COGNITWIN – Hybrid and Cognitive Digital Twins for the Process Industry

Sailesh Abburu
SINTEF Industry, SINTEF AS
Trondheim, Norway
sailesh.abburu@sintef.no

Arne J. Berre
SINTEF Digital, SINTEF AS
Oslo, Norway
arne.berre@sintef.no

Michael Jacoby
Fraunhofer IOSB,
Karlsruhe, Germany
michael.jacoby@iosb.fraunhofer.de

Dumitru Roman
SINTEF Digital, SINTEF AS
Oslo, Norway
dumitru.roman@sintef.no

Ljiljana Stojanovic
Fraunhofer IOSB,
Karlsruhe, Germany
ljiljana.stojanovic@iosb.fraunhofer.de

Nenad Stojanovic
Nissatech
Niš, Serbia
nenad.stojanovic@nissatech.com

***Abstract* — The concepts of Hybrid and Cognitive Digital Twin are introduced as elements of the next level of process control and automation in the process and manufacturing industry. We propose an architecture for the implementation of Hybrid and Cognitive Twins as part of the COGNITWIN software toolbox. The toolbox is designed to cover cognitive capabilities for optimal operations and maintenance of process equipment and assets, thereby minimizing production overheads and increasing efficiencies for the process industry. Furthermore, we identify a set of relevant use cases in the process industry and discuss the possible applicability and use of the toolbox.**

Keywords—Process Industry, Intelligent factories, Digital Twins, Software architecture, Software toolbox

I. Introduction

The process industries and the manufacturing sector have already embraced the digitalisation wave and have reaped much dividends from the investments made in the technological and the scientific forefronts. This also raises an important question: have we reached the pinnacle of digitalisation or is there still scope for further innovations on the progress made thus far? This notion is easily thwarted by looking at other sectors of the economy such as the electronics and the service industry. For instance, the ubiquitous use of chat-bots by service industries to provide optimal answers to address customer concerns. This ability of the machines to learn and almost have humanlike ability to think is the next step in the digitalisation scheme and gives machines the cognitive elements that they previously lacked. For example, IBM's Watson cognitive system in healthcare helps identify or diagnose a disease using several inputs that the span several factors that include knowledge about the condition, patient health history, knowledge from journal articles, best practices, diagnostic tools, etc., which the system collates and analyses to provide the best recommendation for treatment [1]. The scope and complexity of the cognitive twin models extend beyond the typical collection of empirical and first principle-based models for processes. Can we take a leaf out of these examples and transfer those ideas to the manufacturing sector where algorithms can teach a machine to think like humans (cognitive capacity) to move towards "*intelligent*" factories?

This article presents a brief overview of our approach towards implementation of cognitive aspects for the process industries. The "cognitive element" will be introduced by learning from historical process data and events to first predict unwanted events in the operation before they happen and offer best possible solutions. We aim to augment existing ideas and technologies that have been validated in controlled environments to arrive at prototype demonstrations in operational environments. The typical set-up includes a sensor network that will continuously monitor and collect data from various plant processes and assets, which will be stored in a database. This data will be used to develop a digital twin of the process and will be used to develop models with cognitive capabilities for self-learning and predictive maintenance, which will lead towards optimal plant operations. The COGNITWIN toolbox proposed in this paper is meant to be customized towards specific challenges in process industry. It features generic aspects that enable interested parties to quickly deploy cognitive systems in their processes.

To realize the envisioned cognitive aspects for the process industries this paper introduces the notions of Hybrid and Cognitive Twins as extensions of Digital Twins and proposes a software architecture for the implementation of Hybrid and Cognitive Twins as part of the COGNITWIN software toolbox. Furthermore, we identify a set of relevant use cases in the process industry and discuss the possible applicability and use of the toolbox. Although the most important contribution of this paper is at the conceptual level, we also identify the research and technical challenges to be addressed in the future in this context. The rest of this paper is organized as follows. Section II presents the COGNITWIN definitions for various twins. Section III addresses standards and relevant approaches for twins. Section IV gives an overview of the COGNITWIN toolbox. Section V presents a set of relevant use cases, and finally Section VI concludes this paper.

II. Definitions: Digital Twin, Hybrid Digital Twin, and Cognitive Digital Twin

In the literature, one can find discussions around various kinds of twins such as "Virtual Twin", "Digital Twin" "Predictive Twin", "Hybrid Twin", "Semantic Twin", "Executable Digital Twin", "Cognitive Twin", "Associative Cognitive Digital Twin", "Contextual Twin", "Experimental Twin", "Process Twin", "System Twin" etc. (see, e.g., [2] [3] [4] [5] [6] [7] [8] for examples of recent papers introducing some of these concepts). Since there is no widespread consensus on the definition of various types of "twins", we provide in the following working definitions for the types of twins relevant in our context – to go from digital replicas of physical assets to "cognitive augmentation" of physical assets via a three-layer approach: digital, hybrid, and cognitive (see Fig. 1). The three-layer separation is motivated by the need to first create isolated models of physical systems (e.g., for detection of anomalies in isolation), then interconnect the

978-1-7281-7037-4/20/$31.00 ©2020 IEEE

2020 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)

Authorized licensed use limited to: Malardalen University. Downloaded on August 07,2021 at 16:28:40 UTC from IEEE Xplore. Restrictions apply.

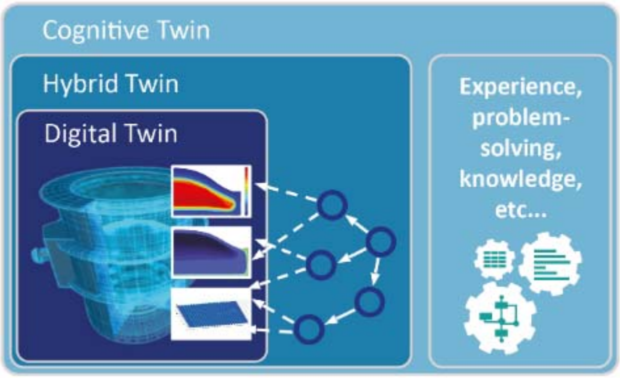

Fig. 1: COGNITWIN 3-layered approach to define "twins".

models (e.g., for integrative prediction of unusual behaviour), and then extend them with expert and problem-solving knowledge (e.g., for dealing with unknown situations).

Digital Twin (DT) A digital replica of a physical system that captures attributes and behaviours of that system. The purpose of a DT is to enable measurements, simulations, and experimentations with the digital replica in order to gain understanding about its physical counterpart. A DT is typically materialized as a set of multiple isolated models that are either empirical or first-principles based.

Hybrid Digital Twin (HT) An extension of DT in which the isolated DT models are intertwined to recognize, forecast and communicate less optimal (but predictable) behaviour of the physical counterpart well before such behaviour occurs. A HT integrates data from various sources (e.g., sensors, databases, simulations, etc.) with the DT models, and applies AI analytics techniques to achieve higher predictive capabilities, while at the same time optimizing, monitoring and controlling the behaviour of the physical system. A HT is typically materialized as a set of interconnected models, achieving symbiosis among the DT models.

Cognitive Digital Twin (CT) An extension of HT incorporating cognitive features that will enable sensing complex and unpredicted behaviour and reason about dynamic strategies for process optimization, leading to a system that continuously evolve its own digital structure as well as its behaviour. A CT is thus a hybrid, self-learning, and proactive system that will optimize its own cognitive capabilities over time based on the data it will collect and experience it will gain. A CT will find new answers to emerging questions by combining expert knowledge with the power of HT. A CT will thus achieve synergy between the HT and the expert and problem-solving knowledge.

III. Standards and Other Relevant Approaches

The general idea of the DT concept is to have a digital representation of a physical system as an (standardized) interface for digital interaction with the system. On a conceptual level, such DTs are envisioned to provide a multitude of functionality, e.g., management of documents related to the system, provide visualization and 3D representation, access to (historical) data about the system, and methods to interact with the system [2].

As DTs are not intended to be used by end-users directly but rather by software systems (agents or visualisation tools), standardization of APIs (Application Programming Interfaces) is essential. In the following, we present current standards and implementations covering the basic set of functionalities of a DT: resource description, resource discovery, and resource access. As these functionalities are essentially the same as the Internet of Things (IoT), we also analyse standards and platforms related to the IoT domain.

A. W3C Web of Things

The goal of the W3C Web of Things (WoT) Working Group is to counter the fragmentation of the IoT by providing so-called "building blocks" that should complete and enhance existing standards. The most important of those building blocks is the WoT ThingDescription (WoT TD or only TD)[1] which became official W3C Recommendation in April 2020. The TD provides a meta-model to describe existing resources in the IoT, e.g., a Digital Twin or a part of a Digital Twin, called *Thing*. A *Thing* comprises a set of properties (which can be read-only or read/write), actions, i.e., methods that can be invoked on the *Thing*, as well as events that can be subscribed to. Unlike others standards, the TD does not define any API how to access those things but for each requires to contain one (or even more) description of an existing API endpoint. This way, resources with existing APIs can easily be described. The TD can be serialized as either JSON-LD or a simplified JSON document. A serialized TD can be published and a consumer will be able to access the described resource. Resource discovery, i.e., how these TDs can be found, is still work in progress but is expected to be published in Q1 2021[2]. One highlight of the TD is that by making use of existing Semantic Web technologies like JSON-LD, is it easily possible to add semantic information to every aspect of a *Thing*.

B. Plattform Industrie 4.0 Specification: Details of the Asset Administration Shell

The "Plattform Industrie 4.0" is a network of companies, associations, trade unions, science and politics in Germany aiming to support development and realization of Industry 4.0. In this context, the specification named "Details of the Asset Administration Shell Part 1: The exchange of information between partners in the value chain of Industry 4.0" was published [9] [10]. The central element of the proposed model is the digital representation of an asset, called *Asset Administration Shell* (AAS). An AAS is composed of so-called submodels, each comprising a set of properties, actions (called operations) and events. AAS descriptions can be serialized using XML, JSON, RDF, OPC UA data model and AutomationML. In contrast to WoT TD, it does not allow mapping properties, actions and events to existing APIs but requires implementation of standardized interfaces per resource. The definition of these interfaces is currently work in progress and will be published as "Details of the Asset Administration Shell Part 2: Interoperability at Runtime – Exchanging Information via the Application Interface".

C. Eclipse BaSyx

Eclipse BaSyx[3] is an open-source implementation based on the AAS concepts. It originated from the BaSys 4.0 research project funded by the German Ministry for Education and Research and is actively supported by BaSys 4.2 (the

[1] https://www.w3.org/TR/wot-thing-description
[2] https://www.w3.org/2020/01/wot-wg-charter.html
[3] https://www.eclipse.org/basyx

Authorized licensed use limited to: Malardalen University. Downloaded on August 07,2021 at 16:28:40 UTC from IEEE Xplore. Restrictions apply.

successor project of BaSys 4.0)[4]. As there are no specifications on interfaces for resource access and resource discovery for the AAS yet, BaSyx tries to fill this gap with its own definitions and implementations. However, as the specifications on the AAS are still being updated and extended, the implementation deviates from the specifications in multiple aspects, e.g., by not supporting events.

D. Eclipse Ditto

Eclipse Ditto[5] is an open-source framework with the objective to bring the TD concept into existing IoT infrastructures. It does use neither existing TD standards nor tries to define a new one but rather focuses on providing stable software that can be used to realize TDs. In Ditto, resources are called *Thing* and are described by a set of features. A feature, in turn, conforms to a definition derived from an information model. Information models are described using the Vorto DSL developed by a previous Eclipse Project called Eclipse Vorto[6]. They are comprised again of a set of properties, actions (called operation) and events, organized as so-called function blocks. Function blocks can reference and re-use other function blocks from the same or even other information models. This functionality is similar to the concept of submodels in the AAS.

IV. The COGNITWIN Toolbox

The architectural blueprint of the COGNITWIN Toolbox (CTT) is shown in Fig. 2. The toolbox essentially has five layers with each of them providing a set of services.

Model-driven services, based on explicit models, such as numerical, first-principle models, delivering data created in rather complex computation processes (e.g., various types of simulations). A model-driven service has as input data from a data source and as output data that is created in real-time processing process and require a CTT-specific adapter in order to use that data in the CTT processing pipeline.

Data-driven services (DDS), relates to the various types of data analytics services, delivering data and models that describe various aspects of the system behaviour, based on data-intensive computation. The DDS falls under two categories, batch processing or real-time processing. In the first case, the data analytics algorithms are implemented on historical data whereas in real-time processing, the models are developed on data packets of small-window-size that are steamed real-time. The input to DDS is either an already available data source, e.g., from existing database, or new data source, e.g., from a recently installed sensor.

The service is responsible to develop a specific data adapter. The outputs from the DDS include: a) empirical models that are developed in a batch/learning process and require a CTT-specific adapter (for models) to store the model in the model repository (e.g. to be used for real-time anomaly detection); and b) Secondary-data that is created during the real-time processing of the data which also requires the toolbox-specific adapter (for data) in order to be used in the toolbox's processing pipeline.

A. Layers in the CTT

The architecture consists of five main layers, as follows.

Data Ingestion and Preparation Layer. The main role of this layer is to enable the integration of relevant data sources

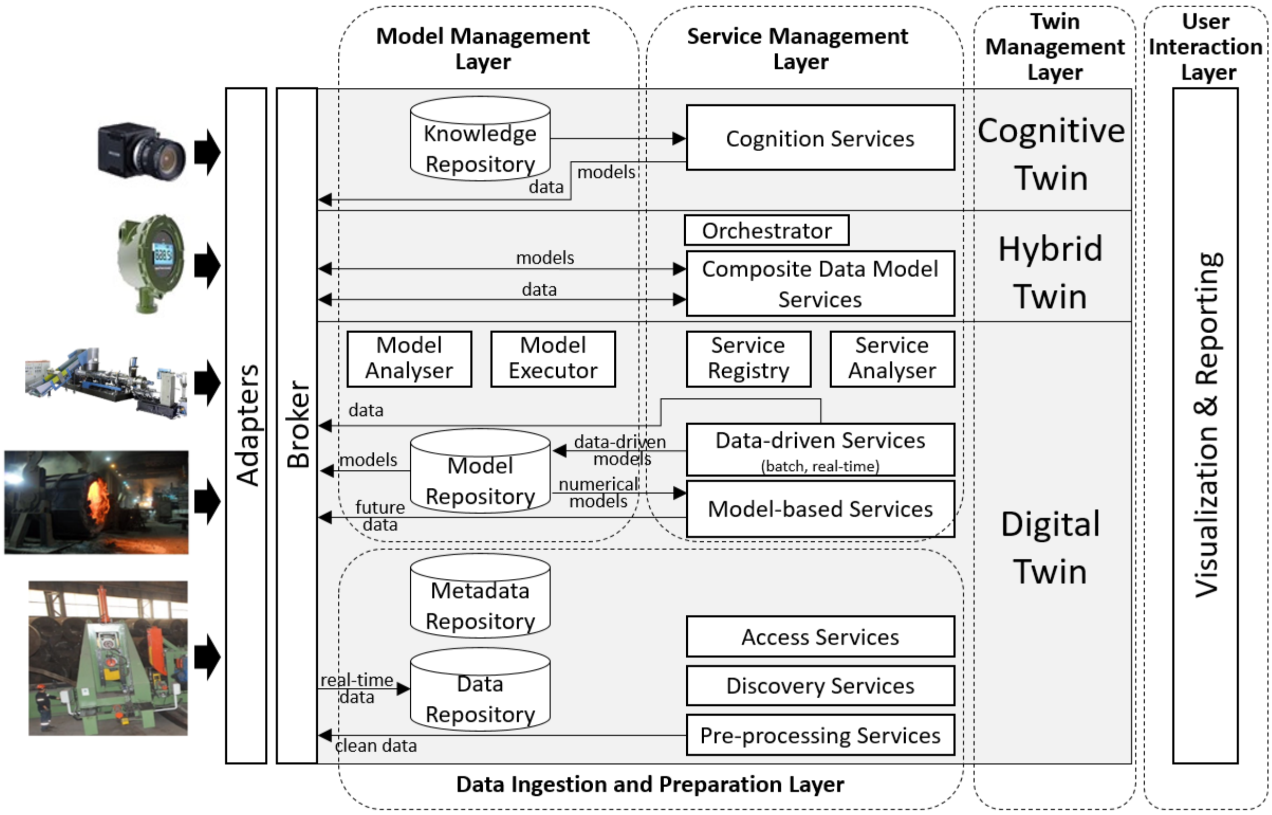

Fig. 2: Cognitive Twin Toolbox conceptual architecture and its five layers.

[4] https://www.basys40.de
[5] https://www.eclipse.org/ditto
[6] https://www.eclipse.org/vorto

978-1-7281-7037-4/20/$31.00 ©2020 IEEE

Authorized licensed use limited to: Malardalen University. Downloaded on August 07,2021 at 16:28:40 UTC from IEEE Xplore. Restrictions apply.

and the preparation of the data for further analysis. Primarily, it collects data, transforms it into suitable CTT data formats, stores it in *Data repositories* and performs various cleansing steps to prepare data for the usage in analytics services. Data is published on the *Broker* that ensures the efficiency and scalability of the data communication. The main components of this layer include: data adapters to enable adaptation various data formats based on the sources, data repository for storage and model development, metadata repository to store metadata relevant for the set-up of analyses/services, data preparation for different approaches to cleanse the data, access services to enable access to data/models, and discovery services.

Model Management Layer. This layer manages three types of models: a) first-principle models for the processes developed using underlying physics; b) empirical models developed using various machine learning and AI methods; and finally c) knowledge-driven models, related to the *tacit knowledge* of the domain experts and human operators, based mainly on their huge working experience. The main role of this layer is to ensure an efficient storage and access to the various types of models, provided by different services (model-based, data-driven or human experts).

Service Management Layer. The main role of this layer is to ensure an efficient usage of all available services for resolving underlying domain problems. It is based on a complex orchestration of services, which creates added-value processing pipelines, combining data-driven and model-based services. It includes a registry of services for enabling an efficient discovery of services required for an orchestration process. Services can publish the results on the Broker that ensures the efficiency and scalability of service communication.

User interaction layer. Since a CT represents a digital replica of a physical system and models its behaviour, it is important to support a user in exploring a CT (data, models) and its characteristics. In other words, an intuitive but explorative user interaction should be enabled.

Twin Management Layer. Since it models the behaviour of a physical system, a twin represents a complex structure that should be managed in an efficient way. Being a digital replica of a physical system, a Twin should reflect its behaviour, e.g., through having model of the normal behaviour of the system. On the other hand, a twin is a digital object that has own life cycle which is influenced by the physical system, i.e., all changes in the behaviour of the physical system should be reflected in twin structure (esp. models), as soon as corresponding data is available in the physical world.

B. Challenges in realising DTs, HTs and CTs

Standardization regarding DTs is currently focused on establishing a common technological ground to exchange information. The main area of interest is resource description. Resource discovery and access is not yet properly addressed in most standards but they are recognised to be important and solutions are currently work in progress. To create a software toolbox that can be used with DTs based on any of the different existing or upcoming standards, the main challenge is to find a generic representation of a DT. This allows integration of the CTT into systems using different standards to represent their DTs.

Thus, the key issue is to define a DT meta model to represent a DT in a way that will allow humans and systems to understand it and easily build applications. A meta model for DTs was proposed in [12]. However, this paper is focused only on the interfaces to the external tools for data exchange and model execution and to other DTs. There is also a need to represent the internal structure of a DT in a standardized way in order to avoid vendor lock-in. The meta-model should include at least the following aspects:

- *Metadata*:
  - Static data: asset, manufacturer, location, etc.
  - Dynamic data: a list of relevant multimodal parameters including the metadata about the parameters (e.g., unit of measurement, frequency, quality, etc.).
  - Models: a list of relevant models including a model type, the input/output parameters, etc.
- *Data*: a local or a remote storage of the multimodal data.
- *Models*: for each model a local or a remote storage of the model as well a service (or a reference to it) for the model execution.
- *References to other DTs*: association, inclusion, etc.

Another challenge is how to create a concrete DT for a given system. The instantiation of the DT metamodel can be done in many ways: (i) manually (e.g., by filling in a template), (ii) by writing a code to extract the needed information from relevant software systems (e.g., planning, E/MCAD, MES, ERP, etc.) or hardware systems (e.g., OPC UA server/client of a PLC), (iii) by combining/reusing existing DTs, etc. Conformance checking should be always applied to ensure that the created DT model conforms to the metamodel.

DTs are active entities that offer different services and collaborate with other entities. Thus, in addition to the provision of the DT metamodel, a standard set of functionalities, available through an API would be mandatory. The DT data, models and services should be provided to others in reactive and proactive mode. The DT model or its parts can be retrieved as needed or they can be distributed. Particular focus should be given to scalability and real-time aspects of the implemented functionalities. Functionalities should be available through an API, hiding the details of implementation. For the implementation of the API, all technical aspects of DTs, identified by IIC WG on DTs [11], should be taken into account.

To make a DT an intelligent entity, it would be beneficial to go beyond the "standard" DT services (such as get/set data/model or visualization service) and provide the services at the higher abstraction level that enhance the capabilities of DTs. This can be achieved by, e.g., incorporating services for anomaly detection, prediction and explanation, etc., or even non-functional services (e.g., to ensure security, privacy, sovereignty). These services should be described in a standardized way to be understood by other twins, software systems or end users.

To enable progressing from DT to HT, DTs must contain (physical) models. One of the challenges is to represent these models as parts of a DT. This should take into account

Authorized licensed use limited to: Malardalen University. Downloaded on August 07,2021 at 16:28:40 UTC from IEEE Xplore. Restrictions apply.

different types of models and interfaces to tools to execute these models.

Managing independent DT models is a challenge because the models are of different types and they are evolving. Combining them is even more complex, because it introduces the interoperability problem at the level of DT models. The models can be directly combined in various ways (e.g., a simulation model could generate the training data for an AI model) or a knowledge graph could be created to intertwine individual DT models. To exploit symbiosis among the models, there is a need to extend the DT metamodel by modelling their dependencies.

Finally, there are many situations where the physical system cannot be properly understood by its models, e.g., due to over-simplification of complex situations. Especially for resolving previously unseen, undesired situations, human involvement is mandatory. Thus, there is a need to provide a knowledge layer on the top of a DT/HT. This layer would allow a set of functionalities for a human user to be able to intervene and improve the DT via written or spoken interventions. To maintain a robustness, such interventions could not only help, but may also harm performance, the DT will use such "knowledge fragments" as recommendations or alternative views to complement understanding of a physical asset. In that respect, external human interventions will offer opportunities to improve a DT.

C. Reference architecture for DTs, HTs & CTs

The COGNITWIN Data Ingestion and Preparation Layer maps into the Data Management area of the BDVA (Big Data Value Association) reference model [12] and the "Data for AI" area of the AI PPP (Artificial Intelligence Public Private Partnership) [13]. The BDVA Data Management includes principles and techniques for data management including both data life cycle management and usage of data lakes and data spaces, as well as underlying data storage services and connection to the cross cutting area of data sharing platforms.

The data input is coming from the COGNITWIN sensors/actuators, which maps further into the BDVA Things/Assets and to the AI PPP "Sensing Measurement and Perception", in order to create information needed for successful decision-making, control, and learning. Of the six main BDVA data types the COGNITWIN architecture is in

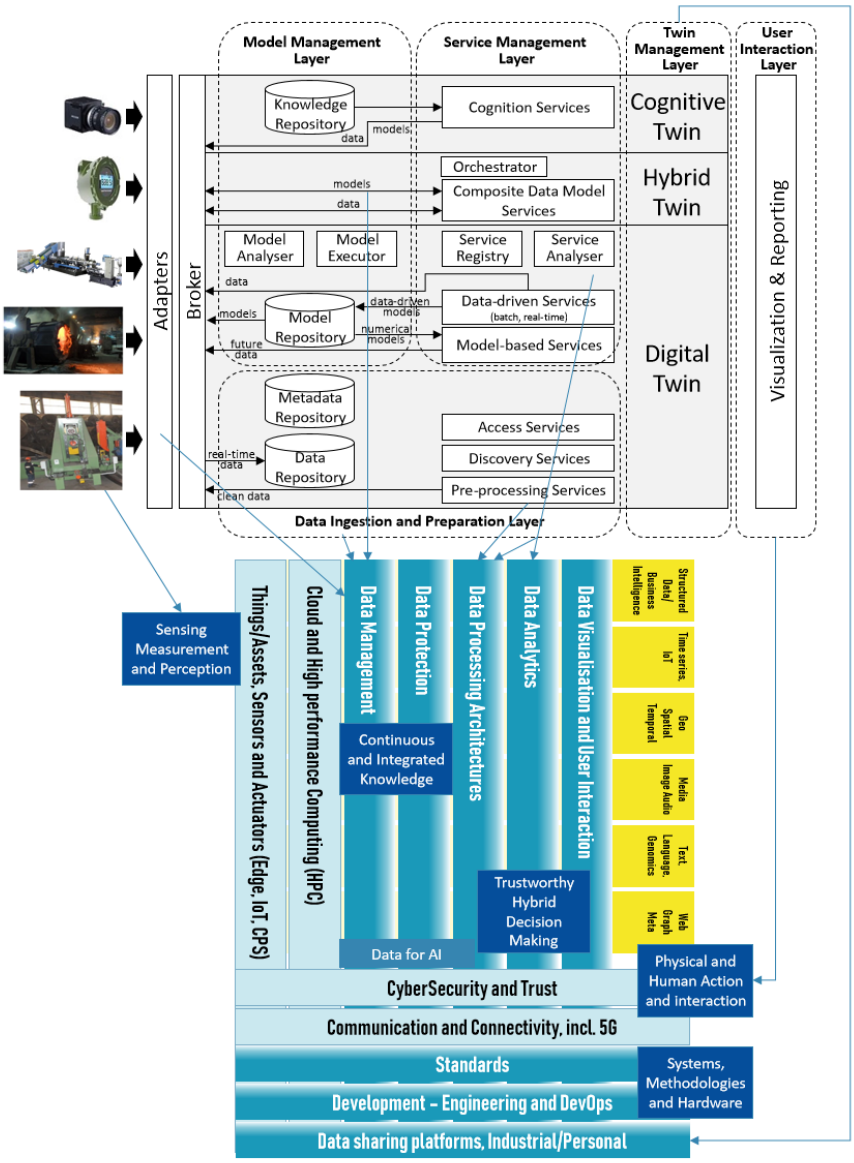

Fig. 3: COGNITWIN architecture mapped onto the BDVA reference model and the AI PPP topic areas.

Authorized licensed use limited to: Malardalen University. Downloaded on August 07,2021 at 16:28:40 UTC from IEEE Xplore. Restrictions apply.

particular focusing on IoT/time series data, spatio-temporal data, media data and graph data.

As shown in Fig. 3, the COGNITWIN Model Management Layer is mapped into the BDVA Data Management and Data sharing platforms, with the specific focus on providing efficient storage and access to the various types of models, provided by different services (model-based, data-driven or human experts). It is further mapped into the AI PPP area of "Continuous and Integrated Knowledge", which makes the sensing, measurement and perception data assets amenable to use in decision-making.

The COGNITWIN Service Management Layer is mapped into the BDVA processing and analytics area. This is supporting both batch and real time analytics for and both data-driven and model-based services and their composition. It is further enhanced with cognitive services and mapped onto the AI PPP area of "Trustworthy Hybrid Decision Making" with decision-making based on data and models combined with learning, optimisation and reasoning.

The COGNITWIN User interaction layer is mapped onto the BDVA user interaction for advanced visualisation approaches for improved user experience. It is mapped onto the AI PPP area of "Physical and Human Action and Interaction" for interactions that occur between machines and objects, between machines, between people and machines and between environments and machines.

The COGNIWTIN Twin Management is mapped onto the BDVA Data Platform for Data Sharing including in particular Industrial Data Platforms (IDPs). The Twin Management Layer is also responsible for data protection, cybersecurity, and trust.

Other relevant areas include Standards (supporting various DTs to facilitate data integration, sharing, and interoperability), Communication (necessary for providing support for data access through efficient communication with sensors and actuators as well as supporting systems) and Engineering (including DevOps. for building DT systems).

V. Hybrid and Cognitive digital twin use cases from the process industry

To highlight the ability and versatility of the COGNITWIN toolbox, we identified a set of use cases from the process industry. In this section, we discuss their respective HT and CT. These use cases represent various hurdles associated with metals production and maintenance of production assets in energy generation that are generic to several manufacturing and process industries and therefore, widens the scope of the toolbox's applicability.

A. Operational optimization of gas treatment centre (GTC) in aluminium production

Aluminium is the second largest produced metal in the world with an annual production reaching 64 million tonnes in 2019 [14]. Aluminium is typically produced by electrolytic reduction of alumina, produced from Bauxite ore, and using carbon as the anode. As a by-product of the reaction, the process also outputs a mixture of gases, Hydrogen Fluoride gas ($HF_{(g)}$) being one of them. Due to the gas's detrimental impact on the environment, the governmental and international regulations require the production plants to scrub the HF gas off before releasing the gaseous mixture to the atmosphere. HF gas is stripped from the gaseous by-products using alumina at the GTC which results in alumina that is now with adsorbed HF (secondary alumina) and is fed back into the cells. For optimal operation of the electrolytic cells, it is highly desirable that the HF concentration in the alumina feed is uniform. The objective in this use case is to continuously monitor the HF concentration raw gas, and feed primary alumina accordingly, hence even out the adsorbed HF. As adsorption of HF on alumina is a very temperature dependant process, temperature control is essential. This can lead to increased cell efficiency, reduced environmental impact and energy requirements of the GTC by varying degrees depending up on the production capacity of the plant [15].

A digital twin of the GTC simulates the gas adsorption phenomenon either using physics based or empirical models. However, individually neither of these approaches truly mimic the underlying processes and therefore combining the two approaches to give hybrid models to explain the process could provide better and more accurate models of the GTC. The first principles models consist of complex differential equations [16] such as kinetic, mass, and energy balance of the adsorption process. The data-based models contain various ML and AI algorithms where the dependent variable would be the HF gas concentration in the raw gas, to be adsorbed by alumina in the GTC. Thus, both DT & HT will provide solutions for optimal quantity of fresh alumina to be used to even the out the HF adsorbed to alumina. The CT of the GTC recommends optimal operating parameters for adsorption based on real-time data gathered about conditions such as the pressure, temperature, humidity, etc., from the sensors. All of these process variables are highly dependent on the external factors that cannot always be controlled such as the surface area of the alumina and ambient conditions (weather). Thus, the HT & CT provide plant operators best suggestions that adapt and evolve continuously.

B. Minimise health & safety risks and maximise the metallic yield in Silicon (Si) production

Silicon metal finds applications in a wide range of industries from electronics and related components, adhesives, coatings and others. With the metal being used in so many different industries, it should not be surprising the production of the metal has been increasing continuously over the past several years. In the year 2016, approx. 7 million tonnes of silicon was produced worldwide [14]. Silicon is typically produced by melting silicon ore in electric arc furnaces and expectedly the furnaces reach very high temperatures of 2000 degree Celsius and above. After the ore has been melted, the molten silicon from the electric furnaces are tapped into ladles. The ladles are subsequently transported to a separate location for refining and alloying of the liquid Si. Finally, the liquid Si-alloy is cast on a pig caster. During the tapping sequence, visual observations are next to impossible due to smoke and flames. Operators still need to execute tap-hole maintenance, as well as sampling of temperature and chemistry from the ladle. In many plants these procedures are performed manually, which presents severe health and safety risks for the workers. Through the COGNITWIN toolbox, we aim to exhibit a scheme that will enable remote operations of the tapping process along with development of sensors that will allow monitoring the silicon melt during the tapping process without manual intervention. Furthermore, a detailed process model for tapping, refining, and alloying relies on a computational approach combined with existing process data as well as new input from novel sensors. We estimate that the post tap-hole silicon yield can be increased, while reducing the

Authorized licensed use limited to: Malardalen University. Downloaded on August 07,2021 at 16:28:40 UTC from IEEE Xplore. Restrictions apply.

overall energy consumption and reduced variation in the end product quality.

A DT that represents the refining/alloying process during post tap-hole operations is generally described by the usual mass, energy, and momentum conservation equations from computational fluid dynamics (CFD). A HT simulates the ladle characteristics by combining the previously mentioned mathematical equations along with empirical models that are developed using various ML algorithms applied to images generated by thermal cameras. A CT of the ladle would include the same models in HT; however, the equations are parameterised CFD equations typically using optimisation routines such as genetic algorithms or particle swarm optimisations for improved accuracy [17]. Combining these representations of the ladle can provide best estimates of when the furnace can be emptied to the ladle for further operations.

C. Condition monitoring of assets in steel and related products production

Just as is the case with many other metals production processes, the first step in steel (recycled) production is melting the scrap steel in an electric arc furnace post, which the molten steel is then tapped into ladles for secondary metallurgy process. In the second stage, the melt is refined and mixed with other substance to produce the required grade of steel. In the final step, the refined molten steel with customer-specific requirements is poured on to a tundish for casting. In addition to the metal flowing through the system, gases and alloy materials are added and off-gases, dust and some slag may be created. The specific aim concerning this challenge relates to the deteriorating refractory bricks in the ladles with each batch of secondary metallurgy process. If refractory lining in the ladle is not sufficiently thick, they run the risk of having a leakage/spillage of the molten steel in the production plant that can cause severe health hazards and production shutdown. To ensure the safe operation of the ladles, and to avoid such a scenario, a technician looks for damages or the health of the refractory lining after each heat. The technician makes the decision for full re-lining of the ladle keeping safety as a priority and works on the “better safe than sorry” philosophy. This however has a drawback, the technician might decide for a full refractory brick replacement even when this is not actually needed. Should this happen more often, it can result in increased production costs for plant can reduced efficiency.

The DT in this case is a mathematical representation of the ladle during each steel refining process that focuses on the refractory brick behaviour. Physics-based models account for refractory wear due to thermal and mechanical stresses on the walls whereas ML models are developed using process parameters to estimate the conditions of the refractory bricks. Combining these two approaches gives us the HT, which can possibly provide reliable estimates of brick's conditions. A CT in this specific case helps the engineers and technicians to make better judgements about the brick health based on the historical data and the brick demolition/repair routines followed.

A closely related application to the above is maintenance of equipment and machinery that handle steel and related products. For example, welded steel pipes are used in several businesses such as oil & gas, water supply and piling. It is typical to use large machines to weld steel pipes especially when the plant is catering to several and diverse customers. However, one major problem when using the spiral welded steel pipes machinery (SWP), a highly sophisticated machinery is the possibility of it breaking down unexpectedly often due to failure of a small component in the system. Therefore, by having a system that predicts the maintenance requirements of the SWP, greatly reduce the downtime and the production time. A typical SWP has several hundreds of moving, vibrating and rotating parts that can breakdown due to mechanical or thermal stress. A HT in this case is represented by mathematical models that simulate for instance the vibration patterns and/or the mechanical stress experienced by the components in the machinery. In addition to this real-time data gathered from various sensors can be used to develop empirical models to better predict the failure mechanisms of the SWP. The CT of the SWP predicts the failure probabilities of the components by monitoring smallest of changes in physical characteristics of the moving parts such as changes in the angular moment or rotatory parts or vibration frequencies of moving parts.

D. Real-time monitoring of finished products for operational efficiency

As is seen in many processes from the process industry, the conditions in steel/metal production sites and rolling mills can be very harsh. Tracking goods and assets under such extreme environmental influences in non-FIFO situations can be challenging. In particular, if goods or assets are heated, RFID-based sensors are no valid option. In addition, if goods are being deformed in the process, physical markings such as stamps are rendered useless as well. However, tracking often is desirable for operational efficiency and in terms of quality management to link data collected at different production steps to the goods or asset.

One solution to this problem is to develop image and object recognition deep learning models. These computer vision-based ML systems are able to track goods or assets throughout the non-FIFO part of the production process.

Based on the DT of the production process section provided by the tracking system, we envision the HT to combine the tracking model with sensor data collected in the production process for example to predict error-prone situations before they occur and allow an operator or operational system to react accordingly, but also to allow for a real-time adaption of process parameters to enable advanced quality management based on predictive model outputs or the linked data from the entire production process directly. A CT of such a system would feature all attributes of a HT together with an innate ability to react on its own to situations requiring an intervention, thus stabilizing the production process further.

E. Improving heat exchanger efficiency

The scaling and fouling phenomenon in the heat exchanger tubes has been a big problem to several industries e.g. power production, chemicals, and food processing industries resulting in significantly reduced operational efficiency. This problem is experienced mostly by power generation companies or municipalities who produce electricity either by waste incineration or by using biofuels in a circulated fluidized bed boiler unit. While it is accepted that in most cases scaling and fouling is given and cannot be fully prevented, there are measure heat exchanger operators can take that will minimise the extent of the fouling phenomenon. With improved heat exchanger maintenance, it is expected

Authorized licensed use limited to: Malardalen University. Downloaded on August 07,2021 at 16:28:40 UTC from IEEE Xplore. Restrictions apply.

that the operating costs of the plant can be reduced significantly while also reducing environmental emissions.

A HT for this use case is represented by mathematical and ML models to take actions to remove the deposits from the heat-exchanger tubes before the operators may have to stop their process and take the equipment out for servicing. The CT of the heat exchanger will predict the deposition of unburnt fuel mixtures, ash and other particles on the heat-exchanger tubes based on both historical practices and real-time process data to suggest clean-up of the tubes in the heat exchanger before the deposits build up more strongly.

VI. Conclusion and Future Work

A DT is more than a container for integrating data and different models. Having a comprehensive understanding of a system requires combining its models. Additionally, there are many situations where the physical system cannot be properly understood by its models, e.g., due to over-simplification of complex situations. To be able to effectively deal with unforeseen situations, additional knowledge (expert, domain, problem solving, etc.) should be exploited.

In this paper, we proposed a conceptual approach to enhance functionalities of Digital Twins by introducing the notions of Hybrid and Cognitive Twins and by describing a toolbox-based solution to realize this vision. Although the proposed approach is general enough to be applied in different domains, in this paper our focus is on the process industry. We identified a set of relevant use cases in the process industry and discussed the needs for and roles of different types of twins as well as the possible use of the toolbox.

Furthermore, we identified the challenges for which we intend to implement solutions in the future. Indeed, our future work will focus on the realization of twins as sketched in Section IV as well as on the management of different levels of twins (DTs, HTs, CTs). Our intension is to conceptualize and develop methods and tools that support the whole lifecycle of a twin, starting from modelling, over generation, connection, operation, reuse, sharing to continuous improvement. We aim to implement a toolbox that is generic and flexible in its design and at the same time comprehensive in its functionality. It should easily connect to and build upon existing DT frameworks and thus forms the basis to be used by a wide range of organizations, even by those who already have some proprietary implementation of DTs.

Acknowledgment

The work in this paper is partly funded by the H2020 project COGNITWIN (grant number 870130). We thank the COGNITWIN consortium partners for fruitful discussions related to HT/CT concepts and the use cases.

References

[1] IBM Research, "The future of health is cognitive," 2016. [Online]. Available: https://www.ibm.com/downloads/cas/LQZ0O1WM.

[2] F. Chinesta, E. Cueto, E. Abisset-Chavanne, J. L. Duval and F. El Khaldi, "Virtual, digital and hybrid twins: a new paradigm in data-based engineering and engineered data," in Archives of Computational Methods in Engineering, 2018, pp. 1-30.

[3] A. Rasheed, S. Omer and K. Trond, "Digital twin. Values, challenges and enablers from a modeling perspective," in IEEE Access 8, 2020, pp. 21980-22012.

[4] A. J. H. Redelinghuys, A. H. Basson and K. Kruder, "A six-layer architecture for the Digital Twin: a manufacturing case study implementation," in Journal of Intelligent Manufacturing, 2019, pp. 1-20.

[5] J. Lu, X. Zheng, A. Gharaei, K. Kalaboukas and D. Kiritsis, "Cognitive Twins for supporting decision-makings of the Internet of Things systems," in Proceedings of 5th International Conference on the Industry 4.0 Model for Advanced Manufacturing, 2020, pp. 105-115.

[6] Schluse, Michael et al., "Expertimentable digital twins - Streamlining simulation-based systems engineering for industry 4.0," in IEEE Transactions on industrial informatics 14.1, 2018, pp. 1722-1731.

[7] D. Hartmann and H. Van der Auweraer, "Digital Twins (in press)," in arXiv preprint, arXiv:2001.09747, 2020.

[8] J. M. Gómez-Berbís and A. de Amerscua-Seco, "SEDIT: Semantic Digital Twin based on industrial data management and knowledge graphs," in International Conference on Technologies and Innovation, Cham, Springer, 2019.

[9] Erich Barnstedt et al., "Details of the Asset: Part 1 - The exchange of information between partners in the value chain of Industrie 4.0 (Version 1.0)," Federal Ministry for Economic Affairs and Energy (BMWi), 2018.

[10] Erich Barnstedt et al., "Details of the Asset: Part 1 - The exchange of information between partners in the value chain of Industrie 4.0 (Version 2.0)," Federal Ministry for Economic Affairs and Energy (BMWi), 2019.

[11] Industrial Internet Consortium, "Digital Twins for Industrial Applications," 2020.

[12] BDV SRIA, "European big data value - Strategic research and innovation agenda," 10 2017. [Online]. Available: http://bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf. [Accessed 12 06 2020].

[13] BDVA and euRobotics, "Joint Vision Paper for an artificial intelligence public private partnership (AI PPP)," 03 2019. [Online]. Available: https://www.eu-robotics.net/cms/upload/downloads/VISION_AI-PPP_euRobotics-BDVA-Final.pdf. [Accessed 12 06 2020].

[14] N. LePan, "All the World's Metals and Minerals in One Visualization," Visual Capitalist, 01 03 2020. [Online]. Available: https://www.visualcapitalist.com/all-the-worlds-metals-and-minerals-in-one-visualization/. [Accessed 12 06 2020].

[15] T. A. Aarhaug and A. P. Ratvik, "Aluminium Primary Production Off-Gas Composition and Emissions: An Overview," JOM, pp. 2966-2977, 2019.

[16] Y. Liu and L. Shen, "From Langmuir kinetics to first-and second-order rate equations for adsorption," Langmuir, pp. 11625-11630, 2008.

[17] M. Raissi, A. Yazdani and G. E. Karniadakis, "Hidden fluid mechanics: A Navier-Stokes informed deep learning framework for assimilating flow visualization data," arXiv preprint arXiv:1808.04327, 2018.

[18] R. Herzog, M. Jacoby and I. Podnar Žarko, "Semantic interoperability in IoT-based automation infrastructures," at-Automatisierungstechnik, vol. 64, no. 9, pp. 742-749, 2016.

[19] M. Jacoby, A. Antonić, K. Kreiner, R. Łapacz and J. Pielorz, "Semantic interoperability as key to iot platform federation," International Workshop on Interoperability and Open-Source Solutions, pp. 3-19, 2016.

[20] P. Schlkwyk, S. Malakuti and S.-W. Lin, "A short introduction to Digital Twins," 2019.

Authorized licensed use limited to: Malardalen University. Downloaded on August 07,2021 at 16:28:40 UTC from IEEE Xplore. Restrictions apply.