downtime information obtained from the experts and NOBIS. With the analyzed data, it has been determined that wire-cutting failure is the most common cause of unplanned downtime. In order to increase the performance of the SWP machine, it is aimed to reduce the unplanned downtimes encountered and to reduce the energy consumption by reducing the downtimes.

COGNITWIN was implemented to the pilot plant in two stages as follows:

- “Before” implementation: November 2019 – March 2022
- “After” implementation: April 2022 – October 2022

In the first phase of COGNITWIN, the final targets for selected KPIs were defined. The improvements in the selected KPIs have been calculated by comparing the measured key performance indicators (KPIs) after the implementation of the COGNITWIN to the measured performance prior to the implementation. The measurable KPIs and their target improvements were set to:

1. Real-time online monitoring of process efficiency with minimal latency < 100 ms

The data transfer time is aimed to be <100 ms in order to instantaneously transfer ML/DL algorithm prediction results and to be monitored in real-time by the end user so that the downtime can be intervened quickly. In the tests carried out in this context, the difference between the time that the data is transmitted and the time it is viewed by the end-user has been measured, and a value of <1 ms has been reached.

2. Accuracy of data analytics, and AI algorithms > 95%

Preprocessing steps (correction/cleaning/merging etc.) were applied to the data obtained from the sensor, energy analyzer, and NOBIS system in order to be used in ML/DL algorithms. The trained model was implemented to predict the type of unplanned wire cut breakdown with the real-time data flow. When the model metrics are calculated, the breakdown is confirmed by comparing it with the NOBIS system, which it correctly predicts 45-60 minutes ago. The model accuracy was calculated with the test data obtained after the model is measured as 99%.

3. No errors in anomaly detection (both Type I and Type II)

For the anomaly detection of the system Type I Error and Type II Error analyses were conducted on the streaming data. The FP and FN values obtained from the generated confusion matrix correspond to

Type I error and Type II error values, respectively. While 37 of the 548 predictions are in Type I (FP), 12 of them are in Type II (FN). This means that out of 548 predictions made by the model and 37 of them were predicted to have breakdown, but no breakdowns occurred (FP). Also, although it was predicted that no breakdown would occur in 12 predictions, the malfunction occurred (FN). According to these numbers, Type I errors make up 6.25% of the total predictions, and Type II errors make up 2.18%. When more data is collected for all production parameters and sensors, the Type I and Type II targets can be approached.

4. Reduction in machine downtime due to conducted predictive maintenance by 10%

The SWP machine works in the production line and the production process consists of many different processing steps. Therefore, in case of unplanned downtime in any of the process steps, the whole production is breakdown. With the prediction of the breakdowns, the disruption in production can be eliminated and the performance of the SWP machine in production can be increased. When the data obtained from the NOBIS system is examined, wire cutting stands out among the other preventable breakdown types. With the predictive maintenance developed in the COGNITWIN project, the predictions created by the ML/DL model are transmitted to the end user before the machine breakdown occurs. Before the DT implementation, a total of 15 wire-cut breakdowns were encountered during ~1200 tons of pipe production (1.25 % per ton). After the commissioning of the model, while more errors were expected during ~5400 tons of pipe production, only 12 wire cut failures were encountered (0,22 % per ton). Thanks to the successful prediction of wire-cut breakdown, the number of unplanned machine downtimes has been reduced and an 82% reduction in unplanned wire-cutting breakdown number has been achieved based on the per ton breakdown analysis. Thanks to the correctly predicted wire-cutting breakdowns by the trained model, the energy and time consumed at the time of the machine breakdown are saved. Before the digital twin integration, the machine could not operate for 2.7% of the operating time due to breakdowns, after the digital twin integration, this rate was reduced to 1% (% 62 reduction in breakdown time).

5. Reduction in energy consumption by 10%

Energy consumption data has been started to be recorded daily since November 2019 at the NOKSEL facility. In order to monitor energy consumption, energy analyzers are placed in the electrical panel of the welding machines and the SWP electrical panel. 30 different data are collected from these energy analyzers. Before November 2019, energy consumption data were recorded monthly.

After the digital twin was integrated, between April 2022 and October 2022, 839 minutes of extra production time was added to the production process by reducing machine downtimes. Approximately 200 tons of extra production is made in ~839 minutes, which corresponds to ~3.6% of the total production time. Considering energy consumption, in the default case, when ~5281 tons of production was made, 12,345 w of energy per ton would be considered to have been consumed. By preventing machine downtimes, ~5481 tons of production was made in the same period and 11,747 w of energy is consumed per ton. As can be seen, while it is expected to consume 12,345 w per ton, the energy consumed per ton is reduced to 11,747 w per ton, thus, 4.84% of energy is saved in energy consumption per ton.

Table 23:NOKSEL Pilot KPIs

NOKSEL – Digital Twin Powered Condition Monitoring (and Control) in Steel Manufacturing Industry		Target	Achieved
KPIs	Real-time online monitoring of process efficiency with minimal latency	< 100 ms	< 1 ms
	Accuracy of data analytics, and AI algorithms	> 95%	99%
	Percentage of Type 1 error in anomaly detection (incorrect rejection of a true null hypothesis)	= 0%	6,75%
	Percentage of Type 2 error in anomaly detection (failure to reject a false null hypothesis)	= 0%	2,18%
	Reduction in machine downtime due to conducted predictive maintenance	10%	62%
	Reduction in energy consumption	10%	4,84%

In the Noksel pilot, following benefits are gained by the COGNITWIN project:

1. Life cycle optimisation of Spiral Welded Machine (SWP) in steel pipe production, where CT of the SWP monitors the condition and health of the machinery, offers early warnings, and suggests optimised predictive maintenance plans for the machinery based on real-time data gathered from sensors such as the pressure, temperature, vibration, etc., and alarm and status information.
2. Improving operational performance of the production process by predicting and identifying the optimal operating parameters based on both historical practices and real-time process and thus improving the overall productivity of the plant.
3. Improving energy consumption efficiency by monitoring and predicting the energy analyser and operational parameters based on both historical practices and real-time process.
4. Enhanced utilisation of computing infrastructure with virtual machines and containerisation technologies to achieve optimised RAM and CPU usage.
5. Minimise health & safety risks and maximise the human operator performance by early warning of machine and system problems.
6. Real-time monitoring of parameters like pipe diameter, pitch angle, belt width, production speed, pipe diameter and wall thickness for semi-finished and finished steel products for ensuring operational efficiency and stabilising the production process.

Noksel pilot associated work and results have been published and/or presented in multiple studies to create further impact. These studies include:

- Albayrak, Ö., & Unal, P. (2021). Digitalization of a Steel Pipe Production Factory: STEEL 4.0- A Family of Products Developed on Routes from Industry 3.0 to Industry 4.0. The Fifth International Iron and Steel Symposium, Data Science in Process Engineering, (pp. 271-274).

- Albayrak, Ö., & Unal, P. (2021). Smart Steel Pipe Production Plant via Cognitive Digital Twins: A Case Study on Digitalization of Spiral Welded Pipe Machinery. Advances in Intelligent Systems and Computing, (pp. 132-143).
- Unal, P., Albayrak, Ö., Jomaa, M., & Berre, A. (2021). Data-driven artificial intelligence and predictive analytics for the maintenance of industrial machinery with hybrid and cognitive digital twins. In Technologies and Applications for Big Data Value. Springer.
- Yallıç, F., Albayrak, Ö., & Ünal, P. (2022). Asset Administration Shell Generation and Usage for Digital Twins: A Case Study for Non-Destructive Testing., (pp. 299-306). Malta.
- Temel, S., Ummak, E., Tokgöz, A., Işık, F., Albayrak, Ö., Ünal, P., & Özbayoğlu, M. (2022). Control System Design and Implementation Based on Big Data and Ontology. Osaka, Japan.

Conclusion and Summary

The initial set of challenges that have been addressed in the previous term includes platform establishment, data modelling and data acquisition. By selecting the optimum place, set of sensors, and deciding the platform architecture these challenges were addressed and required data set was started to be collected. Collected data did not include as much failure data as needed by the machine learning/deep learning algorithms to be trained and tested at the targeted levels. To address this challenge, physical models of the system elements were developed, and executed to generate synthetic data. There are at least three success stories associated with the COGNITWIN project:

1. Anomaly detection: High temperature is detected to be one of the data that results in machine failure. The data measured by the system presented an anomaly in temperature and warned the operators, who disagreed that there was no anomaly and the heat was normal. When external tools are used to present that the heat was actually high as suggested by the digital twin, the operators took required actions and their trust in the digital twin increased resulting in increased level of technology acceptance.

2. Energy consumption: The energy consumption is important for the SWP machine. The data analytics has shown that the energy consumption was positively corelated with the hydraulic power unit pressure filter value. As a result, it was suggested to change the filter, and the energy consumption was decreased.

3. Preventive maintenance: The analysis performed on data collected presented that one of the main causes for the machine failure is about wire-cut, which has been found to be directly associated with the high temperature level in the room where the weld generators have been located. A control system has been developed in order to control the temperature in the room. For the system related ontology has been developed, stored, queried and a hardware control component has been designed and developed. The control system manipulated the temperature parameter by controlling the air conditioner, as a result preventive maintenance at the pilot site has been supported.

The challenges to be addressed are related to data modelling, optimization and deep learning related activities have been cope with. Expert knowledge of the operators and the relevant users of the system have been developed using ontology models. The serialization (persistence) of the developed ontologies have been realized. Following ontology development for fast semantic query performance developed ontology has been stored in a relational database. It was found that one of the major factors of the welding machines breakdown is the abnormal increase in the environmental heat degree in the welding machine's generator's room which was formerly being controlled manually. New system which automatically control the room's degree and whenever reach a critical level system automatically starts to cool down the environment was installed (Figure 77). To control one of the main causes of machine breakdowns, a control panel has been designed, implemented and installed. The hybrid twin that has been generated by M18 has been enhanced, and by adding required cognitive elements to the hybrid twin a cognitive, self-learning digital twin has been accomplished at NOKSEL. The developed cognitive digital twin controlled the temperature and hence prevented unplanned machine failures in a proactive manner. It was experienced that the Cognitive Digital System is alert the abnormalities in SWP system just 15 minutes before its breakdown so that the system enables not to halt to SWP production system.

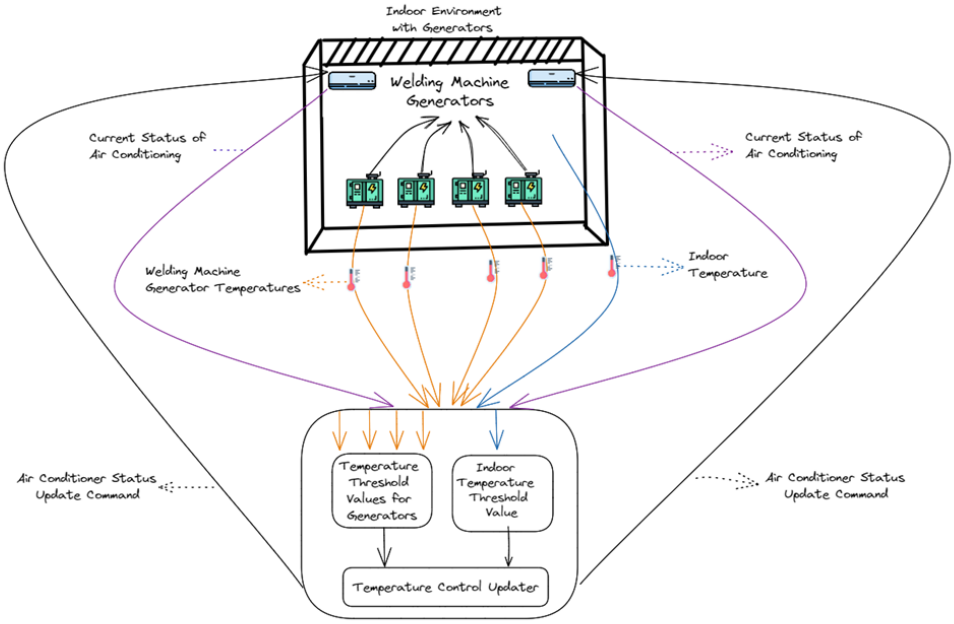


Figure 77: Incoming/Outgoing Data Loop of Indoor Temperature Control System (Ref: Temel et. al., IEEE Big Data 2022)

The final Noksel pilot demonstrator is described in the COGNITWIN Toolbox [1] with the Noksel pilot digital twin pipeline description [8] and the final Noksel pilot demonstrator videos [9]. This is also further described in the final public deliverable D6.4 Best "Digital Twins" practices report [2].

5 Summary

SAG has focused on billet tracking in the Nauweiler rolling mill and identified lacking sensors and required actions. To ensure the correct and automatized identification of billets upon entrance into the mill train, an improved billet identification system was installed. The more demanding task of providing a seamless tracking of billets in the blooming train was addressed in close collaboration with technology partners by means of a Computer Vision tracking system. To this end, 3 FullHD cameras were installed overlooking the entire blooming train. A delay in the deployment occurred because of a delay in part of the development, handled through an internal project at SAG completing the operational deployment of the technologies after the end of the COGNITWIN project.

In SIDENOR the focus has been on refractory management for replacement in steel ladles. The relevant process data and type of available measurement of the ladle profiles were also established considering the requirements of the rest of the partners for the tasks to be developed. The definition of the data requirements for a data pipeline has been provided to the technological partners. With a basis in data collection cognitive digital twin prediction models were created and realised and, after a deployment delay at the end of the project, Sidenorhas installed a cognitive and hybrid digital to support the future refractory wear decision making.

NOKSEL has had a focus on the development of a cognitive digital twin of Steel Pipe Welding machinery. The current system architecture was studied. The data needed to observe and evaluate contributions and impact of the new system was collected and analyzed. The current sensors of the SWP machine were listed. The desired new sensor installations have been integrated into the current plant's system. Data from sensors and PLCs were collected and stored in databases to use in machine learning. Developed Toolbox interoperability has been in use. At the end, the predictive maintenance system supported by the cognitive digital twin for SWP system was built up and is in use.

The final Steel pilot demonstrators are described in the COGNITWIN Toolbox [1] with the pilot digital twin pipeline description [4, 6, 8] and the final pilot demonstrator videos [5, 7, 9] . This is also further described in the final public deliverable D6.4 Best "Digital Twins" practices report [2].

6 References

Bibliographic References

Albayrak, Ö. U. (2021). Digitalization of a Steel Pipe Production Factory: STEEL 4.0- A Family of Products Developped on Routes from Industry 3.0 to Industry 4.0. *The Fifth International Iron and Steel Symposium, Data Science in Process Engineering*, pp. 271-274.

Albayrak, Ö. U. (2021). Smart Steel Pipe Production Plant via Cognitive Digital Twins: A Case Study on Digitalization of Spiral Welded Pipe. *Advances in Intelligemy Systems and Computing*, pp. 132-143.

Faltings, U., Bettinger, T., Barth, S., & Schäfer, M. (2022). Impact on Inference Model Performance for ML Tasks Using Real-Life Training Data and Synthetic Training Data from GANs. *information, 13*(9).

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., . . . Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems, 27*.

HYDAC Brand Sensors. (n.d.). Retrieved February 25, 2021from http://www.hydac.com.tr/allproducts.html . (n.d.).

IFM Brand Sensors. (n.d.). Retrieved February 25, 2021from https://www.ifm.com/tr/en. (n.d.).

Isola, P., Zhu, J., Zhou, T., & Efros, A. (2018). Image-to-Image Translation with Conditional Adversarial Networks. *arxiv*(arXiv:1611.07004).

Mirza, M., & Osindero, S. (2014). Conditional Generative Adversarial Nets. *arxiv*(arXiv:1411.1784.).

PHOENIX CONTACT Brand Sensors. (n.d.). Retrieved February 25, 2021 from https://www.phoenixcontact.com/online/portal/tr?1dmy&urile=wcm%3apath%3a/trtr/web/main/products/entry_page/entry_page . (n.d.).

Preventable machine failures are costing UK manufacturers up to half a million every year. (2018, February 11). Retrieved February 25, 2021, from https://www.oneserve.co.uk/blog/preventable-machine-failures-costing-uk-manufacturers-half-million-every-year. (n.d.).

Temel, S. U. (2022). *Control System Design and Implementation Based on Big Data and Ontology.* Osaka, Japan.

Unal, P., Albayrak, Ö., Jomaa, M., & Berre, A. (2021). Data-driven artificial intelligence and predictive analytics for the maintenance of industrial machinery with hybrid and cognitive digital twins. In *Technologies and Applications for Big Data Value.* Springer.

Yallıç, F. A. (n.d.). Asset Administration Shell Generation and Usage for Digital Twins: A Case Study for Non-Destructive Testing. Malta.

COGNITWIN References

[1] COGNITWIN Toolbox: https://cognitwin.github.io/toolbox/

[2] COGNITWIN Deliverable D6.4 Best "Digital Twin" practices report (Public)

[3] COGNITWIN Deliverable D2.3 "Hybrid models with cognitive elements for the Steel pilots" (Confidential)

[4] Saarstahl Digital Twin Toolbox pipeline: https://cognitwin.github.io/toolbox/pipelines/saarstahl.html

[5] Saarstahl Demonstrator: (2) SAARSTAHL Pilot (Final COGNITWIN Demonstrator - D2.4) - YouTube

[6] Sidenor Digital Twin Toolbox pipeline: https://cognitwin.github.io/toolbox/pipelines/sidenor.html

[7] Sidenor Demonstrator: (2) Sidenor Pilot (Final COGNITWIN Demonstrator – D2.4) - YouTube

[8] Noksel Digital Twin Toolbox pipeline: https://cognitwin.github.io/toolbox/pipelines/noksel.html
[9] Noksel Demonstrator: (2) NOKSEL Pilot (Final COGNITWIN Demonstrator - D2.4) - YouTube