

## Regular Article

Stein Tore Johansen\*, Perin Unal, Özlem Albayrak, Enso Ikonen, Kasper J. Linnestad, Sudi Jawahery, Akhilesh K. Srivastava and Bjørn Tore Løvfall

# Hybrid and cognitive digital twins for the process industry

<https://doi.org/10.1515/eng-2022-0418>

received August 17, 2022; accepted February 19, 2023

**Abstract:** In a Europe that is undergoing digital transformation, the COGNITWIN project is contributing to accelerate the transformation and introduce Industry 4.0 to the European process industries. The opportunities here can be illustrated by the SPIRE 2050 Vision document ([https://www.spire2030.eu/sites/default/files/users/user85/Vision\\_Document\\_V6\\_Pages\\_Online\\_0.pdf](https://www.spire2030.eu/sites/default/files/users/user85/Vision_Document_V6_Pages_Online_0.pdf)), which states that “Digitalisation of process industries has a tremendous potential to dramatically accelerate change in resource management, process control and in the design and the deployment of disruptive new business models.” The process industries are characterized with harsh environments where sensors are either costly, not available, or may be subject to costly maintenance. The development of digital twins that can exploit the combinations of data-based and physics-based models is often found to be a preferred path to robust digital twins that can help cutting costs and reduce energy consumption. In this article, we present 5 out of 6 industrial pilots that are developed in the COGNITWIN project. We discuss the commonalities and differences between the selected approaches and give some ideas about how cognition can be incorporated into the digital twins. The aim of this article is to inspire similar projects in related industries.

**Keywords:** cognitive plants, digital twins, cognition, industry 4.0, process industry

\* **Corresponding author: Stein Tore Johansen**, Department of Process Technology, SINTEF Industry, S. P. Andersens Veg 15, N-7494, Trondheim, Norway, e-mail: [stein.t.johansen@sintef.no](mailto:stein.t.johansen@sintef.no)

**Perin Unal, Özlem Albayrak:** Teknopar, Ankara, Turkey

**Enso Ikonen:** University of Oulu, Oulu, Finland

**Kasper J. Linnestad:** Cybernetica, Oslo, Norway

**Sudi Jawahery:** Cybernetica, Trondheim, Norway

**Akhilesh K. Srivastava:** Department of Process Technology, SINTEF Industry, Porsgrunn, Norway

**Bjørn Tore Løvfall:** Department of Process Technology, SINTEF Industry, Trondheim, Norway

## 1 Motivation: Needs of COGNITWIN

The European Green Deal [1] has the ambition to transform Europe into the first climate neutral continent by 2050. This requires development of a sustainable EU economy, and significant transformation of both production and consumption practices.

One of the key cornerstones of this goal relies on the successful transformation of the process industries. These industries have a strong impact on energy and material consumption [2]. In addition, these industries are an integral part of the value chains of the European economy. These heavy industries do realize that digitalization of their processes can offer great advantages. However, their challenges are often quite different from what is seen in the manufacturing sector, where advances on the digitization of their processes have advanced fast. The process industry sector is different, often dealing with very high temperatures, limited or no access to the core of the process, chemically harsh environments and where appropriate sensors are often very expensive to purchase, have limited lifetime and require short maintenance intervals. Currently, the sensor needs are defined to meet traditional process control procedures, and the process industry is lagging behind other industries in digitalization level. Therefore, the further digitalization in the process industry represents an unexploited potential and may enable acceleration of the transformation to more sustainable future industries while supporting [3] digital plant operations, intelligent material and equipment monitoring, and autonomous integrated supply chain management.

The Horizon 2020 project COGNITWIN<sup>1</sup> is aiming at answering these challenges and demonstrate how the process industry can be lifted through new sensors and application of the best available technologies in model development, combining data-based and physics-based

<sup>1</sup> [www.cognitwin.eu](http://www.cognitwin.eu).

approaches. In addition to developing hybrid digital twins [4], the project will introduce cognitive concepts into the digital twins.

In Section 2 of this article, we present the COGNITWIN industrial pilots and the approaches applied to develop hybrid and cognitive digital twins. Different approaches to the cognitive plants and digital twins are represented. Each approach is based on the industrial challenges in the different pilots, and the selection of the best methods to support the pilots, within the timeframe of the project, is crucial. Another important element is to develop methods and tools which may be used to support future industrial developments. Section 3 discusses the similarities and differences between the approaches applied in the different COGNITWIN pilots. Finally, the conclusions and next steps are given in Section 4.

## 1.1 COGNITWIN

The primary goal of the COGNITWIN project is to support European process industries and exploit the advantages of digitalization. Digital twins<sup>2</sup> (DTs) are key elements in achieving this goal. Referring to the Wikipedia definition, “A **digital twin** is a virtual representation that serves as the real-time digital counterpart of a physical object or process. Digital twins are the result of continual improvement in the creation of product design and engineering activities.” When a DT has been realized, this will have numerous applications, such as better process understanding, better understanding of the sensor data, data collection and processing, supporting optimization of the overall process as well as subprocesses. In the COGNITWIN project, the ambition is to further add cognition<sup>3</sup> to the digital twins. DTs may take many shapes and involve a huge forest of different techniques and methods, such as reduced order models [5,6], assimilation methods [7], recurrent CFD [8], and pragmatism in industrial modeling [9]. As discussed in [10] and [4], the COGNITWIN project has introduced three different layers of the DTs. We follow the IIC definition of DTs [11] and assume that a digital twin is a formal digital representation of an asset that captures attributes and behaviors of that asset. In

another valuable contribution to defining DTs [12], it is discussed how a model differs from a DT.

DTs may be developed from data (data-driven), from physics-based models or from a combination of the two approaches (one type of hybrid approach). In process control, this type of hybrid DT has been extensively used since the 1960s [13]. Lately, nonlinear model tuning to data has been successfully applied by means of extended Kalman filters [14].

The DTs may be built to handle a limited universe of data and models and where the interaction with generic and up till now unknown data and models are not foreseen. In other cases, it is known a priori that the specific DT may be a system of multiple subsystems, and we may want to profit in the future from new models and data becoming available. To deal with such cases, several approaches to DT asset management have been developed [15–18]. Generic access to data or models may be realized through semantic interoperability methods [19–21].

In the COGNITWIN project, a Toolbox [22] is being developed for use in the project but also for use in coming projects with similar types of process industries. Based on the requirements from the pilots, we develop DTs and apply our present Toolbox [22] together with an overall orchestration pipeline (Digital Twin Pipeline, Figure 1).

It must be noted that the majority of COGNITWIN pilots have challenges that deviate significantly from the manufacturing industries, mostly due to high temperatures ( $T > 1,600^{\circ}\text{C}$ ), fumes, dust, operations with heavy equipment and lack of affordable and robust sensors. In many cases the measurement challenges are formidable and there are no sensors in the development pipeline. Available measurement data may have inconsistencies due to on-the-fly operational changes. The data were not saved with the thought of being the source for future machine learning and DT building. Therefore, we have already learned that it is crucial to have physics-based models at hand to assess the data and make critical corrections. When data violate basic principles of mass, momentum, and energy conservation, we may use the physics-based model not only to validate the data but also to correct by model-based virtual data. The latter point may be important when the amount of data is scarce.

Even with very accurate DTs in operation, human intervention is still required to respond to rare, sudden, unknown events. A cognitive DT extends a hybrid DT by marrying expert knowledge with the power of hybrid analytics (Figure 2). The synergy with expert knowledge makes it possible to find solutions to previously unforeseen situations. In some cases, the outcome from the DT

<sup>2</sup> [https://en.wikipedia.org/wiki/Digital\\_twin](https://en.wikipedia.org/wiki/Digital_twin).

<sup>3</sup> Cognition elements, such as applying machine learning techniques to generate hybrid, self-learning and proactive systems or exploiting the interaction between model and humans, where the models support the decisions that are made by humans.

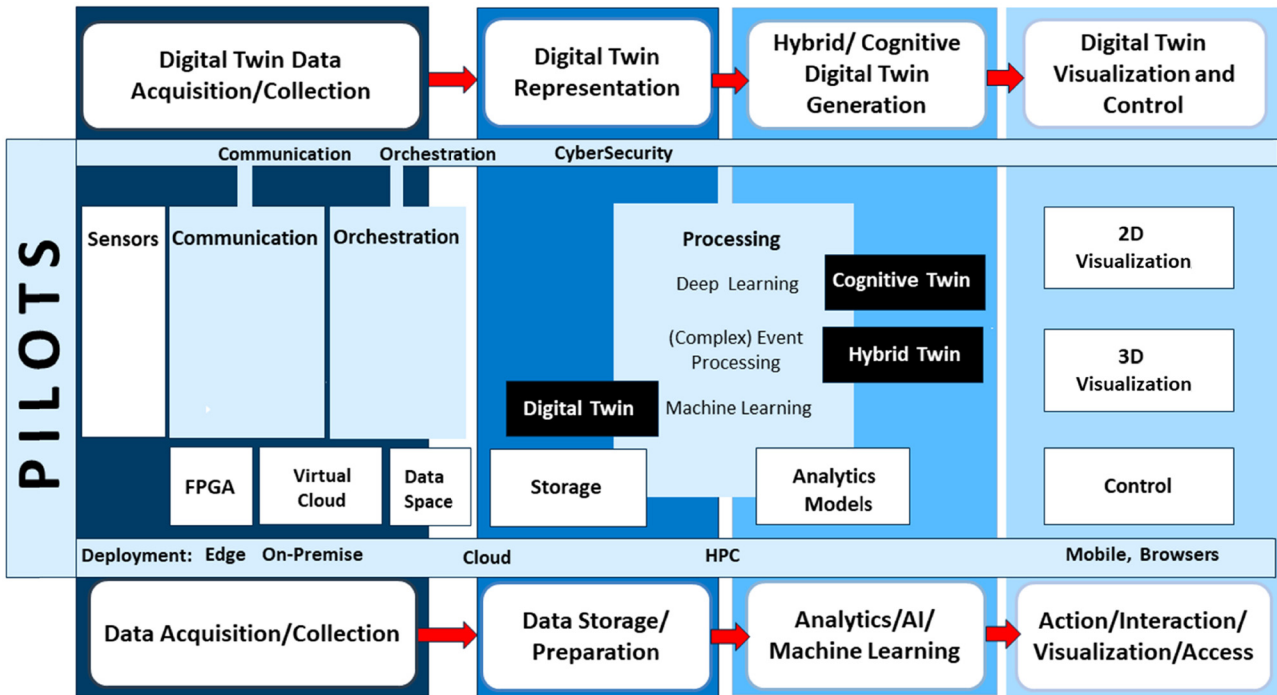


Figure 1: COGNITWIN Digital twin pipeline [21].

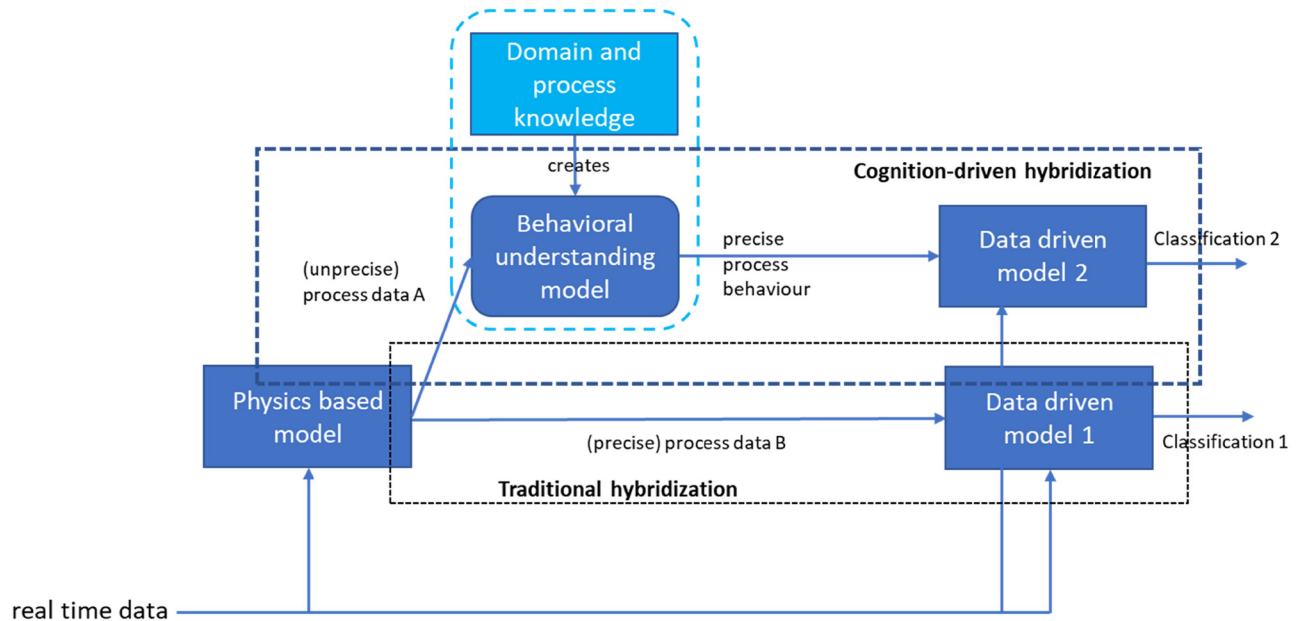


Figure 2: Cognition as an enabler of behavior understanding.

is of such a critical nature that an action will be a synthesis of DT prediction and an operator's approval.

In a nutshell, the cognition process involves comprehensive modeling of various knowledge elements that can support interpretation of process changes, including detection and interpretation of unexpected process variations.

The proposed DT abstraction layer [4,10] contains knowledge on (a) how to model dynamics in particular models, and (b) how to interpret variations in the models. Assuming that all models created are correct in principle (albeit with some limitations), the output values can therefore be interpreted on the level of the behavior of the process

by understanding the output data in the context of the process variations. In simple terms, the cognitive layer may be seen as primarily a qualitative assessment, but may change the models, methods, and output, by exploiting quantitative data and experiences.

## 2 Overview of pilots related to cognitive digital twins

In this section, we briefly describe five different pilots for cognitive plants by explaining their approaches to digital twins and cognition. An overview over the pilots is provided in Table 1.

### 2.1 Hydro

The Hydro pilot aims to increase insight into the current status of the Gas Treatment Centre (GTC), with the end goal of evening out the fluoride content of the raw alumina feed proceeding to the core electrolysis process. The GTC supports the core electrolysis process that produces Hydro's main product: aluminum. Hazardous, high HF (hydrogen fluoride)-content fumes are a byproduct of the aluminum production process. The GTC has multiple functions: (1) to remove continuously generated fumes from the electrolysis cells using powerful fans, (2) to clean these fumes of toxic species (mainly HF) such as to avoid releasing hazardous gases into the workplace and the environment, and (3) to return fluoride (emitted HF) to the electrolysis cells, where it has an important role in maintaining optimal conditions for aluminum production.

The COGNITWIN project works with data from Hydro's Karmøy Technology Pilot (KTP) because of the improved digital infrastructure and greater availability of logged data signals. Fresh alumina is shipped to Karmøy and enters the GTC via filter compartments. While traveling through the baghouse filters and adsorption reactors, the alumina is exposed to and adsorbs the HF present in the fumes before being sent to a silo for secondary (HF-rich) alumina. The secondary alumina is distributed to the individual electrolysis cells where it serves as both the raw material for aluminum production and a source of elemental fluoride, which is necessary for the optimal performance of the electrolysis process.

Multiple factors affect the levels of HF present in the fumes and the subsequent HF-content of the secondary

alumina returned to the electrolysis cells. This variable HF-content can result in fluctuating amount of fluoride species in the cell and influence the efficiency of the aluminum production process. Close monitoring as well as compensatory measures to stabilize the alumina composition are therefore needed to optimize cell performance. The fluoride content of alumina is, however, difficult to determine. Relevant measurements are occasionally taken but are not as of yet with high enough frequency to be used directly in process control. In addition, no tools for anticipating upcoming changes to HF-content currently exist.

The COGNITWIN project aims to develop a digital twin technology that can predict the current HF-content of secondary alumina, anticipate future disturbances to the HF-content based on forecasted process inputs, and help operators adjust GTC operation to minimize variation in alumina composition.

The model-based digital twin is currently running online in monitoring mode (not controlling the process) with data from many different sources, including the GTC process, electrolysis cells, raw alumina certificates, and the weather. The digital twin produces real-time predictions of the GTC process. The accuracy of the digital twin can be improved by self-adaption using data-driven estimation techniques based on model agreement with process measurements. The resulting hybrid digital twin is shown to follow the dynamics of the process measurement well, as demonstrated in Figure 3. By combining model predictions with process measurements, the hybrid digital twin enables soft-sensing of the otherwise unpredictable composition of alumina.

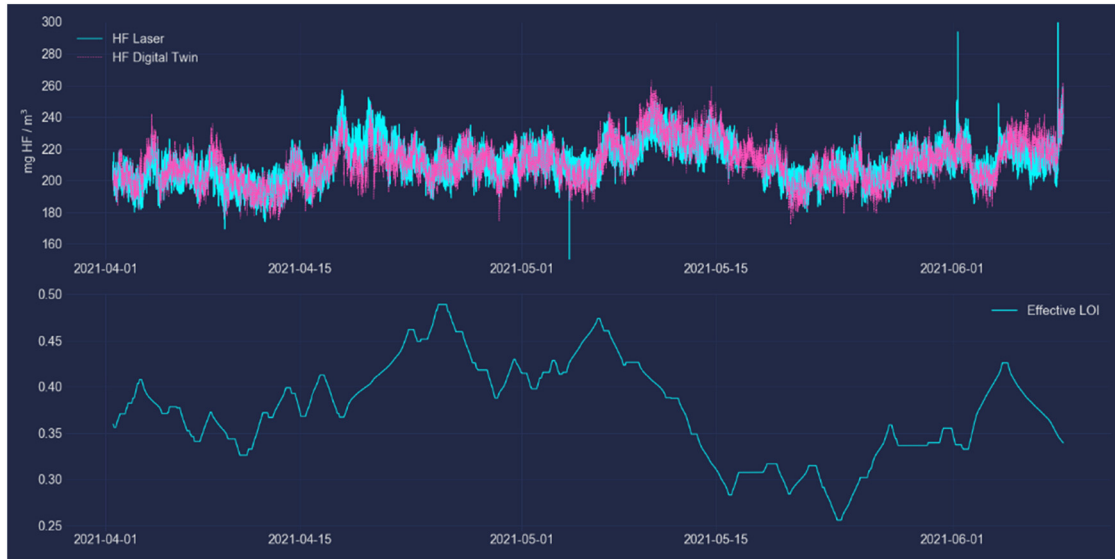
The digital system can be extended with information from the weather forecast in order to anticipate the future HF-content of alumina. The ability to predict disturbances to the alumina composition in advance allows for the effective optimization of GTC operation. Operator knowledge and experience will be used to define optimization schemes, hence adding a cognitive aspect to the twin.

### 2.2 ELKEM

The goals for the Elkem pilot are to increase product hit rate and maximize the use of recycled material for the ferrosilicon refining process. Ferrosilicon is produced in a submerged arc furnace and tapped into a transportable ladle at regular intervals. Typically, a batch of liquid ferrosilicon at approximately 1,600°C is tapped. The composition of the tapped metal varies with variations in the

Table 1: Overview over five COGNITWIN pilots

| Pilot                              | Hydro   | Elkem  | Sidenor  | NOKSEL   | Sumitomo SHI FW   |
|------------------------------------|---|--|--|--|---|
| Type of industry                   | Aluminum producer   | Silicon producer   | Steel producer   | Large diameter welded steel pipe producer                                  | Technology provider. Delivers complete combustion plants for energy production  |
| Pilot objectives                   | Increase raw material stability and reduce energy costs in the Gas Treatment Center | Metal tapping and refining: Improved control on metal alloying and refining, reduced energy consumption and increased ladle lifetime | Improve steel ladle refractory life-time without compromising safety   | Anomaly detection, reduced machine downtime and reduced energy consumption | Improved boiler efficiency, reduce operating costs, reduce emissions, improve both reliability and availability of plants |
| Primarily physics-based models?    | Yes   | Yes  | Yes  | No, only sub-models are physics based                                      | Yes   |
| Model calibration by sensor data   | Yes   | Yes  | Yes  | Yes  | Yes   |
| Primarily data-based models?       | No. However model optimization from data (anomaly detection)                        | No, but model for slag composition is purely from data   | No, but alternative data-based models were developed   | Yes  | No. The complete model uses a combination of physics-based and data-based models  |
| Can model generate real-time data? | Yes   | Yes  | “Yes” – the model can run and generate new data as soon as data are available, which is some time after the event has happened | Yes  | Yes   |
| Hybrid twins                       | Yes   | Yes  | Yes  | Yes  | Yes   |
| Cognitive twins                    | Yes   | Yes  | Operator decisions are supported by model predictions and application of knowledge graphs                                      | Yes  | Yes   |



**Figure 3:** Results from the digital twin technology for the GTC. Digital twin predictions and agreement with measurement can be monitored by operators. (Top) The hybrid digital twin predictions for HF-concentration in the fumes closely follow the dynamics of the process measurements made via HF-laser. (Bottom) Self-adaptation of a model parameter to improve digital twin prediction accuracy.

raw materials used in the furnace. After tapping, the ladle is transported to the refining station where the alloying and refining takes place. The composition of the metal is adjusted during refining and alloying to achieve the target composition of the specific grade. The tapped metal carries excess heat, and recycled ferrosilicon is utilized to cool the metal down to the correct temperature for casting. Currently, the registered amount of each addition is calculated by the control system and corrected manually by the operators based on their experience.

Sampling the process is highly challenging due to the high temperature of the metal and is therefore only performed once or twice for each batch. The processing time for the samples is also significant, making it unsuitable for direct use in an automatic control system.

The COGNITWIN project aims to arrive at a cognitive digital twin that will aid the operators with choosing the correct amount of each addition which will both achieve the grade specification, minimize cost, and maximize the use of recycled ferrosilicon.

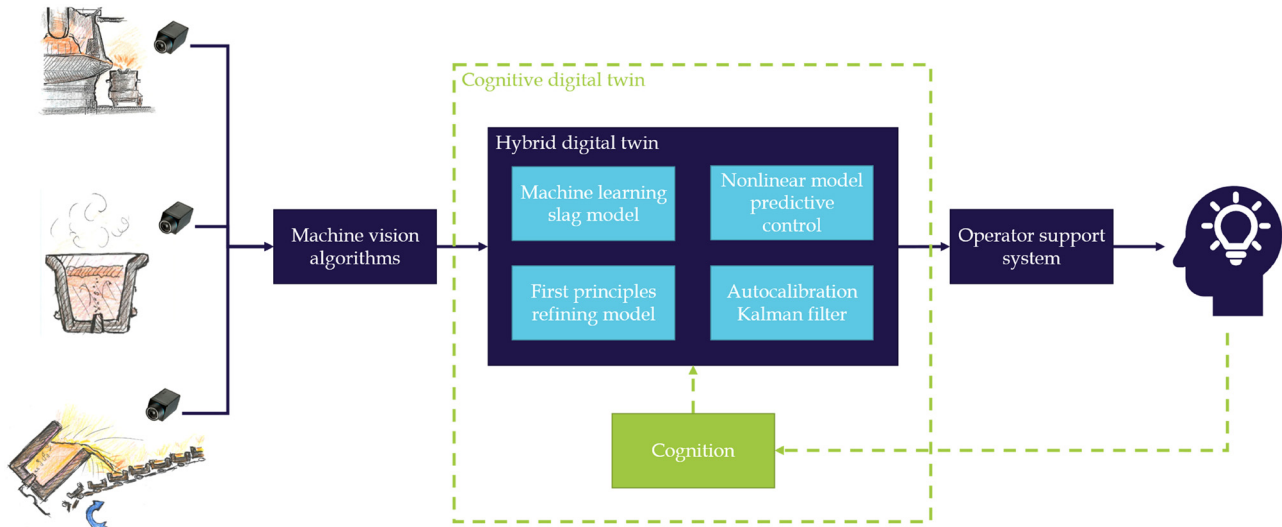
The foundation of the cognitive digital twin is a hybrid digital twin. A model based on first principles is currently installed at the plant and running online as a digital twin. The digital twin can predict the evolution of the temperature, the composition, and mass of slag/metal for the refining process with good accuracy ( $\sim 1\%$  standard deviation). The prediction accuracy can be further improved by extending the digital twin to a hybrid digital twin, utilizing a data-driven model to calculate the amount of slag that is tapped from the furnace into the ladle. The

digital twin will also employ a self-adapting algorithm, e.g., an augmented Kalman filter, to correct the first-principles model in real time. The hybrid digital twin will use new thermal cameras combined with a set of machine vision algorithms to extract additional information from the process. The optimum amount of each addition will be calculated by a nonlinear model predictive control algorithm. A flowchart for the hybrid digital twin is shown in Figure 4.

Once the hybrid digital twin is running online, it will be extended with cognitive elements, including self-adapting algorithms with the possibility to learn from the actions of the operators (shown in green in the figure). When the system has learned from a wide enough range of scenarios, operator intervention should become less frequent.

## 2.3 SIDENOR

The Sidenor pilot is dealing with optimization of the refractory life-time of ladles in steel production. The steel, produced from scrap smelting in an electric arc furnace, is tapped into a ladle where alloying and refining of the steel takes place. The ladle is a cup-shaped vessel (see Figure 5) that may typically contain 100–140 tons of steel. The ladle wall contacting the steel is made of refractory bricks that are designed to cope with the high temperatures of and the corrosiveness of the steel and slag. The slag is made of liquid and viscous metal oxides and is often designed



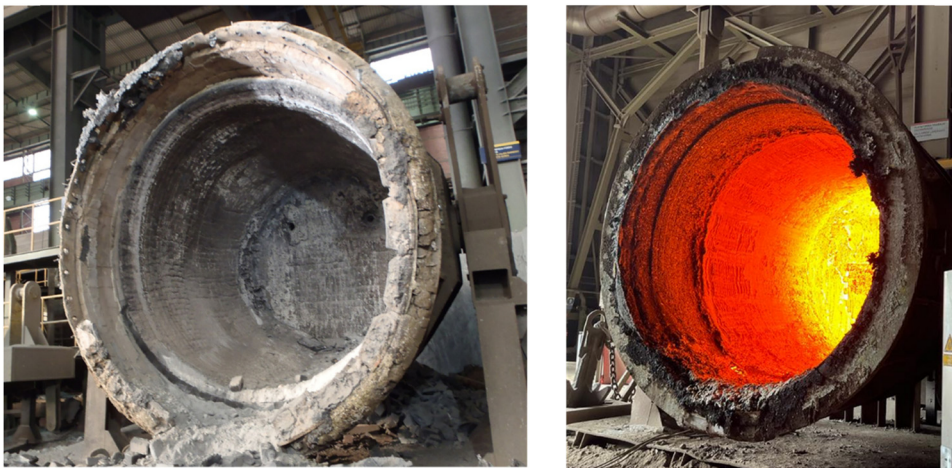
**Figure 4:** A flowchart for the hybrid digital twin. A thermal camera is observing each part of the process, tapping, refining and casting. The thermal image data are processed by a set of machine vision algorithms before being used by the hybrid digital twin. The optimized results from the hybrid digital twin are shown to the operators via a graphical operator support system.

to support the removal of unwanted impurities in the steel. The inner refractory lining is being consumed by each heat, and after a number of heats, the lining is so thin that it is deemed unsafe to continue to use the ladle without risking leakages into the plant, risking the life of people and making costly damages on building and equipment.

Currently, the assessment of having one more safe use (one more heat cycle) is done by visual inspection by experienced operators. As an outcome of the COGNITWIN project, we aim at creating a cognitive digital twin that will be able to assist the operator to decide if it is safe to use the ladle for one more heat cycle.

To build the model, operational data for several years are available, as well as information of the actual wear of the refractory at the time the ladle was deemed ready for relining. Direct machine learning on the data has been attempted, but the available data are not sufficient to create a data-based model that alone is able to reliably predict the remaining lifetime of a ladle. This is due to not only lack of relevant data, but also data points with various problems that need a lot of attention to penetrate.

A physics-based model with proper predictive capabilities might be possible, but with the complexity of the process and the limited time and resources available, this is also infeasible. Therefore, a hybrid approach is sought.



**Figure 5:** Left: Used steel ladle taken out for demolition of lining. Right: Hot steel ladle, after visual inspection and waiting for next heat.

The physics-based model is a transient model able to predict the thermal evolution in the refractory walls and the erosion of the refractory lining. The results are depending on the history of the ladle to be accurate, which means we need to be able to follow the entire lifetime of the ladle.

The physics-based model will assist the data-driven model in two different ways. First, the process data sometimes have data points that are not valid for various reasons. The physics-based model can be used to find these so they can be disregarded by the machine learning model. Next, the results from the physics-based model will be used to generate additional data that can be used in a machine learning model to ultimately create a cognitive digital twin.

The system will be realized in a StreamPipes<sup>4</sup>-based application. The cognitive additions will include self-adapting algorithms, but even more important the interaction with the experienced operator. Clearly, the model and the operator must come to some agreement if the model allows one more heat while the operator says no! A long-term goal will be to have a cognitive twin that can make such assessments alone, at least under certain conditions.

## 2.4 NOKSEL

The NOKSEL pilot aims to enable predictive maintenance of Spiral Welded Steel Pipe (SWP) machinery to reduce the energy consumption in steel pipe manufacturing [22,23]. The SWP machine is composed of electromechanical systems. Because of the multi-step, serial, and interdependent nature of the production process followed by the SWP machine, even a single malfunction can stop the whole production, where the cost of unplanned machine breakdown is very high.

To avoid unexpected machine breakdowns, reduce maintenance costs, and lessen energy consumption of the machine, TEKNOPAR has enhanced TIA Platform elements and provided TIA AssetHealth solution, a cognitive digital twin, at the NOKSEL facilities in İskenderun [<https://tia-platform.com>, <https://tia-platform.com/product/tia-assethealth.html>]. The selected sensor set installed on their optimum places, coupled PLCs, and control applications at the IoT level provided data acquisition. TIA AssetHealth uses acquired past and stream data for

functions/applications: to monitor the real-time condition of the machine (TIA MONITORING) and to enable predictive maintenance (TIA PREMA). The online condition monitoring functionality of TIA AssetHealth includes descriptive statistics analysis (TIA STATISTICS and TIA METRICS), and overall equipment effectiveness calculations (TIA OEE). To enable predictive maintenance remaining useful life (RUL), estimations have been performed (TIA PREMA). Visualization was implemented by TIA UX applications.

TIA CONTROL subsystem has realized a cognitive and proactive dimension of the CT. Based on big data analytics and developed ontologies, TIA CONTROL is being used to prevent one of the common reasons for SWP machine failure. TIA CONTROL system enables autonomous decision-making to conduct preventive maintenance [23].

Figure 6 presents the TIA AssetHealth applications for the NOKSEL pilot.

During the development of the TIA AssetHealth for the NOKSEL pilot, some challenges have been met and solved. Table 2 lists these challenges, associated modules, and methods/tools to solve these challenges.

## 2.5 Sumitomo SHI FW

The Sumitomo SHI FW pilot deals with optimization of operation of circulating fluidized bed (CFB) energy boilers. The work has developed a digital twin-based system for management of fouling (Figure 7) at the flue gas heat exchange surfaces, exploiting physical and data-driven models, online process data, and novel sensing. The aim of the fouling management system is to help the operator of the power plant to optimize the boiler controls in such a way that boiler operation economy is closer to optimal and the emissions and downtime of the boiler are reduced. The focus is on the optimization of soot-blowing sequence timing, via improved monitoring of fouling and better monitoring of incoming fuel properties.

Utilization of biomass and waste fuels in CFB boilers is often connected with an elevated risk of bed agglomeration, fouling, and corrosion. A better knowledge of the incoming fuel characteristics enables to monitor fuel quality changes and study its impact on fouling at heat exchange surfaces. Direct online measurement of fuel characteristics is expensive and laborious. Hence, a physical model-based on-line state estimation problem has been formulated. The developed approach provides an estimate of the distribution of different fuel fractions in

<sup>4</sup> <https://streampipes.apache.org/>.



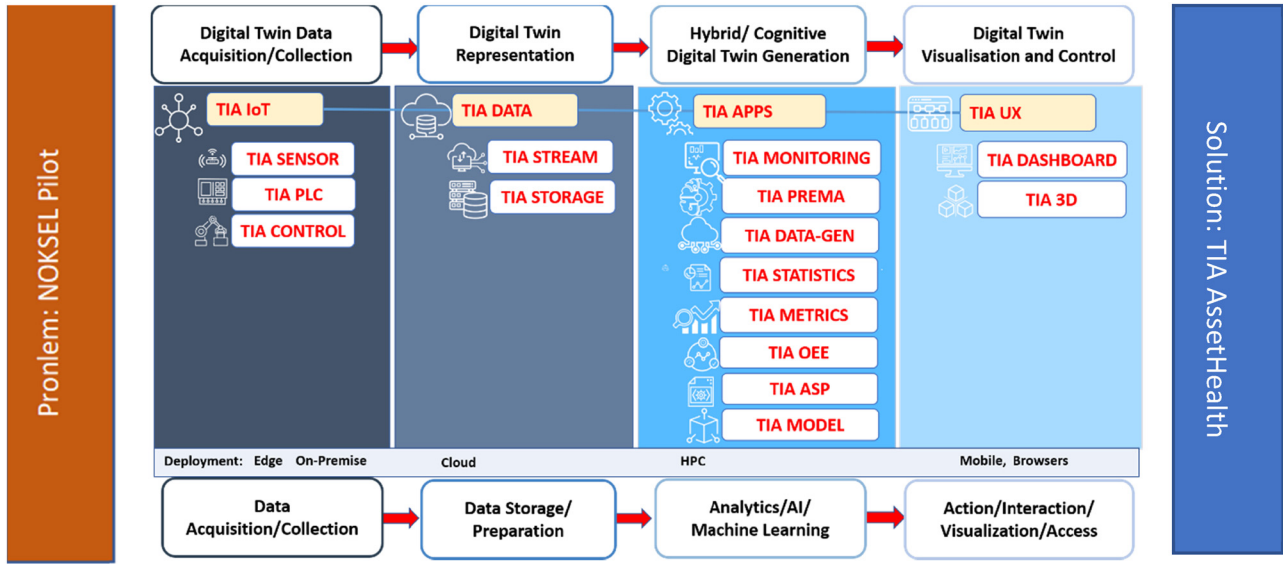


Figure 6: The cognitive digital twin developed and deployed for the NOKSEL pilot.

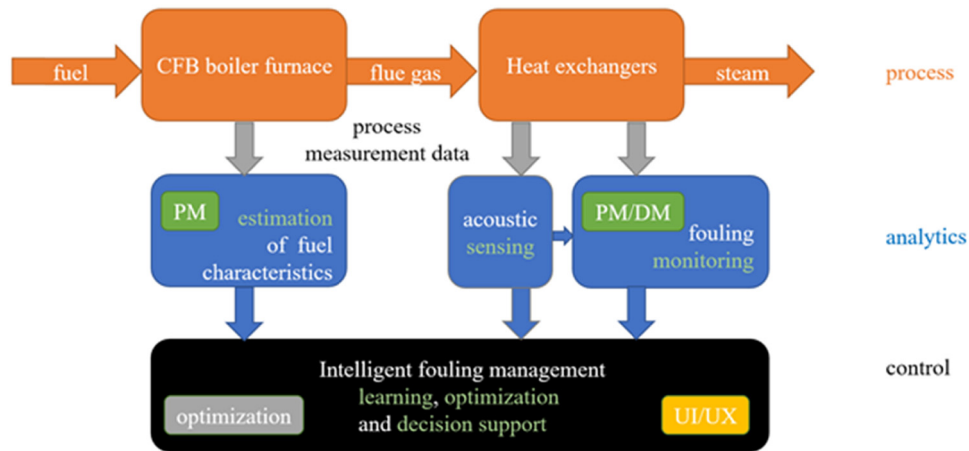
Table 2: Challenges and methods/Tools to solve them for DTs in the NOKSEL pilot

| Related Challenges                   | Related Module(s)          | Methods/Tools   |
|--------------------------------------|----------------------------|---|
| Harsh environment for sensor network | TIA IOT                    | Noise reduction   |
| Data cleaning                        | TIA DATA                   | Pre-processing  |
| Bottleneck in data transfer          | TIA DATA                   | Optimization of Kafka topic numbers   |
| Data synchronization                 | TIA DATA                   | Pre-processing  |
| Noise in real data                   | TIA IOT, TIA DATA          | Filters   |
| Latency in big data visualization    | TIA DATA, TIA APPS, TIA UX | Creation of database views  |
| Inefficient RAM utilization          | TIA DATA                   | Using Dask dataframes to scale Pandas workflows   |
| Lack of real data                    | TIA APPS                   | Synthetic data generation: TIA DATA-GEN   |
| Vibration sensor modeling            | TIA APPS                   | TIA DATA-GEN tool   |
| Model calibration                    | TIA APPS                   | Iterative model comparisons/ updates  |
| Including expert knowledge           | TIA APPS                   | Interviews and meetings with experts to determine model parameters, their associations and to evaluate results generated by the models, use of deep learning algorithms |

the feed. The solution uses a detailed nonlinear dynamic physical model for the boiler furnace, fusing the model predictions and process on-line measurement data in the framework of unscented Kalman filtering.

Fouling is a phenomenon where material accumulates on the surface of heat exchangers. To ensure efficient heat transfer, the surfaces need to be cleaned regularly. A common practice at industrial plants is to use soot blowing with a predetermined fixed time interval. It would be beneficial to guide the soot blowing based on online monitored information, however. Slag buildup can be monitored by estimating the heat transfer efficiency using direct process data over the heat exchangers, but this method is error-prone due to the many variables involved.

Therefore, several approaches of applying advanced model-based state estimation techniques for improved estimation have been developed, making use of both physical and data-driven models, and fusing predictions with online data in model-based state estimation. The ensemble Kalman filtering was applied together with a physical heat exchanger model to monitor the heat transfer coefficient online. A computationally less demanding approach was developed for optimizing the next soot-blowing sequence start time, based on adaptive subspace identification for determination of prediction models used in mixed-integer optimization. In addition, the possibilities for direct monitoring of fouling on the heat exchanger tubes have been investigated, based on acoustic sensing and signal analytics in the frequency domain.



**Figure 7:** Intelligent fouling management. PM stands for physics-based modeling, DM for data-driven modeling, and UI/UX for user interface/experience.

The industrial goal is to help the power plant operator to optimize the boiler controls, such as soot-blowing, and potentially also longer-term maintenance planning. In addition to efficient optimization, the fusion of the human operator in the decision-making loop is essential, to ensure that the improved knowledge is fed back to impact the plant operation.

### 3 An analysis of the different approaches

In this section, we discuss the similarities and differences between the approaches applied in the different COGNITWIN pilots.

#### 3.1 Hybrid digital twins

The following discussion will be based on the Wikipedia definition<sup>5</sup> of a DT, referred to in the Motivation section: “A digital twin is a virtual representation that serves as the real-time digital counterpart of a physical object or process. Digital twins are the result of continual improvement in the creation of product design and engineering activities.”

In the Hydro pilot, we introduced hybridization through improving DT accuracy by self-adaptation, using data-driven estimation techniques based on model agreement

with process measurements. By combining physics-based model predictions with process measurements, the resulting hybrid<sup>6</sup> digital twin (HT) enables soft-sensing of data that is otherwise unpredictable.

In the Elkem pilot, a first principle online model is installed at the plant. The DT prediction accuracy is improved by utilizing a data-driven model to calculate the amount of slag that is tapped from the furnace into the ladle. The digital twin will also employ a self-adapting algorithm, e.g., an augmented Kalman filter, to correct the first-principles model in real time. The hybrid digital twin will use thermal cameras combined with a set of machine vision algorithms to extract additional information from the process. The optimum amount of each addition will be calculated by a nonlinear model predictive control algorithm.

For the Sidenor pilot, a first-principle model was developed, in parallel with development of ML models. These developments were initially independent. However, as the physics model was evolving, some crucial data were questioned as those did not make sense according to the model. When those points were investigated, reasonable explanations were provided by the pilot and alternative data were obtained and applied. It was observed that a ML model can be developed to reproduce data acceptably, even if data that explain the results, in hindsight, can be found to be not what was assumed. It will be interesting to explore if the model is still correct (providing new explanations) or the model have prediction issues due to limited amount of data. Hybridization is achieved by (i) Systematically tuning of the

<sup>5</sup> [https://en.wikipedia.org/wiki/Digital\\_twin](https://en.wikipedia.org/wiki/Digital_twin).

<sup>6</sup> Here, hybrid is referring to a physics based model that is tuned and improved by learning from process data. Hybridization may also be based on exploiting different data-based models.

model against the data, (ii) “repairing” data in the cases where lack of consistency is found in the data, (iii) using the model to provide new simulated data for the ML model which have not been measured, and (iv) combined model that exploits the physics-based prediction, ML/AI methods and plant data.

In the Noksel pilot, a Hybrid Twin (HT) is developed, where the first-order models and ML/AI models are integrated into a model for a HT. An iterative process is used, similar to the approaches applied in a majority of the pilots. As explained above, also here the predictions from the data-driven model will be taken as the data that will be applied by the hybrid algorithms to improve the overall model prediction power.

In the Sumitomo SHI FW pilot case, multiple physics-based models are combined with extensive plant data using AI/ML in adaptive/learning methods for data-driven model construction and tuning, and population-based state estimation techniques. Hybridization is introduced by continuously estimating unmeasured process quantities, exploiting non-linear physics-based models and process operating data, applying nonlinear Kalman filtering. The approach of DT hybridization here is strongly related to the approaches in the Hydro and Elkem pilots. The applied methods and tools have application in a large range of similar industrial problems.

Several of the pilots are using or exploring StreamPipes to orchestrate the DT. This is an efficient platform for data collection, further expansion of the models and their interaction.

### 3.2 Cognition

The following Wikipedia definition of Cognition<sup>7</sup> is used for this analysis:

Cognition refers to “the mental action or process of acquiring knowledge and understanding through thought, experience, and the senses.” It encompasses many aspects of intellectual functions and processes such as perception, attention, the formation of knowledge, memory and working memory, judgment and evaluation, reasoning and “computation,” problem solving and decision making, comprehension and production of language. Cognitive processes use existing knowledge and discover new knowledge.

Cognition elements are yet not fully introduced in the COGNITWIN development. However, this will be faced in

the coming months. Herein, we have however investigated some possibilities.

Involving humans to make critical decisions, but with recommendation from a numerical prediction, is a possibility that will certainly be explored in all the pilots. Similarly, self-adapting algorithms will be explored. In the Hydro case, optimization schemes will be based on operator experience and knowledge. The Noksel pilot has already developed sensing, storing, processing, and decision-making. One goal here is to develop capability to handle events that previously has not been experienced.

## 4 Conclusions

In the previous section, we have presented how the COGNITWIN project has approached the different pilots (industrial applications) in order to arrive at strong digital solutions for the involved industries. The solutions are hybrid and cognitive digital twins that are built on merges of existing and, or, new developments, applying both existing basic tools and adaptations hereof.

The toolbox [22] that is developed will be helpful for similar future projects with the industry. As the industry experience the usefulness of these digital solutions, optimizations on higher level may be wanted. This increases the complexity, but also the reward. For optimization at “mother company” level, we need to have good digital solutions for all sub levels. To do this, it becomes fundamental to realize the developments within asset management shells that enable effective and flexible communication and sharing between levels.

A major learning from COGNITWIN is that heavy process industries have limited data. Sometimes, there is no sensor technology available that can do the job, while in other cases, the sensor is available but cannot be applied due to excessive costs. It is critical that the data are well understood before use. In a majority of cases, for this industry, hybrid twins involving physics-based models is a must. The combination of physics and data-based models can offer better forecasts than each by themselves. Different strategies to arrive at hybrid twins were applied, ranging from constantly tuning the physics-based model by data, to running primarily data-based models and where virtual sensors<sup>8</sup> are explored based on physics-based models.

<sup>7</sup> <https://en.wikipedia.org/wiki/Cognition>.

<sup>8</sup> The sensor data are predicted from a model, based in indirectly measured data (see also [https://en.wikipedia.org/wiki/Soft\\_sensor](https://en.wikipedia.org/wiki/Soft_sensor)).

The cognitive elements being explored seems to involve operators as well as adaptive algorithms.

## List of acronyms and explanations

|       |  |
|-------|--|
| AI    | Artificial Intelligence  |
| CFB   | Circulating Fluidized Bed  |
| CFD   | Computational Fluid Dynamics   |
| CT    | Cognitive digital Twin   |
| DT    | Digital Twin ( <a href="https://en.wikipedia.org/wiki/Digital_twin">https://en.wikipedia.org/wiki/Digital_twin</a> ) |
| GTC   | Gas Treatment Center   |
| HT    | Hybrid Digital Twin  |
| IIC   | <a href="https://www.iiconsortium.org/">https://www.iiconsortium.org/</a>  |
| IoT   | Internet of Things   |
| ML    | Machine Learning   |
| PLC   | Programmable logic controller  |
| SPIRE | <a href="https://www.aspire2050.eu/spire/about-spire">https://www.aspire2050.eu/spire/about-spire</a>                |
| SWP   | Spiral Weld Pipes  |
| TIA   | Teknopar Industrial Automation   |
| WER   | Word Error Rate  |
| 3D    | Three dimensional  |

Other acronyms are found at <https://www.allacronyms.com/data/abbreviations>.

**Acknowledgments:** Sidenor (Bilbao), by Tamara Rodriguez Duran, is acknowledged for providing the images of steel ladles.

**Funding information:** This research was funded by the H2020 COGNITWIN project, which have received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 870130.

**Conflict of interest:** The authors declare no conflict of interest.

**Data availability statement:** Due to the nature of this research supporting data is not available.

## References

- [1] EU Commission, The European Green Deal, Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions, COM(2019) 640 Final; 2019.
- [2] Birol F. Key world energy statistics 2021 [Internet]. International Energy Agency; 2021. Available from: <https://www.iea.org/reports/key-world-energy-statistics-2021>.
- [3] A.SPIRE, Processes4Planet Roadmap: Sustainable Process Industry through Resource and Energy Efficiency; 2021.
- [4] Abburu S, Berre AJ, Jacoby M, Roman D, Stojanovic L, Stojanovic N. COGNITWIN – Hybrid and Cognitive Digital Twins for the Process Industry. In Proceedings of the 2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC). Cardiff, United Kingdom: IEEE; June 2020. p. 1–8.
- [5] Lucia DJ, Beran PS, Silva WA. Reduced-order modeling: New approaches for computational physics. *Prog Aerosp Sci.* 2004;40:51–117. doi: 10.1016/j.paerosci.2003.12.001.
- [6] Iuliano E. Towards a POD-Based Surrogate Model For CFD Optimization. In: CFD and Optimization 2011, An ECCOMAS Thematic Conference [Internet]. Antalya, Turkey; 2011. p. 21. <https://www.researchgate.net/publication/259580406>.
- [7] Ștefănescu R, Sandu A, Navon IM. POD/DEIM reduced-order strategies for efficient four dimensional variational data assimilation. *J Comput Phys.* 2015;295:569–95. doi: 10.1016/j.jcp.2015.04.030.
- [8] Lichtenegger T, Pirker S. Recurrence CFD – A novel approach to simulate multiphase flows with strongly separated time scales. *Chem Eng Sci.* 2016;153:394–410. doi: 10.1016/j.ces.2016.07.036.
- [9] Zoric J, Johansen ST, Einarsrud KE, Solheim A. On pragmatism in industrial modeling. *Progress in Applied CFD : Selected papers from 10th International Conference on Computational Fluid Dynamics in the Oil & Gas, Metallurgical and Process Industries.* vol. 1 Issue. 1, 2015. p. 9–24.
- [10] Stojanovic L, Bader SR. Smart Services in the Physical World: Digital Twins. In: Maleshkova M, Kühl N, Jussen P, editors. *Smart Service Management.* Cham: Springer International Publishing; p. 2020. 137–47. ISBN 978-3-030-58181-7.
- [11] Van Schalkwyk P, Lin SW, Malakuti S. A short introduction to digital twins. *IIC Journal of Innovation - Novemb. 2019 Ind. Internet Consort.* (iiconsortium.org). 9, 2019.
- [12] Wright L, Davidson S. How to tell the difference between a model and a digital twin. *Adv Model Simul Eng Sci.* 2020;7:13. doi: 10.1186/s40323-020-00147-4.
- [13] Kalman RE. A new approach to linear filtering and prediction problems. *Trans ASME–J Basic Eng.* 1960;82:35–45.
- [14] Simon D. *Optimal state estimation: Kalman, H infinity, and nonlinear approaches.* Wiley; 2006.
- [15] Arm J, Benesl T, Marcon P, Bradac Z, Schröder T, Belyaev A, et al. Automated design and integration of asset administration shells in components of industry 4.0. *Sensors.* 2021;21:2004. doi: 10.3390/s21062004.
- [16] Jacoby M, Jovicic B, Stojanovic L, Stojanović N. An approach for realizing hybrid digital twins using asset administration shells and apache streampipes. *Information.* 2021;12:217. doi: 10.3390/info12060217.
- [17] Jiang H, Qin S, Fu J, Zhang J, Ding G. How to model and implement connections between physical and virtual models for digital twin application. *J Manuf Syst.* 2021;58:36–51. doi: 10.1016/j.jmsy.2020.05.012.
- [18] Bradac Z, Marcon P, Zezulka F, Arm J, Benesl T. Digital twin and AAS in the industry 4.0 framework. *IOP Conf Ser Mater*

- Sci Eng. 2019;618:012001. doi: 10.1088/1757-899X/618/1/012001.
- [19] Preisig HA. Ontology-based process modelling-with examples of physical topologies. *Processes*. 2021;9:592. doi: 10.3390/pr9040592.
- [20] Mir ZM, Friis J, Hagelien TF, Svenum I-H, Ringdalen IG, Konchakova N, et al. Interoperability architecture for bridging computational tools: Application to steel corrosion in concrete. *Model Simul Mater Sci Eng*. 2020;28:025003. doi: 10.1088/1361-651X/ab6209.
- [21] Hagelien T, Preisig H, Friis J, Klein P, Konchakova N. A practical approach to ontology-based data modelling for semantic interoperability. In *Proceedings of the 14th WCCM-ECCOMAS Congress*. CIMNE; 2021.
- [22] COGNITWIN Toolbox Portal. Available online: <https://cognitwin.github.io/toolbox/> (accessed on 31 January 2023).
- [23] Unal P, Albayrak Ö, Jomâa M, Berre AJ. Data-driven artificial intelligence and predictive analytics for the maintenance of industrial machinery with hybrid and cognitive digital twins. In: Curry E, Auer S, Berre AJ, Metzger A, Perez MS, Zillner S, editors. *Technologies and applications for big data value*. Cham: Springer International Publishing; 2022. p. 299–319. ISBN 978-3-030-78306-8.