

# A Generic DigitalTwin Model for Artificial Intelligence Applications

1<sup>st</sup> Oliver Niggemann

*Institute of Automation Technology,  
Helmut Schmidt University, Hamburg, Germany*  
oliver.niggemann@hsu-hh.de

2<sup>nd</sup> Alexander Diedrich

*Fraunhofer IOSB, Fraunhofer Center  
for Machine Learning, Lemgo, Germany*  
alexander.diedrich@iosb-ina.fraunhofer.de

3<sup>rd</sup> Christian Kühnert

*Fraunhofer IOSB, Fraunhofer Center,  
for Machine Learning, Karlsruhe, Germany*  
christian.kuehnert@iosb-ina.fraunhofer.de

4<sup>th</sup> Erik Pfannstiel

*Institute of Automation Technology,  
Helmut Schmidt University, Hamburg, Germany*  
erik.pfannstiel@hsu-hh.de

5<sup>th</sup> Joshua Schraven

*Institute of Automation Technology,  
Helmut Schmidt University, Hamburg, Germany*  
jushua.schraven@hsu-hh.de

**Abstract**—Services for Cyber-Physical Systems based on Artificial Intelligence and Machine Learning require a virtual representation of the physical. To reduce modeling efforts and to synchronize results, for each system, a common and unique virtual representation used by all services during the whole system life-cycle is needed—i.e. a DigitalTwin. In this paper such a DigitalTwin, namely the AI reference model AITwin, is defined. This reference model is verified by using a running example from process industry and by analyzing the work done in recent projects.

## I. INTRODUCTION

DigitalTwins are a key concept for Cyber-Physical-Systems (CPS): By maintaining a collection of relevant information about a physical entity, a digital shadow is created which can be used for tasks such as monitoring, diagnosis or optimization. Most publications about the DigitalTwin focus on engineering- and process-oriented aspects such as a continuous enrichment of the twin during its life-cycle [30], on simulation scenarios [28], [20] or on modeling questions such as optimal meta-levels [29], hierarchies [30] or engineering chains [20]. Even the few publications written with a clear focus on Artificial Intelligence (AI) and Machine Learning (ML) fail to connect the content and capabilities of the DigitalTwin with AI/ML methods. A review of the related work confirming this impression is given in section II. Under the bottom-line, the DigitalTwin is seen as an information sink into which all available information is poured—hoping for a benefit by AI/ML applications at some later point in the life-cycle.

On the other hand, AI/ML methods have always used models of the environment and domain knowledge. So the DigitalTwin concept and AI/ML should be compatible. But two main contradictions remain: First, AI/ML are very heterogeneous, and each AI/ML method comes with a specialized

model formalism to capture relevant aspects of the environment and the application domain. Hence, the question is how a DigitalTwin can provide the correct model to each AI/ML method. The second contradiction is that AI/ML requires explicit, i.e. by an algorithm processable, knowledge since compiled knowledge in form of simulation libraries, raw data or executables does not help. But most publications refer to these kind of information.

In this paper we follow the key idea in Edward E. Lee’s seminal paper [16] in which he claims that “The intellectual heart of CPS is in studying the joint dynamics of physical processes, software, and networks.”, i.e. from an AI/ML perspective the main task of a DigitalTwin is to use models to predict the behavior of the corresponding CPS subsystem, comprising cyber and physical parts. Here we add an additional constraint: The prediction knowledge must be explicit and processable by an AI/ML algorithm.

So the paper addresses the following research questions (RQs):

*RQ 1:* Can we develop, as shown in figure 1, a common API of a DigitalTwin which is suitable for a set of heterogeneous AI algorithms? I.e. here in this paper neither the AI algorithms nor the internal structure of the DigitalTwin is addressed but only a common DigitalTwin API.

*RQ 2:* Can AI algorithms use a common information base, stored in the DigitalTwin? Can AI algorithms even exchange information that way?

*RQ 3:* Can such a DigitalTwin be used for both machine learning algorithms (i.e. sub-symbolic AI) and symbolic AI algorithms? Bridging sub-symbolic and symbolic AI is currently a major topic in the AI community [15], [3].

In section III, an analysis of the requirements of AI/ML methods to the corresponding models is made and in section

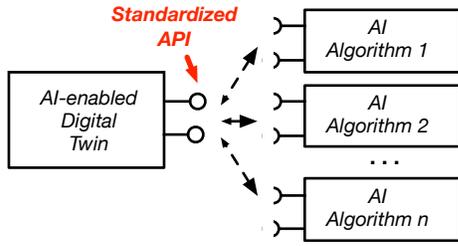


Fig. 1. Research Questions 1: Common AI-enabled API for DigitalTwin.

IV the main solution approach is presented. As a result, a DigitalTwin AI reference model, called the **AITwin**, will be derived. If such a reference model is used by several AI/ML methods, multiple advantages arise: First of all, the results of one method could be used by another. Second, redundant information used by different methods are avoided, easing the modeling efforts and preventing contradicting behavior. Case studies can be found in section V, they are used to evaluate the solution approach. A conclusion is given in section VI.

## II. RELATED WORK

As described in the introduction, the DigitalTwin with its enabling technologies plays a decisive role for the success in several areas and is an ongoing field of research. One of the most comprehensive reviews for the current-state-of-the-art has been made by Fuller [13] covering the areas manufacturing, smart cities and healthcare. As one of his conclusions he states that there is still the lack of DigitalTwin reference models, which leads to discrepancies between similar projects and hence slows down the progress of this technology. Proposals for DigitalTwin reference models have already been made in [4] with the focus on risk control and prevention on process plants, in [18] for smart-manufacturing and in [2] for cloud-based CPS.

Recent publications, taking into account DigitalTwins and AI/ML are given by Min [22] focussing on production optimization in petrochemical industry or [10] who proposes a machine-learning enhanced DigitalTwin for Human-Robot-Collaboration. Still, most of the research is dealing with DigitalTwins from a modeling perspective. For example [19] develops the DigitalTwin of a CNC-machine, in [34] a aircraft has been reengineered and [26] explains the challenges and enablers for DigitalTwin in modeling.

Hence, current research combining AI with DigitalTwins results always in individual solutions for specific application fields. The herein presented work tries to close this gap by proposing the DigitalTwin AI reference model AITwin.

## III. AI/ML METHODS AND THE DIGITALTWIN

Here we categorize AI/ML methods and derive requirements to the DigitalTwin. The number of used AI/ML methods are legion and so are the formalisms used to model domain knowledge and environment models. Still for simplicity sake, we focus on those methods mainly used for CPS and categorize

these AI/ML methods into three main categories shown on the left hand side of figure 2. To illustrate these applications, a running example is used in this paper (figure 3):

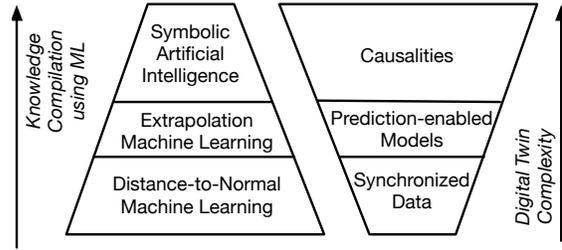


Fig. 2. Levels of AI/ML and the matching complexity of digital twins.

The system consists of four water tanks  $t$ , seven electric valves  $p$  with integrated flow sensors, an unlimited water source and an unlimited water sink (not shown). Valve  $v_0$  controls water from the unlimited water source, for example the public water mains, into tank  $t_0$ . From there, three pipes with an equal diameter divide the water flow. Finally, valve  $v_6$  drains tank  $t_3$  into the unlimited water sink, for example a river or a processing facility.

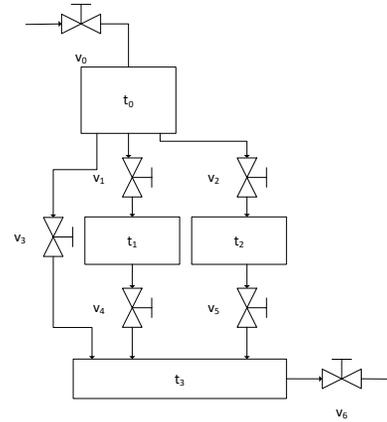


Fig. 3. The running example in the form of a four tank model.

**Distance-to-Normal Machine Learning:** Anomaly detection is a typical example for this approach: The machine learner memorizes historical data points and computes a distance measure between a current data point (e.g. production operation point) and memorized points. Often, the historical data points are interpreted as “normal” and too distant data points are classified as “anomalous”. To some small extent, interpolation is used to deal with data points close to previous data points. To implement “Distance-to-Normal Machine Learning”, the Digital Twin (see right hand side of figure 2) must comprise a common timing model to synchronize all data.

In our running example, the vector of all flows can be used to detect unusual flows, i.e. anomalies.

**Extrapolation Machine Learning:** Using a DigitalTwin in autonomous systems for optimization and self-configuration

means using it in a closed control loop, e.g. for model-predictive control tasks. I.e. the DigitalTwin must predict output values such as resource consumption, positions, etc. for new input values. Please notice that this extrapolation capability means to predict values, in time or distance, far away from observed data points. In terms of ML this is especially a problem for technical systems where often only few individual data points are available which are also clustered in only a few regions.

In our running example, the vector of all flows can be predicted (i.e. extrapolated) for new valve positions. This can be used to optimize the system configuration.

**Symbolic Artificial Intelligence:** Many AI tasks such as diagnosis, reconfiguration or planning require causal information, i.e. symbolic AI models: Diagnosis uses root cause-symptom relations (“if the tube leaks, the tank fills slowly”) and planning needs information about process steps (“Modul 1 takes a board and drills a hole”). Such causalities always describe effects resulting from causes where both effects and causes are symbolic predicates in some formal logic. So the Digital Twin (see right hand side of figure 2) must comprise such logic-based models.

In our running example, the effect of switching a valve such as an overflowing container can be modeled, e.g. to identify root causes of errors.

#### IV. AI REFERENCE MODEL

In this section, the three research questions from section I are discussed and, as a result, the DigitalTwin AI reference model, **AITwin**, is derived.

##### A. Timing Model

While time is the essential feature of any physical process (the “P” in CPS), software (the “C” in CPS) does not come with any build-in concept of time. In software parallel events may occur at various points in time and time spans between commands are not defined but vary according to process behavior. So first, a common time model of the DigitalTwin must be established. Following an idea from [16], one may differentiate between two general solutions for this:

*Cyberizing the Physical:* Normally such approaches define signal values  $x$  as  $x(t, k)$  where  $t \in \mathbf{R}$  is the physical time and  $k \in \mathbf{N}$  the number of an ordered sequence of events/instructions in the software. This is called a superdense model of time.

*Physicalizing the Cyber:* New means for the time synchronization in distributed systems, e.g. the Precision Time Protocol [14], allow for a consistent concept of physical time especially in controllers. Hence, software now may refer to precise points in time and time spans.

“Cyberizing the Physical” has the drawback that most AI algorithms need to be adapted, since they require a clear time model and concept of causality. Therefore, the usage of “Physicalizing the Cyber” approach should be preferred, i.e. all information  $\mathbf{x}(t)$  in the DigitalTwin refers to a unique point

in time  $t$ . This requires from the underlying technical systems time synchronization between all devices and an implemented concept of simultaneity.

So on this level only two interfaces are needed for the DigitalTwin AI reference model (AITwin): Function  $getData(i, t)$  returns the  $i$ 'th signal  $x \in \mathbf{R}$  at a specific point in time  $t$  and function  $getData(t)$  returns all signals  $\mathbf{x} \in \mathcal{X}$  at a specific point in time  $t$ . We refer with  $\mathcal{X}$  to the set of all possible values  $\mathbf{x}(t)$ .

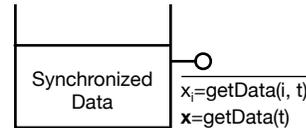


Fig. 4. Interfaces for the first level of the DigitalTwin AI reference model AITwin.

Please note that the vector  $\mathbf{x}(t)$  captures the complete observable state of the system, e.g. comprising sensor/actuator signals, observable product/object locations and states, configuration parameters and so on.

##### B. Extrapolatable Models

While superficially much data seems to be generated by Cyber Physical Systems, a different picture can be seen when ML is applied: Data intensive ML such as deep learning often fails or delivers unusable results in production settings while in other domains it shows break-thru performance. This is mainly due to the following two reasons: (1) Most data comprises repetitive patterns –mirroring the repetitive structure of typical production processes, while the pure amount of data is sufficient. (2) The data hardly comprises information about failures or other problems.

Therefore, learned models often only have limited extrapolation capabilities—which is essential for CPS. Furthermore, even if learned models such as neural networks predict output values for new input values, the prediction quality quickly decreases for values far away from observed data points in time and distance. Hence, the AI reference model, AITwin, must not only be able to extrapolate but also to assess the quality of the extrapolation. So far the best approach is the usage of probabilistic models, e.g. probabilistic or Bayesian Neural Networks [23] which can also be applied to time series analysis [12]. So all predictions should not only be implemented as a point estimator but should also compute the probability that the prediction is correct. In general, we have to differentiate between two types of analyses and predictions:

**Static Predictions:** For tasks such as condition-monitoring or anomaly detection, only the signal values  $\mathbf{x}(t) \in \mathbf{R}^n$  at some point in time  $t$  are used. In other words, for the analysis a static feature vector [11] is taken into account. Thereby, the assumption is that no information is coded in the sequence of value and all necessary information is contained in the current signal values. This assumption is true for many CPSs, even for systems which have a dynamic nature them-self. Typically

three use cases are implemented here by machine learning applications:

*Anomaly Detection:* For a new data point  $\mathbf{x} \in \mathbf{R}^n$  its probability  $p(\mathbf{x}|X)$  given some historical data  $X$  is (at least approximately) computed.

*Prediction:* Given a partial feature vector  $\mathbf{x}' \in \mathbf{R}^n$ , i.e. a feature vector with missing entries, those entries are computed in form of a complete feature vector  $\mathbf{x}$ , best with their probability (e.h. to handle ambiguous predictions). The given entries then become inputs and the computed entries become outputs. Of course, if some input to output configuration is not possible, this must be signaled to the user.

*Optimization:* The reader may note that the partial gradient over the signals can be created by extrapolating in the vicinity of an operation point, so no special functionality of the DigitalTwin is needed here for the optimization tasks.

**Dynamic Predictions:** A totally different situation arises when important information is coded in the sequence of signal values over time. Again, three use cases exist:

*Anomaly Detection:* Given historical data  $X = \{\mathbf{x}(t = t_0), \mathbf{x}(t = t_1), \dots, \mathbf{x}(t = t_k)\}$ , for a time step  $\Delta t$  and a new data vector  $\mathbf{x}(t = t_k + \Delta t)$  and its probability is computed so that  $\mathbf{x}(t = t_k + \Delta t)$  is the logical continuation of the series.

*Prediction:* Given historical data  $X = \{\mathbf{x}(t = t_0), \mathbf{x}(t = t_1), \dots, \mathbf{x}(t = t_k)\}$  and a time step  $\Delta t$  the next value  $\mathbf{x}(t = t_k + \Delta t)$  is computed, best with its probability distribution.

*Optimization:* The reader may note that the temporal gradient over the signals can be created by extrapolating in the vicinity of the point in time  $\Delta t$ , so no special functionality of the DigitalTwin is needed here.

As mentioned earlier, for reliable predictions, the model must not only capture the behavior in the normal but also in the failure modes. For example predictions for a failure mode, like a broken drive or a blocked pipe, are helpful for diagnosis tasks. Hence, the DigitalTwin must be able to activate specific failure modes by setting some components to failure states.

Figure 6 shows the resulting API methods with extrapolation capabilities. The static case is solved by a function *extrapolateStatic(x')* where  $\mathbf{x}'$  is a partially filled signal vector. The function returns the missing values  $\mathbf{x}$  and a corresponding vector of probabilities, both with *null* entries if not computable. This solves all three use cases discussed above.

For the dynamic case, a function *extrapolateDynamic(X, Δt)* where  $\mathbf{X}$  is a sequence of historical signal vectors up to the point in time  $t$  is defined. The function returns the next signal vector  $\mathbf{x}(t + \Delta t)$  and a corresponding vector of probabilities, both with *null* entries if not computable. Again, this solves the three use case discussed above.

The failure modes are set by a function *setFailedComps(C')* where  $\mathbf{C}'$  is a subset of all components. If a failure mode is set, the functions *extrapolateStatic* and *extrapolateDynamic* predict the behavior assuming that the components are not working correctly. Another function *getComps()* returns all components.

### C. Causalities

No definition of causality exists which is accepted by the majority of researchers. We therefore first analyze the special usage of causality models in two main AI application fields of DigitalTwin and derive from there a suitable causality definition.

AI Application 1 - Consistency-based Diagnosis: Diagnosis is the task of computing the error cause based on incomplete symptoms, i.e. observations. Therefore, knowledge about the causalities between root causes (i.e. failures) and symptoms must be known since failures such as sensor faults or machine degradation often lead to complex symptom patterns, e.g. alarms, error messages and plant stops. Currently users often see complex alarms and symptom patterns but fail to identify quickly the correct root cause and therefore can not repair the factory efficiently.

Often models predict the behavior only for the normal mode. Consistency-based diagnosis [8] is able to use so-called weak fault models, i.e. models which capture the system behavior only for the OK-mode. In most cases, it uses partial system models such as  $OK(C_{i_1}) \wedge \dots \wedge OK(C_{i_k}) \rightarrow (s_1 \rightarrow s_2 \wedge \dots \wedge s_l), k, l \in \mathbf{N}$  where  $OK(C)$  denotes a correct functioning of component  $C$  and  $s_i \in \{0, 1\}$  is a binary/symbolic representation of the current system status, e.g.  $s_i \equiv$  "Sensor Value 42 > 10" or  $s_j \equiv$  "Motor 3 is on". So normally  $s_i \in S$  where  $S$  describes all system states. In other words, if components  $C_{i_1} \dots C_{i_k}$  work correctly, the implication describes the normal system causality. Consistency-based Diagnosis then redraws OK-assumptions until no contradictions between predictions and observation exist anymore.

AI Application 2 - Planning: Planning is the task of computing a sequence of process steps which transforms given raw objects into an also given final object. Let  $P$  be the set of all possible objects. Then a process step  $q_i$  is an implication  $q_i : \mathbf{x} \rightarrow \mathbf{y}$ , which takes  $u$  input objects  $\mathbf{x} \in P^u$  and transforms them into  $v$  output objects  $\mathbf{y} \in P^v$ . We denote the set of all process steps as  $Q$ . Please note that parameters of process steps can be implemented by defining several individual process steps.

Based on these application scenarios, a common causality model for the AI reference model **AITwin** is derived. Figure 5 shows two types of causalities, namely System State Causalities (handling system changes) and Product State Causalities (handling product changes) are defined:

**System State Causalities:** Such causalities model the transition from a system state  $s_i \in S$  to state  $s_j \in S$ . The occurrence of an event  $e_1$  triggers, denoted by Pearl's **do** operator [24], this transition. Please note that also empty events, i.e. immediate transitions without external triggering events, are possible.

A system state  $s \in S$  corresponds to a subset of  $\mathcal{X}$ , with  $\mathcal{X}$  being the set of all possible values  $\mathbf{x}(t)$ . Here we restrict this to convex subsets of  $\mathcal{X}$ , corresponding to linear inequalities on  $\mathbf{x}(t)$ . I.e. a system state is defined as a subset of the values.

An event  $e_1$  is then defined as the crossing of a threshold in  $\mathcal{X}$ . Such thresholds correspond to equations of the form  $\mathbf{f} \cdot \mathbf{x}(t) < \mathbf{c}, \mathbf{f}, \mathbf{x}(t), \mathbf{c} \in \mathbf{R}^n$  which define a halfspace of  $\mathcal{X}$ . Such a system state causality is only valid under the condition that specific components  $C_1, \dots, C_l$  are functioning correctly, e.g.  $OK(C_1) \wedge \dots \wedge OK(C_l)$  is true (see section IV-B). Please note that with these information the diagnosis task from above can be implemented, e.g. consistencies between predictions  $\mathbf{x}(t)$  from the predictions API level and causalities predicting system states  $s_i$  can be checked as  $s_i$  is a subspace of  $\mathcal{X}$ .

**Product State Causalities:** Such causalities model the transition from one set of objects  $p_i, \dots, p_j$  to another set  $p_k, \dots, p_l$ . A product  $p_i$  might be any product, partial product or raw material, e.g. a chemical, a half-finished car or a complete car.

Again, the occurrence of an event  $e_2$  triggers this transition and empty events are possible. Events are defined as above. Such a system state causality is only valid under the condition that specific components  $C'_1, \dots, C'_l$ , see section IV-B, are functioning correctly, e.g.  $OK(C'_1) \wedge \dots \wedge OK(C'_l)$  is true. The resulting API extensions are shown in figure 6. Please note that with these information the planning task from above can be implemented.

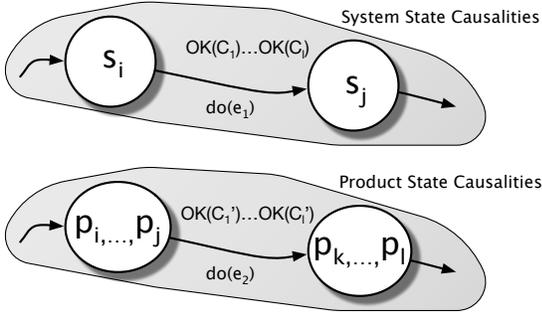


Fig. 5. The main causality concept of the DigitalTwin AI reference model AITwin,  $s_i$ s are system states,  $C_j$  are components,  $p_k$  are products.

An implication API allows for the handling of causalities in form of transitions or implications: The function  $getSystemCausalities(s_i)$  returns all causalities  $s_1, e, s_2, I, n, OK$ .  $s_1$  denotes the starting state,  $s_2$  the end state,  $e$  is the event,  $n$  the name of the causality and  $I$  stores additional information such as probabilities, timing etc.  $OK$  is the set of  $OK$  assumptions  $OK = OK(C_1) \wedge \dots \wedge OK(C_l)$  which are a precondition for the validity of the causality.

The function  $getProductCausalities(p_i)$  returns all causalities  $p_1 = (p_1^1, \dots, p_1^l), e, p_2 = (p_2^1, \dots, p_2^k), I, n, OK$  where  $p_1$  denotes the original objects,  $e$  is the event,  $p_2$  denotes the resulting objects,  $n, I, OK$  are defined as above.

$e = defineEvent(halfspace)$  defines an event  $e$  which corresponds to a halfspace. The parameter  $halfspace$  is defined using a mathematical inequality in the language MathML [21]. The function  $getEvent(\mathbf{x}, \mathbf{x}')$  returns all events which are triggered when moving from the operation point  $\mathbf{x} \in \mathcal{X}$  to  $\mathbf{x}' \in \mathcal{X}$ .

$s = defineConcept(region)$  defines a system state  $s$ , namely a concept, as a convex region  $region$ . The parameter  $region$  is part of in  $\mathcal{X}$  using a mathematical inequality in the language MathML [21]. The function  $getConcepts(\mathbf{x})$  returns all system states  $s_i$  (aka concepts) with  $\mathbf{x} \in s_i$ .

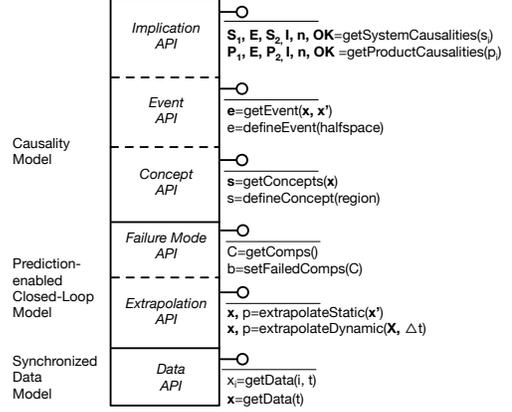


Fig. 6. The complete DigitalTwin AI reference model AITwin.

#### D. The AITwin Reference Model

Figure 6 shows the complete DigitalTwin AI reference model AITwin. Please note that the focus is on those parts of a DigitalTwin which are required by AI algorithms. Further information for domains such as production, automation and logistics must be added. AITwin may be used in parallel on different system granularities, e.g. for machine modules in production plants or for production lines.

Please note that the three levels of figure 2 correspond to the three levels of figure 6 and form a holistic concept: Data  $\mathbf{x}$  from the "Synchronized Data Model"-API is used also in the "Prediction-enabled Closed-Loop Model"-API. And these data points can also be correlated to system states from that "Causality Model"-API. And all APIs use the same components for modeling failure modes.

Besides those APIs, the DigitalTwin is based on three state spaces which must be part of the DigitalTwin definition:  $\mathcal{X}$  denotes all possible signal values,  $P$  denotes all possible object/product configurations and  $COMPS$  is the set of components which may fail. Furthermore, two derived state spaces exist, which normally are defined via API methods during the operation phase:  $S$  is a set of system states where  $s \in S$  is a subset of  $\mathcal{X}$  and  $E$  is the set of all events where  $e \in E$  is a halfspace of  $\mathcal{X}$ .

Please note that the APIs here do not cover the way how the information are added to the DigitalTwin.

#### V. EVALUATION

Two ways exist to evaluate such a new API. In section V-A, AITwin is used by several AI algorithms for one running example. In section V-B, previous research projects in the field of AI in production are analyzed whether AITwin would have been suitable.

Config.	Fault	LSTM Autoen.		SVM	
		AUC	F1	AUC	F1
a_tank	stable	1.000	1.000	1.000	1.000
a_tank	pumpSlow	0.500	0.406	0.997	0.997
a_tank	tank1_leak	0.500	0.402	0.994	0.994
a_tank	valve0Block	0.989	0.977	0.999	0.999
a_tank	valve1Block	0.984	0.950	0.990	0.970
a_tank	valve1Stuck	0.989	0.977	0.500	0.666
4_tanks	stable	1.000	1.000	1.000	1.000
4_tanks	pipe4_jam	0.996	0.996	0.992	0.992
4_tanks	tank2_leak	0.960	0.959	0.992	0.992
4_tanks	valve3_jam	0.996	0.996	0.983	0.983
4_tanks	valve6_jam	0.996	0.996	0.701	0.770
3_tanks	stable	1.000	1.000	1.000	1.000
3_tanks	pumpFast	0.915	0.833	0.727	0.588
3_tanks	pumpSlow	0.993	0.990	0.929	0.562
3_tanks	tank1Leak	0.994	0.992	0.517	0.554
3_tanks	tank2Leak	0.994	0.992	0.959	0.960
3_tanks	valve2Closed	0.988	0.978	0.985	0.895
3_tanks	valve3Closed	0.986	0.975	0.505	0.057

TABLE I  
RESULTS OF THE MACHINE LEARNING METHODS

### A. Applying AITwin to a Running Example

In a next step, the "synchronized data" and especially the "extrapolation" APIs have been used to detect anomalies. Anomalies can be inserted via the "failure mode" API.

For this, first data is extracted via the "data" API and used to train a model of the normal behavior via unsupervised learning. Later, non-normal data is fed to the models and the models are supposed to detect these anomalies. As learning algorithms, Support Vector Machines (SVM) and Long-Short-Term-Memory (LSTM) neural network are used. Table I shows the results for different systems and different faults. Both Area-Under-Curve (AUC) and also F1 measures show that anomalies are detected with a high probability.

In a second step, the AITwin is applied to symbolic AI tasks. For this, concepts are extracted from the DigitalTwin via the "concept" API and causalities via the "implication" API. This information can be used to create a model for Consistency-Based Diagnoses (see section IV-C). Failures can be added via the "failure" API. Table shows the results, most failures can be identified correctly.

We also used a quantitative simulation of the Tennessee Eastman process. The implementation of Downs et al. [9] contains 20 different injected faults (process disturbances). However, the instrumentation of the simulated process is such that not all faults will be identified exactly. Table II shows the results for six experiments. The injected faults for IDV 16 through 20 have an insufficient description for validating results. The other faults not described contain fault modes that are not associated with single components and thus cannot be evaluated. This leaves us with the experiments shown in Table II.

An *ok* indicates that the faulty component was found as part of the minimal cardinality diagnosis and a *not ok* shows that the faulty component was not found as part of the diagnosis. We were able to find all faults that had an identifiable component fault as the cause of the process disturbance. The change of input ratios can only be detected indirectly, since no observations are available at the inputs.

IDV	Fault isolated	Injected Fault
1	not ok	Feed ratio changed
6	ok	Pipe A feed loss
8	ok	Feed ratio changed
13	not ok	Reactor kinetics fault
14	ok	Reactor cooling fault
15	ok	Condenser cooling fault

TABLE II  
EXPERIMENTAL RESULTS WITH THE TENNESSEE EASTMAN PROCESS

These two result tables show clearly that the AITwin concept can be used for a wide range of AI applications—from data-based machine learning problems to symbolic AI diagnosis problems.

### B. AITwin and Research Projects

In the following, some projects and publications from the field of "AI for Production" are listed and analyzed with regard to the AITwin. For this we split those references into the field of condition-monitoring, predictive maintenance, optimization and the field diagnosis and planning.

### C. Condition Monitoring, Predictive Maintenance, Optimization

There is a huge number of projects and research works in the field of machine learning regarding anomaly detection, condition monitoring, predictive maintenance and optimization and the following review can only cover a small part of this work. Starting with the project SMARTPas [31], which deals with the prediction of beverage characteristics, the data & prediction API can be used to predict new data points for control and optimization. The projects IMPROVE [32] and PrognoseBrain [33], dealing with condition-monitoring and anomaly detection in production systems with ML, could have used the data & prediction API as well. In terms of research works e.g. Li [17], who predicts the surface roughness in extrusion-based additive manufacturing could have used the proposed DigitalTwin. Eiteneuer [12], performing an anomaly detection with LSTMs, can use the data & prediction API as well. Other examples in which the DigitalTwin could have been used can be found in [35], giving a comparative survey with ML in manufacturing, in [7] using ML in digital manufacturing or in [6] who does an on-line defect recognition. Within all these projects and research work, instead of having several incompatible data models, possibly resulting in not comparable results, one could have used a single DigitalTwin and the data & prediction API.

### D. Diagnosis and Planning

This section reviews several typical projects and research work in the field of symbolic AI, mainly covering diagnosis and planning tasks. Especially Diagnosis algorithms can, if consistency-based, get all necessary information from the AITwin reference models. It needs to be kept in mind that other diagnosis methods can be added, but within the AITwin concept consistency-based diagnosis is especially suited for CPS—please refer to [8] for details. Regarding the projects,

e.g. EfA [5], which deals with the development of intelligent assistance systems, can use the causality API for the planning of automation systems. The project IMPROVE [32], in which innovative modelling approaches for production systems are developed, can use the causality API for the detection of system causalities. In terms of recent publications, e.g. [27], aiming to do an automated process planning for CPS and [25] doing non-convex hull based anomaly detection in cyber-physical production systems could use the AITwin. Other examples are [1] who gives a detailed review of ML for diagnoses in production or [36] who analysis gas wells. In every work the causality API of the AITwin can be used. Regarding planning tasks, projects and research work can also derive most of its needed information from the AITwin. So far, descriptions of machines must be implemented individually but can refer to process step descriptions as capability models. Furthermore, the computed overall production process is also stored individually, but could be added in a straight forward manner. The bottom-line is that the analyzed consistency-based diagnosis activities for CPS can be implemented with the AITwin model. The planning activities can to a large extent be implemented with the AITwin model, some proprietary information must be added.

## VI. CONCLUSION

In section I, three research questions have been identified as key challenges for the definition of a common DigitalTwin reference model, called AITwin. For this, section IV analyzes conceptual requirements of CPS to timing models, section IV-B analyzes conceptual requirements of CPS to machine learning models and section IV-C analyzes conceptual requirements of CPS to symbolic AI models. Based on these requirements, concepts and APIs are derived in section IV-D. This AITwin answers the three research questions.

Section V uses a running example and projects to empirically verify the AITwin reference model leading to the point that most AI and ML activities can be implemented using it. The authors do not claim that the reference models is the ultimate definition of a common AI DigitalTwin. But it is to a significant extent able to meet the requirements and may serve as a crystallization point for further research.

Assuming that both DigitalTwins and AI/ML methods are crucial elements of future CPS architectures, which is a rather safe bet, both worlds need to be harmonized. The AITwin reference model can prevent unnecessary redundant modeling efforts and allows for a reuse and sharing of computed results. Finally, this can be a starting point for a standardization of an AI-enabled DigitalTwin for CPS.

**Acknowledgements:** This work was partly developed within the Fraunhofer Cluster of Excellence "Cognitive Internet Technologies".

## REFERENCES

[1] T. T. Ademuji, M. P. Brundage, and V. V. Prabhu. A review of current machine learning techniques used in manufacturing diagnosis. In *IFIP International Conference on Advances in Production Management Systems*, pages 407–415. Springer, 2017.

[2] K. M. Alam and A. El Saddik. C2ps: A digital twin architecture reference model for the cloud-based cyber-physical systems. *IEEE Access*, PP:1–1, 01 2017.

[3] T. R. Besold, A. S. d'Avila Garcez, S. Bader, H. Bowman, P. M. Domingos, P. Hitzler, K.-U. Kühnberger, L. C. Lamb, D. Lowd, P. M. V. Lima, L. de Penning, G. Pinkas, H. Poon, and G. Zaverucha. Neural-symbolic learning and reasoning: A survey and interpretation. *ArXiv*, abs/1711.03902, 2017.

[4] M. Bevilacqua, E. Bottani, F. E. Ciarapica, F. Costantino, L. Di Donato, A. Ferraro, G. Mazzuto, A. Monteriù, G. Nardini, M. Ortenzi, and et al. Digital twin reference model development to prevent operators' risk in process plants. *Sustainability*, 12(3):1088, Feb 2020.

[5] B. Böttcher, J. Badinger, N. Moriz, and O. Niggemann. Design of industrial automation systems—formal requirements in the engineering process. In *Emerging Technologies & Factory Automation (ETFA)*, 2013.

[6] A. Caggiano, J. Zhang, V. Alfieri, F. Caiazzo, R. Gao, and R. Teti. Machine learning-based image processing for on-line defect recognition in additive manufacturing. *CIRP Annals*, 68(1):451–454, 2019.

[7] I. M. Cavalcante, E. M. Frazzon, F. A. Forcellini, and D. Ivanov. A supervised machine learning approach to data-driven simulation of resilient supplier selection in digital manufacturing. *International Journal of Information Management*, 49:86–97, 2019.

[8] A. Diedrich and O. Niggemann. Model-based diagnosis of hybrid systems using satisfiability modulo theory. In *Thirty-Third AAAI Conference on Artificial Intelligence (AAAI-19)*, Hawaii, USA, Jan 2019.

[9] J. J. Downs and E. F. Vogel. A plant-wide industrial process control problem. *Computers & chemical engineering*, 17(3):245–255, 1993.

[10] K. Dröder, P. Bobka, T. Germann, F. Gabriel, and F. Dietrich. A machine learning-enhanced digital twin approach for human-robot-collaboration. *Procedia Cirp*, 76:187–192, 2018.

[11] B. Eiteneuer, N. Hranisavljevic, and O. Niggemann. Dimensionality reduction and anomaly detection for cps data using autoencoder. In *20th IEEE International Conference on Industrial Technology (ICIT)*, Melbourne, Australien, Feb 2019. IEEE.

[12] B. Eiteneuer and O. Niggemann. Lstm for model-based anomaly detection in cyber-physical systems. In *Proceedings of the 29th International Workshop on Principles of Diagnosis*, Warsaw, Poland, Aug 2018.

[13] A. Fuller, Z. Fan, and C. Day. Digital twin: Enabling technology, challenges and open research. *arXiv preprint arXiv:1911.01276*, 2019.

[14] <https://www.nist.gov/el/intelligent-systems-division-73500/ieee-1588>. www accessed 1/23/20.

[15] G. Lample and F. Charton. Deep learning for symbolic mathematics. *ArXiv*, abs/1912.01412, 2020.

[16] E. A. Lee. Cps foundations. In *Proceedings of the 47th Design Automation Conference, DAC '10*, New York, NY, USA, 2010. ACM.

[17] Z. Li, Z. Zhang, J. Shi, and D. Wu. Prediction of surface roughness in extrusion-based additive manufacturing with machine learning. *Robotics and Computer-Integrated Manufacturing*, 57:488–495, 2019.

[18] Y. Lu and X. Xu. A digital twin reference model for smart manufacturing. In *48th International Conference on Computers and Insturial Engineering, CIE*, 12 2018.

[19] W. Luo, T. Hu, C. Zhang, and Y. Wei. Digital twin for cnc machine tool: modeling and using strategy. *Journal of Ambient Intelligence and Humanized Computing*, 10(3):1129–1140, 2019.

[20] A. M. Madni, C. C. Madni, and S. D. Lucero. Leveraging digital twin technology in model-based systems engineering. *Systems*, 7(1), 2019.

[21] <https://www.w3.org/TR/REC-MathML>. www accessed 1/23/20.

[22] Q. Min, Y. Lu, Z. Liu, C. Su, and B. Wang. Machine learning based digital twin framework for production optimization in petrochemical industry. *International Journal of Information Management*, 49:502–519, 2019.

[23] V. Mullachery, A. Khera, and A. Husain. Bayesian neural networks. *CoRR*, abs/1801.07710, 2018.

[24] J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, New York, NY, USA, 2nd edition, 2009.

[25] L. Peng and O. Niggemann. Non-convex hull based anomaly detection in cpsps. *Engineering Applications of Artificial Intelligence*, 87, 2019.

[26] A. Rasheed, O. San, and T. Kvamsdal. Digital twin: Values, challenges and enablers from a modeling perspective. *IEEE Access*, PP:1–1, 01 2020.

[27] A. Rogalla and O. Niggemann. Automated process planning for cyber-physical production systems. In *22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, 2017.

- [28] R. Rosena, J. Jäkel, M. Barth, and et. al. Simulation und Digitaler Zwilling im Engineering und Betrieb automatisierter Anlagen. In *VDI Tagung Automation*, 2019.
- [29] R. Stark and T. Damerau. Digital twin. In L. Laperrire and G. Reinhart, editors, *CIRP encyclopedia of production engineering*. Springer Publishing Company, 2019.
- [30] F. Tao, Q. Qi, L. Wang, and A. Nee. Digital twins and cyber-physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering*, 5(4), 2019.
- [31] <https://www.th-owl.de/init/en/forschung/projekte/b/filteroff/449/single.html>. www accessed 1/10/20.
- [32] <http://www.improve-vfof.eu>. www accessed 1/10/20.
- [33] <https://www.th-owl.de/init/en/forschung/projekte/b/filteroff/291/single.html>. www accessed 1/10/20.
- [34] E. J. Tuegel, A. R. Ingrassia, T. G. Eason, and S. M. Spottswood. Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, 2011, 2011.
- [35] D. Wu, C. Jennings, J. Terpenney, R. X. Gao, and S. Kumara. A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests. *Journal of Manufacturing Science and Engineering*, 139(7), 2017.
- [36] J. Zhu, G. Cao, W. Tian, Q. Zhao, H. Zhu, J. Song, J. Peng, Z. Lin, H.-Q. Zhang, et al. Improved data mining for production diagnosis of gas wells with plunger lift through dynamic simulations. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 2019.