2022 IEEE 19th International Conference on Software Architecture Companion (ICSA-C) | 978-1-6654-9493-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICSA-C54293.2022.00024

Towards Trust-preserving Continuous Co-evolution of Digital Twins

[*] M. M. Bersani, [†] C. Braghin, [‡] V. Cortellessa, [§] A. Gargantini, [¶] V. Grassi, [¶] F. Lo Presti,
[*] R. Mirandola, [‡] A. Pierantonio, [†] E. Riccobene, and [§] P. Scandurra
[*]Politecnico di Milano, Italy, [†]Università degli Studi di Milano, Italy, [‡]Università degli Studi dell'Aquila, Italy,
[§]Università degli Studi di Bergamo, Italy, [¶]Università di Roma Tor Vergata, Italy,

***Abstract*—Physical systems are complex, and their behaviour is typically affected by uncertainty caused by incomplete knowledge about their state and operations. The emergence of Digital Twins (DTs), i.e., virtual copies of physical systems, leverages the opportunity of supporting *in-vitro* diagnoses and decision-making. However, the adoption of DTs cannot neglect the *trust* that stakeholders have in them and in their insights. Here, we present a preliminary framework for architecting DTs where appropriate modeling notations, trust assurance, and co-evolution between DTs and physical systems are part of a holistic method aiming to guarantee that a DT conforms, in terms of configuration and behaviour, to its corresponding Physical Twin (PT).**

Introduction

Physical systems are complex, as they are made of heterogeneous components, and their behaviour is affected by uncertainty caused by incomplete knowledge about their state and operations [1]. The concept of Digital Twin (DT) has recently emerged as a virtual high-fidelity machine-processable representation of a physical system. By coupling physical entities (i.e., objects, processes, humans, or human-related features) to their virtual copies through a continuous and bidirectional flow of data, it leverages the benefits of both the virtual and physical environments to support diagnoses, and decision-making [2], such as:

- *safety/performance analysis*, e.g., by means of a *what-if* state-space exploration, which enables control and adaptation;
- *maintenance*, e.g., through the identification of an effective system evolution paths leading to a desired to-be state from the current as-is state;
- *model-based design*, as the DT can be seen as a blueprint of a system under construction that can, e.g., help in observing in advance the interactions with other systems.

For these reasons, the concept of DT can represent the core of an architectural framework and paradigm for designing, analyzing, implementing, controlling, and adapting complex physical systems with components belonging to different domains (e.g., cyber-physical, business, and societal systems). The co-presence of the physical and virtual parts of the system and their continuous interaction over time underpin the different activities throughout their entire lifecycle [3]. Since its inception, the concept has increasingly gained momentum, driven by advances in related technologies such as IoT, big data, and machine learning. It is now considered a key innovation enabler and strategic technology trend, already successfully adopted by leading organizations and enterprises [4].

However, the broad adoption of DTs and their full exploitation heavily depend on the *trust* that stakeholders have in DTs as well as in the insights they provide. While the research in this area has so far primarily focused on privacy and security [5], one of the significant gaps identified by a recent systematic review of literature is represented by the fidelity [6]. Indeed, the accuracy with which the virtual copy reflects the physical counterpart depends on the quantitative and qualitative selection of the parameters and their abstraction levels exchanged by the virtual and physical environments. In this respect, the critical role of DTs in a physical system lifecycle requires firm trust being based on a broader set of concerns. A stakeholder will hardly rely on a DT that guarantees acceptable levels of privacy and security but fails to meet required performance, dependability, and safety levels. More importantly, a DT must faithfully reflect the physical system's behaviour to avoid out-of-time or wrong feedback. Therefore, the following concerns appear essential:

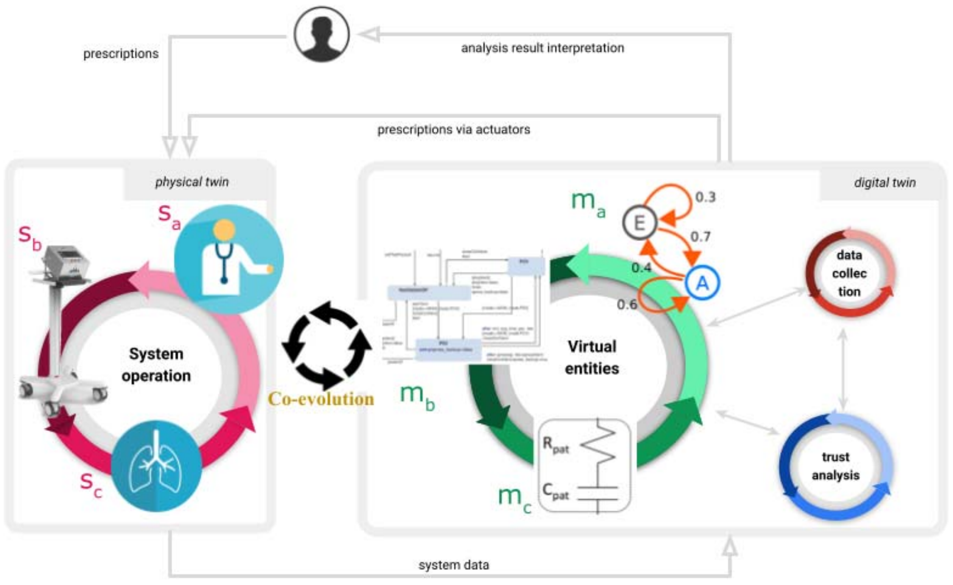
i) the degree of compliance between the behaviour of a PT and the corresponding behaviour modeled by its DT;
ii) DT-based performance, dependability, and safety of the physical system, even in presence of uncertain human behaviour and environment.

In this paper, we propose a conceptual framework that enhances the current state-of-art in the following major areas:

- *modeling* using notations that dominate/govern the complexity arising from the heterogeneity of the system components and the environmental uncertainty, through multi-view, multi-paradigm, multi-dimensional models;
- *trust assurance* by means of techniques and approaches rendered into explicit DT components that continuously check/predict trust-related properties, like behavioural conformance, safety, performance, dependability;
- *co-evolution* of DTs and physical systems through traceability mechanisms, and continuous impact analysis of the runtime changes applied to the DT, the physical system, or to the human in the loop feedback.

In the next section, our conceptual framework is introduced through a case study in the medical appliance domain.

978-1-6654-9493-9/22/$31.00 ©2022 IEEE
DOI 10.1109/ICSA-C54293.2022.00024



Physical Twin - System operation		Digital Twin - Virtual entities	
s_a	One or more members of the *medical staff* (e.g., doctors, nurses) who take care of the patient undergoing assisted mechanical ventilation. Specific workflow defining the procedures for caring that are applied by the medical staff are inherently considered.	m_a	The *stochastic models* that capture the behaviour of the medical staff members, the uncertainty of human actions and the workflow of medical procedures, together with spatio-temporal models representing the historical and near future movements of medical staff within the care unit.
s_b	The mechanical *ventilator* and the entire breathing circuit.	m_b	The *digital copy of the ventilator*, which includes the software adopted to control the physical sensors and actuators, as long as all the virtual copies of the sensors and actuators.
s_c	The sensors and the data collection system that collects data and enables the definition of the behaviour and the physiological characteristics of the *lungs of the patients*.	m_c	The *model of the lungs* of the patients, which can be as simple as a resistance-compliance model or as complex as an elaborated neural-network.

Fig. 1. Architecture of our conceptual framework for a Digital Twin applied to a mechanical ventilation plant

CONCEPTUAL FRAMEWORK

DTs are typically employed in the digital transformation of manufacturing systems and are today envisioned as the enabler means to improve personalized medicine, healthcare organization performance, and new medicines and devices [8]. We outline our conceptual framework through a case study in the medical domain and illustrate that it enables the design and development of a new generation of medical devices where trust and co-evolution play a central role.

Our case study is a lung mechanical ventilator. During the Covid 19 pandemic, our research groups were involved in the realization of the Mechanical Ventilator Milano (MVM) [7]. It is a low-cost and fast-to-develop lung ventilator that has been successfully designed, certified, and is currently built and delivered (especially to emerging countries). The MVM provides pressure-regulated ventilation support for patients in intensive care.

The architecture of our conceptual framework is illustrated in Figure 1, whereby our case study has been contextualized. The underlying table illustrates the correspondences between system elements and virtual entities of our case study.

Digital and Physical Twins exchange data through two traditional channels that are indicated in the figure as *system data* and (direct or human-interpreted) *prescriptions*. The former transfers data collected in the field to DT to update physical entity models and trigger decision-making processes. The latter brings the interpretation of the outcomes of the physical world analysis in terms of actions to be executed, which will align the current system state and behaviour to a targeted goal. In order to add trust-awareness to this consolidated paradigm, we envision the following improvements, all detailed in the next section. First, the virtual entities in the DT are represented by multi-view, multi-paradigm, multi-dimensional modeling notations to effectively represent the different aspects of physical entities and to capture the multi-form uncertainties that intrinsically occur in complex systems. Second, a *Trust analysis* task is introduced in the DT for assuring our extended concept of trust. Finally, we introduce

TABLE I
MODELING NOTATIONS AND DATA COLLECTION FOR THE CASE STUDY

Modelling Virtual Entities	**Data collection**
Discrete Event Formalisms: they model the evolution of the state of the ventilator-patient system that depends on the occurrence of events. Examples are Mealy/Moore/State machines, Petri Net, and their stochastic and hybrid extensions. *State machines model the controller of the ventilator.*	**Medical/forensic analysis**: it provides logs, including the physiological temporal profile of a patient, rendered into timestamped sequences of data. *Breathing parameters of interest can be the airway pressure, the tidal volume, the air flux, etc. [7]*
Stochastic: they model the uncertainty of the human behaviour in the ward, the variability of the working conditions, the human free will, etc. Examples are Stochastic Petri Net and Timed/Hybrid Automata, Markov decision processes, Queuing networks, etc. *Markov processes model work processes and the medical procedures of the staff.*	**System testing**: it provides numerical information about the correctness/safety/performance of the runtime behaviour of the ventilator. *Suitable dashboards are used by the medical staff, but also by the designers of the ventilator, to carry out automatic procedures for verifying the correctness of the ventilator functionality, such as the pump activity or the opening/clamping of the valves.*
Hybrid: they model physical phenomena that can be described through Ordinary Differential Equations such as the breathing of the patients. Examples are Hybrid automata, ODE systems, etc. *Human lungs are modelled as a system resulting in a resistor, which resists the airflow pumped through the ventilator, and a capacitor, which determines the mass that can be stored in the lungs.*	**Model training**: it provides all the required numerical information needed to refine (the models realising) the virtual copies of the patients and medical staff. *Physiological data are used to refine the parameters of the model of the lungs (the resistance or the capacity) or to train completely new models that might replace the current ones.*

a new process called *Co-evolution* that involves DT and PT, and that is activated when these two parts are not “aligned”. Misalignment intuitively occurs when DT is not an up-to-date representation of the physical system, or when the DT is modified to meet specific goals that do not coincide with the current state/behaviour of the physical system.

Based on our framework, a system like MVM can benefit from the use of a DT, starting from the architectural design throughout its entire lifecycle: at the level of *analysis*, with enhanced logging and monitoring capabilities that help clinicians improve the care; at the level of *maintenance*, with the identification of possible or potential failures of the machines (as also mandated by law). Concerning the *design*, DTs can help the development and testing of more powerful and safer ventilators before their use with actual patients.

STEPS AHEAD OF OUR DT VISION

Hereafter, we illustrate the application of our envisioned conceptual framework to the running case study.

Modeling heterogeneous systems with uncertainty

The heterogeneous nature of physical system components, often realized by different vendors with specific technologies and operational parameters, calls for a coordinated family of modeling notations to be used at any stage of the DT lifecycle. The system description shall consist of multiple views, possibly based on different notations, to help designers capture uncertainty aspects as well as the aleatoric variability of parameters (e.g., resource demands, failure rates) and indices (e.g., response time, reliability on demand). Besides the static and behavioural aspects, the adopted modeling notations have to allow the *digital shadow* specification, i.e., the rigorous description of (historical and real-time) data that enables trust analysis and predictive/prescriptive functions. Table I reports a list of possible modelling notations and mechanisms for data collection that may be used for engineering a DT of the next generation of ventilators.

Trust assurance

Trust is a vital concern in the adoption of digital systems. We intend to widen the scope of trust that in the DT context has been limited, up today, to privacy, safety, and security issues (e.g., [9], [5], [10]). Provided that a DT giving late or wrong feedback is less trustable than a timely and correct one, we target performance and availability/reliability as extra attributes to become first-class ones in the trust assurance process. However, a premise to the mentioned non-functional guarantees is that a DT (in terms of configuration and behaviour) conforms to its corresponding PT. Thus, the DT-PT conformance represents the basis for trust assurance. The left column of Table II reports examples of trust analysis techniques that may be applied to our case study.

Co-evolution of DTs and physical systems

Evolution is an inevitable aspect that affects the whole lifecycle of systems, and it occurs because of changing requirements, detected software defects, or new insights emerging from the domain. A more complex form of evolution is co-evolution, which occurs whenever two or more artefacts mutually affect each other’s evolution because of implicit or explicit dependencies. PTs and DTs are strictly coupled entities that co-evolve according to continuous data exchange. The DT monitors and records the PT’s operational states and can give stakeholders useful feedback if such information is truthful and up-to-date. A significant benefit in using DT-based analysis is the ability to carry experiments (e.g., what-if analysis) that would be costly on a PT, such as analyzing the performance sensitivity to system parameter variations. Based on a solid connection between runtime information and DT design, changes in the PT can be suggested to meet

TABLE II
TRUST ANALYSIS AND CO-EVOLUTION FOR THE CASE STUDY

Trust Analysis	Co-evolution
Simulated training: it allows medical staff to learn patient behaviour before treatment, using a virtual model, to avoid damage to the patients. *Newly recruited doctors to learn the use of the ventilator in managing different situations where patients show various pathologies. They use suitable dashboards on the virtual side to control the simulated patient-ventilator system.*	**From physical entities to virtual ones**. Modification of the physical side reflects on the virtual side: *a new sensor is introduced in the system for the measurement of a new physiological information about the patient (e.g., pulse oximeter for oxygen saturation); then, a new virtual copy of the sensor state is integrated with the current digital representation of the ventilator-patient system and the sensor data is used to update the digital representation of the device's state.*
Runtime QoS analysis and monitoring: it focuses on reliability and safety analysis of the ventilator coupled with the patient's breathing system. *Doctors continuously monitor patient's health status and run predefined procedures to check, for instance, if and when a breathing crisis might occur and if the ventilator can safely intervene with a prompt action. Moreover, the designers of the ventilator can carry out formal analysis to determine if the ventilator satisfies the safety/performance requirements using up-to-date models and set up, together with doctors, runtime procedures that guarantee integrity of the patients.*	**From virtual entities to physical ones**. Results of the analysis at the virtual side reflects on the physical side. • **Continuous refinement**: new requirements/scenarios imply a change in the virtual counterpart that is used to evaluate the system in these new contexts. *The ventilator has been used only for mechanical ventilation of adult patients, but emerging needs demand the use of the ventilator on children. This implies the addition of: the model of children's lungs and the model of a paediatrician together with the workflow that is suitable for the care of young patients. Moreover, the software for the control of the ventilator is modified.* • **Quality assessment** (if-what analysis): the Trust Analysis loop executed on the virtual counterpart allows for detecting violations of the quality requirements. Hence, changes to the virtual models are evaluated and possibly reflected into the physical world. *The digital models show a limited reactivity of the medical staff due to low response time in emergency management. Hence, a modification in the model of the medical staff is applied and validated.*
Data protection enforcement during the transmission of data from the physical world to the digital counterpart and viceversa. *The Data collection service can export logged data but only granted doctors/users can perform data exchange and read protected information. Hence, authentication, authorization mechanisms are enforced automatically.*	• **Prediction** (what-if analysis): the Trust Analysis loop applied on logs detects forthcoming criticalities. Hence, countermeasures must be considered to secure the physical world (no changes are applied to the virtual entities). *Symptoms that might lead to a respiratory distress are identified. Hence, the virtual counterpart enforces proactively a new respiratory profile in the patient by applying suitable corrective actions through the ventilator controller.*

(functional or non-functional) requirements before the system faces certain scenarios (e.g., some specific workloads). The right column of Table II reports how the DT-PT co-evolution may be managed in our case study.

CONCLUSIONS

In order to increase the level of trust in the results and indications coming from a DT, and to reduce the fidelity gap, we here sketched three interrelated research directions: modeling notations for heterogeneous systems with uncertainty (multi-view, multi-paradigm, multi-dimensional models); trust assurance techniques (behavioural conformance, safety, performance, dependability); co-evolution of DTs and PTs (traceability, impact analysis of the DT changes).

Besides problems arising from modeling approximations and uncertainties caused by incomplete or imprecise collected data, one of the open problems we identify is the integration of the results that different DT-based analyses might produce, and the synthesis of a set of non-conflicting prescriptive actions to be transferred to PT. This process of decision-making and integration could be subject to the *paradox of choice* [11] that might make a decision becomes overwhelming due to the many potential outcomes.

REFERENCES

[1] B. Schleich, N. Anwer, L. Mathieu, and S. Wartzack, "Shaping the digital twin for design and production engineering," *CIRP Annals - Manufacturing Technology*, vol. 66, pp. 141–144, 04 2017.

[2] Q. Qi, F. Tao, Y. Zuo, and D. Zhao, "Digital twin service towards smart manufacturing," *Procedia CIRP*, vol. 72, pp. 237–242, 2018.

[3] M. Grieves, *Origins of the Digital Twin Concept*, 08 2016.

[4] MarketsandMarket, "Digital twin market worth 35.8 billion by 2025," 2020. [Online]. Available: https://www.marketsandmarkets.com/PressReleases/digital-twin.asp

[5] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges and open research," *IEEE Access*, vol. 8, pp. 108 952–108 971, 2020.

[6] D. Jones, C. Snider, A. Nassehi, J. Yon, and B. Hicks, "Characterising the digital twin: A systematic literature review," *CIRP Journal of Manufacturing Science and Technology*, vol. 29, pp. 36–52, 2020.

[7] A. Abba *et al.*, "The novel mechanical ventilator milano for the covid-19 pandemic," *Physics of Fluids*, vol. 33, no. 3, p. 037122, 2021.

[8] B. R. Barricelli, E. Casiraghi, and D. Fogli, "A survey on digital twin: Definitions, characteristics, applications, and design implications," *IEEE Access*, vol. 7, pp. 167 653–167 671, 2019.

[9] A. Bécue, Y. Fourastier, I. Praça, A. Savarit, C. Baron, B. Gradussofs, E. Pouille, and C. Thomas, "Cyberfactory1 — securing the industry 4.0 with cyber-ranges and digital twins," *2018 14th IEEE International Workshop on Factory Communication Systems (WFCS)*, pp. 1–4, 2018.

[10] V. Damjanovic-Behrendt, "A digital twin-based privacy enhancement mechanism for the automotive industry," 09 2018, pp. 272–279.

[11] B. Schwartz, *The Paradox of Choice: Why More Is Less.* Harper Perennial, January 2005.