ORIGINAL ARTICLE

Towards AI-assisted digital twins for smart railways: preliminary guideline and reference architecture

Lorenzo De Donato[1] · Ruth Dirnfeld[2] · Alessandra Somma[1] · Alessandra De Benedictis[1] · Francesco Flammini[2,3] · Stefano Marrone[1] · Mehdi Saman Azari[2] · Valeria Vittorini[1]

Abstract

In the last years, there has been a growing interest in the emerging concept of digital twins (DTs) among software engineers and researchers. DTs not only represent a promising paradigm to improve product quality and optimize production processes, but they also may help enhance the predictability and resilience of cyber-physical systems operating in critical contexts. In this work, we investigate the adoption of DTs in the railway sector, focusing in particular on the role of artificial intelligence (AI) technologies as key enablers for building added-value services and applications related to smart decision-making. In this paper, in particular, we address predictive maintenance which represents one of the most promising services benefiting from the combination of DT and AI. To cope with the lack of mature DT development methodologies and standardized frameworks, we detail a workflow for DT design and development specifically tailored to a predictive maintenance scenario and propose a high-level architecture for AI-enabled DTs supporting such workflow.

Keywords Digital twin · Railway · Artificial intelligence · Machine learning · Cyber-physical system · Internet of things

1 Introduction

A Digital Twin (DT) is an accurate model of a physical entity which is kept alive at run-time and updated with real-time data collected from monitoring devices. DTs can be framed into general conceptual models including, among others, data collection management, model execution, and re-configurations planning [27]. The importance of the DTs is witnessed by the increasing number of ongoing standardization activities, such as the ones carried on by the ISO/IEC

Technical sub-committee on IoT and DT,[1] and the ISO/IEEE 11073 (Standardized DT Framework for Health and Well-Being in Smart Cities) [37]. In the railway sector, DTs can be used at different stages of the system life cycle and for different assets, including rolling stock (e.g., trains, trams, metros), signalling systems, infrastructures and manufacturing systems. They can be used to monitor physical assets, supervise train movements, provide information on passenger behaviour onboard trains, and detect or predict failures.[2]

Several surveys are available in the literature covering different aspects of DTs. Semeraro et al. [57] conduct a systematic review on the DT paradigm in manufacturing enterprises. The authors discuss questions such as what a DT is, where it is used, when a DT has to be used and why, and what the challenges are when implementing a DT. Perno et al. [49] perform a systematic review of the enabling technologies and barriers of DTs in the process industry. Kaewunruen et al. [35] discuss DTs for railways with a focus on Building Information Modelling (BIM) applications highlighting that BIM demonstrates a strong DT potential for railway maintenance and resilience. Bao et al. [6] first discuss the relationship and differences between DTs and

---

[1] https://www.iso.org/committee/6483279.html.

Lorenzo De Donato, Ruth Dirnfeld and Alessandra Somma have contributed equally to this work.

✉ Stefano Marrone
stefano.marrone@unina.it

1 Department of Electrical Engineering and Information Technology, University of Naples Federico II, Via Claudio 21, 80125 Naples, Italy

2 Department of Computer Science and Media Technology, Linnaeus University, Växjö, Sweden

3 School of Innovation, Design and Engineering, Mälardalen University, Eskilstuna, Sweden

---

[2] https://www.globalrailwayreview.com/topic/digital-twins.

traditional traffic simulation, then, they propose a three-layer technical architecture and analyse important technologies of DTs in traffic scenarios for intelligent transportation systems.

In a previous work [19], we provided the identified challenges and opportunities for DTs *in the rail sector* with a particular focus on maintenance applications and their combination with artificial intelligence (AI) techniques due to the potential that this combination can bring to railway digitalization [28, 58]. For example, DTs could be used to safely generate (synthetic) data related to assets' failures, by means of, for example, fault injection techniques [46], to improve the performance of AI systems oriented at detecting possible defects/faults [1]. On the other hand, AI could help to build more intelligent DT models [41] that could be used, among other applications, to understand how the physical asset would behave when parameters change. In summary, in our view, DT and AI have the potential to shift the current maintenance and inspection activities from scheduled and corrective to continuous and predictive. However, realizing an AI-assisted DT (i.e., a DT in which some components are enhanced by AI methodologies) is not a trivial matter, as its design, as well as its implementation and characterization, need to be aware of benefits and side effects that the use of AI can bring. Thus, in this paper, we want to make a step forward by introducing a preliminary guideline and reference architecture to be used when realizing an AI-assisted DT. In particular, with respect to our previous work we:

- Extended the review by adding the most relevant recently published papers. All schemes and charts have been updated accordingly;
- Improved the guidelines by introducing decision and task nodes taking into account factors like the aim of the DT, the decisions it has to make, which functionalities it has to provide, etc.;
- Realized a preliminary reference architecture to support the design of AI-Assisted DTs with a particular focus on predictive maintenance applications.

The remainder of this paper is organised as follows. Section 2 briefly summarizes the needed background on the DT paradigm and related technologies as well as on maintenance and smart maintenance approaches. In Sect. 3, we provide an overview of the adoption of DTs and AI in the railway sector. Challenges and open issues are discussed in Sect. 4. Section 5 provides our guidelines for the design and development of DTs with a specific focus on a predictive maintenance scenario. Section 6 sketches the proposed high-level conceptual architecture for AI-enabled DTs and, finally, Sect. 7 contains some final remarks and ideas on future works. Please note that as the number of analysed aspects resulted in the massive

Table 1 List of acronyms used in this paper

Acronym	Description
AI	Artificial Intelligence
AR	Augmented Reality
B5G	Beyond Fifth Generation
BIM	Building Information Modelling
CNN	Convolutional Neural Network
CNs	Connections
DD	Digital Twin Data
DL	Data Layer
DT	Digital Twin
gDT	Geometrical DT
HMI	Human–Machine Interface
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IIoT	Industrial IoT
IoT	Internet of Things
ISO	International Organization for Standardization
IT	Information Technology
KNN	K-nearest Neighbour
MAE	Mean Absolute Error
ML	Machine Learning
MQTT	Message Queue Telemetry Transport
PCD	Point Cloud Data
PD	Primary Data
PdM	Predictive Maintenance
PL	Physical Layer
PLM	Product Lifecycle Management
PS	Physical Space
PT	Physical Twin
RANSAC	Random Sample Consensus
RMSE	Root Mean Squared Error
RUL	Remaining Useful Life
SD	Secondary Data
SL	Service Layer
Ss	Services
VL	Virtual Layer
VR	Virtual Reality
VS	Virtual Space

use of several technical terms and acronyms, for the sake of readability we reported in Table 1 the list of the acronyms we adopted throughout the whole manuscript.

2 Background

To better frame the proposed guidelines and reference architecture, in this section we provide a brief background on DTs (Sect. 2.1) and on Predictive Maintenance 2.2, highlighting in
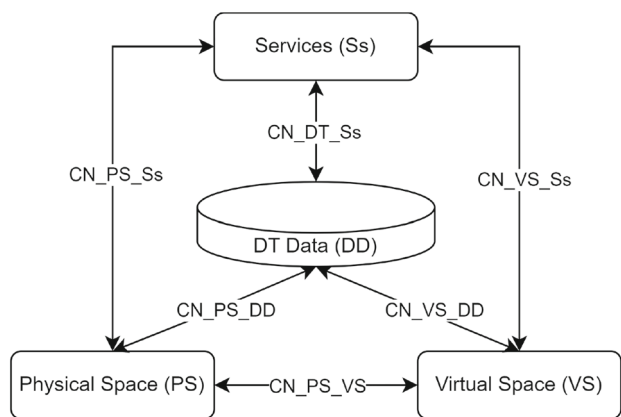
Fig. 1 Illustrative schema of the Digital Twin (DT) five-dimensional model according as described in Qi et al. [51]. CN_XX_YY indicates the connection that holds between the components XX and YY

both cases the connection existing with artificial intelligence (AI).

2.1 Digital twins

The Digital Twin is a *dynamic* and *self-evolving* virtual replica of a physical object and/or process, characterized by a bi-directional seamless communication that allows real-time data sharing between physical and digital world [7, 21, 63]. With DT technology, it is possible to implement several services (e.g., simulation, monitoring, prediction).

Despite DT informal introduction can be dated back to 2002 during Michael Grieves's presentation about product lifecycle management (PLM), at present there is still a lack of a formalized definition and a standardized architecture for DT implementation [7, 38, 59]. Nonetheless, one of the most recurring definitions in academia and industry is the five-dimensional model [51, 59], reported in Fig. 1, according to which a Digital Twin Model (M_{DT}) is described by the following quintuple:

$$M_{DT} = (PS, VS, Ss, DD, CN) \tag{1}$$

where the *Physical Space* (PS) includes systems and/or processes and their internal and external interactions; the *Virtual Space* (VS) contains the digital replica fed with *Digital Twin Data* (DD), i.e., operational, environmental or historical data; the *Services* (Ss) are the applications that can be realized leveraging the DT technology (e.g., predictive maintenance); Finally, there are the *Connections* (CN) that create the communication link (data and control) between the four parties.

A DT is often depicted as a "mixture" of Industry 4.0 technologies, as it owes its industrial and academic diffusion to the massive exploitation of innovative information and communication technologies [49, 51]. For example,

Internet of Things (IoT) and *Industrial Internet of Things* (IIoT) are among the enablers of the seamless communication between physical and virtual replicas: sensors capture events or environmental changes that can be further processed and analyses; actuators allow the physical system to respond to a set of inputs or commands received from the virtual replica. Data exchange is clearly achieved through *communication technologies*, such as MQTT, 5G/6G networks and so on.

Recently, *artificial intelligence* is more and more playing a key role in DT design and implementation, not only for data analytics purposes but also for the development of data-driven DT-based services (e.g., predictive maintenance). It is worth noting that AI and DT can leverage each other in two main different ways (not necessarily mutually exclusive): AI can be used to support DT modules (and this is what we will analyze in this paper) or a DT can be used to generate data to train an AI model.

Finally, *Virtual and Augmented Reality* technologies (VR/AR) better allow users to interface with the digital replica, thus enabling the human–machine interfacing with a computer-generated 3D visual environment.
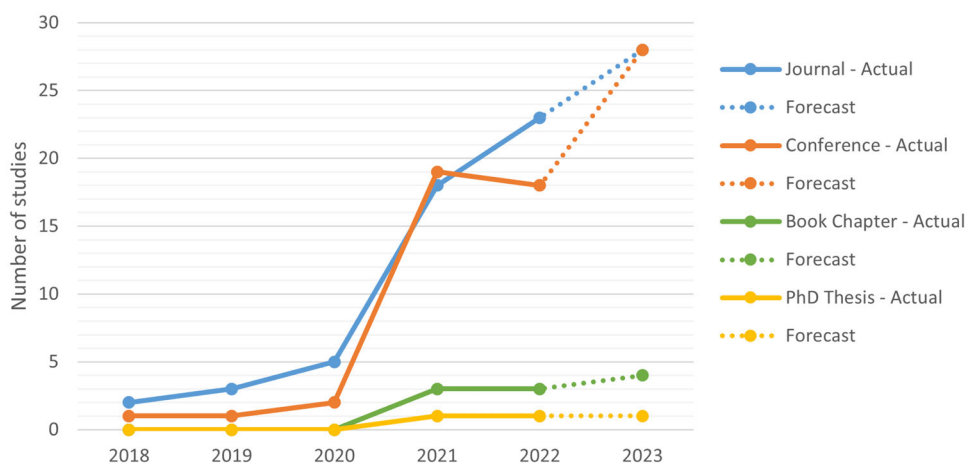
2.2 Maintenance approaches

As pointed out in Yokoyama [67], 21st-century innovation in terms of services for customers and technology supporting railway operations have enabled the shift from time-based maintenance to condition-based maintenance, leveraging large volumes of data directly acquired from monitoring to predict deterioration and failures of rolling stock and electrical equipment. The main maintenance paradigms and related actions can be distinguished in the following three categories [73]: *corrective maintenance* occurs when the fault is detected; *preventive maintenance* programs the repair actions at specific times according to past experiences; *predictive maintenance* (PdM) schedules the repairing activities predicting the failure occurrence through asset data analysis. According to Errandonea et al. [22], predictive maintenance also provides plans or strategies for taking action when the fault is predicted. Major benefits of PdM are maintenance costs and downtime decreasing, productivity and quality increasing.

The approaches used for prediction and thus for predictive maintenance are [22, 73]: (i) *physical model-based*, in which mathematical modelling reflexes component conditions; (ii) *data-driven*, relying on statistics-based, pattern recognition or AI models based on large amounts of data collected from the asset; (iii) *knowledge-based* approaches, aimed at reducing the complexity of physical modelling (typically used as a hybrid strategy).

The DT ecosystem provides a great opportunity for incorporating predictive models to evaluate the current state of a physical asset, analyse historical and operational data

Fig. 2 Distribution of studies over the years and by publication type (please note that some papers in 2022 might be still missing for indexing reasons)

Table 2 Mapping DT references to relevant AI categories

AI category	Literature works
Machine learning	Liu et al. [40], Gan et al. [30], Boscaglia et al. [11], Zhang et al. [69], Consilvio et al. [14]
	Ikeda [33], Sahal et al. [53], Milosevic [45], Yang et al. [66], Wang et al. [62], Zhou et al. [71]
	Lumban-Gaol et al. [42], Zhang et al. [68], Dimitrova and Tomov [17], Chen et al. [13], Wu et al. [64]
	Ariyachandra and Brilakis [4], Febrianto et al. [24], Peng and Zheng [48], Patwardhan et al. [47]
	Flammini et al. [28], Yang et al. [65], Zhang et al. [70], Torzoni et al. [60], Ferdousi et al. [25]
Intelligent Sensor Data Integration	Errandonea et al. [23], Jiang et al. [34], Bustos et al. [12], Zhou et al. [72], Gürdür Broo et al. [31]
Data mining	Malek et al. [44], Du et al. [20]
Adversarial search	Ambra and Macharis [2]
Computer vision & Image Processing	Avizzano et al. [5], Futai et al. [29]
Evolutionary computing	Ricondo et al. [52]
Operations research	Magnanini and Tolio [43]
Other	Boockmeyer et al. [9]

obtained from the sensors of the physical asset, and finally predict the degradation of a specific component.

3 An overview of DTs and AI in railways

We performed a literature review including 100 papers addressing DTs in railways. The selected papers obtained via the literature review are categorised by year in Fig. 2. We then reduced the set to 38 papers specifically addressing DTs in combination with AI. We refer to the AI categories presented in reference [8], which are reported in Table 2. In the next sections, we will report some consideration on the analysed papers, grouped according to the main topic into *Machine Learning*, *Intelligent sensor data integration*, *Data Mining*, *Computer vision and image processing*. It is worth noting that the aim is not to provide a systematic literature review but just to describe the most relevant recently published papers, based on the experience we matured working on the topic.

3.1 Machine learning

Machine Learning (ML) applied to railway DTs mainly focuses on predictive maintenance [14, 40, 53] as a natural application of the former to the latter. Liu et al. [40] propose a framework enabling the application of industrial AI technologies for a Prognostic Health Management system for high-speed railways. The authors discuss the use of dynamic clustering of historical data for the identification of health conditions. Specifically, in a case study on traction motor condition monitoring, they discuss the use of an Artificial Neural Network for modelling the vibration between sensors that are close to the four wheels. Similarly, Consilvio et al. [14] propose a generic framework and build a Decision Support System by combining ML, data analysis, and simulation techniques for analysing real-time data to make automated decisions. The manuscript takes into account two case studies, one at the strategic level and the other at the operational level. At the strategic level, the authors describe a data-driven model based on the K-means

clustering method, Petri net models, and Monte Carlo simulations for decision support in order to evaluate a sustainable mix of interventions, such as renewals and refurbishments. On the other hand, at the operation level, they describe a data-driven model based on a one-class support vector machine, the Bayesian network approach, and a linear programming model in order to improve maintenance operations and minimise costs caused by sudden failures and service disruptions.

Sahal et al. [53] introduce a conceptual framework to fulfil the DTs collaboration requirements and to enhance an intelligent DT based on blockchain technology and predictive analysis. By using distributed ledger technology, several DTs can be connected through the blockchain network and provide data analytics and management across the railway sector. To minimise maintenance costs, the authors explain that the framework consists of predictive models that use data-driven ledger-based historical DT operational data and a distributed decision-making algorithm for improving the DTs collaboration. The authors argue that the framework can predict potential failures from rail-based operational data by applying ML algorithms and training the offline predictive model with this data. Then, an online predictive model can be evaluated to predict the failures in railways. The outcome of these models serves as input for the consensus algorithm to make the best decision for data anomalies or to help decision-makers.

One of the most important issues in DTs development is their design and generation. Ariyachandra and Brilakis [4] argue that generating a DT of an existing railway from its Point Cloud Data (PCD) is a time-consuming and tedious task. Automation processes of this task face the challenge of detecting masts from air-borne LiDAR data since masts when scanned from above, are visible as thin lines. Thus, it is difficult to differentiate objects of similar shapes, such as signal poles, from masts. The authors propose a method that begins with tools that clean the PCD and identify its positioning, then, the resulting data sets are processed and masts are detected using the RANSAC algorithm. To summarise, the article focuses on DT generation to save time and modelling costs and proposes a method to enable the rapid adoption of a geometric DT.

Lastly, Zhou et al. [71] introduce a DT framework for automatic train regulation and control with the goal of reducing train delay times and energy consumption. The authors achieve this goal by using a Convolutional Neural Network (CNN) to map the relationship between the train running time and the energy consumption of the speed profile and switch points.

3.2 Intelligent sensor data integration

Jiang et al. [34] discuss the railway industry's demand for on-site data collection, control, and management. The authors propose a monitoring platform design and architecture for intelligent high-speed railways. Then, they also discuss the DT architecture including aspects such as the collection of real-time data from the sensors and the integration of sensor monitoring data to provide optimal solutions for failure handling. The aim of the DT is to develop a high-intelligence stage to intervene during abnormal working conditions through automatic monitoring and to prevent accidents.

Similarly, Errandonea et al. [23] propose an IoT approach for intelligent data acquisition to generate DTs in the railway industry with the goal of creating an onboard system for maintenance prediction in trains. The authors detail the architecture of the approach in three aspects: (1) The communication module for near-real-time communication and batch information transfer; (2) The functional design of the system, where the processing and integration of the sensor data are done in modules so that the system remains independent; (3) Data ingestion technologies, where the authors make use of Apache NIFI since it allows different data sources and parallel executions.

3.3 Data mining

Du et al. [20] introduce a DT framework and its implementation method for urban rail transit. The proposed implementation method for the DT consists of four components, where: (1) the sensor network collects data; (2) onboard data is saved to a ground data storage device; (3) an analysis process is performed for mining and visualising data; and (4) AI algorithms are used for data mining, evaluating and predicting faults. The authors point out the necessity to pre-process the data by cleaning, integrating, transforming and reducing the data before the actual data mining process. For the analysis, the authors suggest using k-means when the labels belong to an unknown category, and for the classification of known data, they suggest using KNN, regression classification, neural networks, or naive Bayes.

3.4 Computer vision and image processing

Avizzano et al. [5] introduce an algorithm for reconstructing rolling stocks from a sequence of images into an image model. The aim is to be able to generate DTs for complete rolling-stock vessels to serve for monitoring, diagnosis and failure prediction. The authors combine Kalman Filters, Gaussian mixture models and specific algorithms to track images of a train, which were taken by a fixed camera, so to provide a light and robust tracking system for image stitching.

4 Challenges and open issues

In this section, we summarise the most significant open issues and challenges to be tackled in the context of intelligent DT in railway applications in order to reach their full potential. For each listed issue and open challenge, we report at least one paper from a European initiative that, in our opinion, addresses the point or provides an interesting view of it.

4.1 Interoperability

To enable DT to analyze and exchange data from heterogeneous systems in real time, interoperability is essential. Dimitrova and Tomov [17] argue that current railway networks are managed by different tools and devices, resulting in too heterogeneous and non-integrated data, which makes interoperability and decision-making especially challenging. Sahal et al. [53] discuss the DT interoperability challenge and propose a data-driven ledger-based predictive model, which ensures intelligent and secure interoperability.

4.2 Connectivity

Despite the large adoption of IoT sensors in different sectors, connectivity challenges still exist especially when dealing with real-time monitoring. When trying to connect all of the sensors in a railway system at the same time, it could face some obstacles and difficulties in obtaining data due to the variety and mutability of (harsh) environmental conditions including, among others, storms, freezing, and extremely high temperature which may disrupt connectivity and/or cause loss of data. To tackle this obstacle, Sahal et al. [53] discuss communication technologies such as Beyond Fifth Generation (B5G) and 6G as opportunities for handling connectivity among the increasing number of smart devices in intelligent transportation systems.

4.3 Lack of standards and frameworks

To be efficient, DTs require well-defined frameworks taking into account the IoT paradigm, data analytics techniques, and many other enablers. Furthermore, within all forms of DT development, also the modelling seems to be still an open issue given that, to the best of our knowledge, there is no standardized approach to model DTs. Standardized procedures enable domain and user comprehension, as well as information flow between phases of a DT's development and implementation. In a such view, Gan et al. [30] discuss the main challenges to developing standards for Industry 4.0 and achieving agreements with the goal to connect the entire supply chain and stakeholders. In Borjigin et al. [10], the authors discuss the evolution of BIM towards an international standard. BIM is described as a method using a shared digital representation of a built asset, which progresses to become a DT. In that context, the authors stress the need for collaboration and shared information among different stakeholders at various stages, so that more accurate decisions can be made.

In the railway sector, DTs can improve operational efficiency, reduce costs, and enhance the passenger experience. However, the lack of standardized frameworks and guidelines for implementing DTs in this sector can make it challenging for railway operators and manufacturers to adopt and implement DT solutions. Therefore, it is crucial to develop standardized frameworks and guidelines to ensure consistency, interoperability, and security when implementing DT solutions in the railway sector. Krishna et al. [36] tackle the lack of frameworks by introducing a train-track multibody simulation model to assess long-term rail surface damage evolution in accordance with European standards. The framework can be turned into a DT and provide guidance for infrastructure managers regarding the condition of the rail surface. Moreover, Shift2Rail[3] and CENELEC WG[4] are initiatives that recognize the significance of DT in this sector and aim to promote the development and implementation of innovative solutions in a coordinated and standardized way. Shift2Rail focuses on integrating new technologies and solutions into the railway sector, including those related to DT, while CENELEC WG has developed standards related to railway applications, signalling, and telecommunications, providing a framework for DT solution development and implementation. Collaborating with these initiatives to develop and implement DT solutions in accordance with established standards can help ensure interoperability, consistency, and effectiveness in the adoption and implementation of DT solutions in the railway sector.

4.4 Data privacy and security

Data privacy and security are other relevant challenges when it comes to railway systems, especially safety-critical ones. The main reasons are related, but not limited, to the fact that data exchanged (e.g., between trains and a zone controller) are extremely critical and any intrusion could potentially cause considerable damages both in terms of costs and human lives. Hence, to overcome this issue, DT-enabling technologies must meet existing security and privacy policies and legislation. Along this line, Boockmeyer et al. [9] point out the importance of IT security and suggest the use of appropriate encryption techniques and software update strategy.

---

[3] https://shift2rail.org/.

[4] https://www.cenelec.eu/.

4.5 Scalability

Scalability refers to the possibility of integrating and managing multiple DTs at the same time since (within manufacturing systems, as an example), it may happen that there are multiple DTs deployed each of which virtually represents a specific machine or workstation. In this context, Sahal et al. [53] identify in the blockchain technology a suitable solution for scalability since large amounts of DTs can work together in a hierarchical and granular manner.

5 Guidelines for DT design and development

Although all the issues presented in the previous section are of paramount importance when it comes to efficiently designing and implementing DTs, in this work, we mainly focus on the lack of guidelines and frameworks as, in our view, these aspects form the basis for the realisation of DTs. To that aim, as mentioned in the introduction, this section provides some guidelines that would be hopefully useful to assist the realisation of DTs, while Sect. 6 proposes a *preliminary* reference architecture for AI-assisted DTs, i.e., DTs encompassing intelligent functionalities to improve services and/or data analysis.

Even if the DT concept has been around for almost 20 years now, there is still a lack of mature methodologies to lead the process from the design to the implementation of an operating DT [56]. Starting from our review findings and the analysis of state-of-art DT literature concerning the design methodology of a virtual replica [3, 7, 50, 55, 56], we provide some guidelines for DT design and development with a focus on *data-driven predictive maintenance*. We decided to focus on this application as, in our opinion, it represents one of the most promising applications for DTs in railways.

As shown in Fig. 3, the workflow is structured in twelve different *tasks*: requirement specification, process planning, architectural design, digital representation, digital twinning, service planning, data flow, conversion, training process, prediction process, decision process and evaluation. The first six tasks are grouped in a blue box as they represent the design guidelines of the DT and the service (i.e., predictive maintenance) upon the DT. The remaining indicate the development and deployment guidelines. It is worth noting that each task involves different *steps*, each modelled by a node in the workflow diagram. The next paragraphs describe each, detailing functionalities and the rationale.

5.1 Requirement specification

The first task lays the groundwork for the design and the implementation of a DT, providing the primary information upon which the rest of the steps will be based. First of all,

it is needed to identify for which *industry* the DT will be developed [50] and this is significant as some industries (e.g., railway, aerospace) have to follow some regulations. Step two aims to extract DT functional requirements and thus to define its *purpose*, such as optimization, prediction, monitoring, security assessment, development improvement and so on [56]. The following step is the identification of the *process* or *asset* to digitally replicate: it may occur that the Physical Twin (PT) does not already exist; in this case, a prototype of the PT can be built to guide the construction of the actual DT [56]. For example, in the railway context, in the design phase engineers can start the creation of the DT of a train to produce an accurate computer model of every single part (inside and out), but also to carefully plan the sensing layer (e.g. sensors, IoT devices, and asset and facility management system). In general, using a DT in the asset's design stage helps designers and engineers to model, simulate, and conduct what-if scenarios analysis to improve and optimise the asset design. Once the physical counterpart leaves production, the DT can also help engineers to monitor it, increasing not only production management and work performance but also the health, safety, and wellbeing of workers and materials [56].

Once the global goal of the DT-based application has been identified, it is needed to define what *kind of use* the DT is expected to deal with [50]. Indeed, we should at least define the *DT usage frequency*, i.e., if the DT will be used when needed or continuously or even in hybrid mode, and the *DT nature*, i.e., static or dynamic, and thus if the DT will adapt to initial conditions alterations or not. In fact, even if the DT definition that we presented in Sect. 2.1 describes the DT as a dynamic and self-evolving virtual replica, a DT can execute in two modes:

- *simulation mode*, namely it simulates/emulates the behaviour of the physical system and thus it executes in a static way. For instance, when the DT is used in the design stage, it is not connected to a physical counterpart (that can not even exist), therefore it has no capability to evolve as the physical system changes with respect to time;
- *Replication mode*, i.e., the DT is used online and its models are fed with real-time data, so the virtual replica evolves along with its physical counterpart throughout its life cycle. This means that any changes in either the physical or Digital Twin are reflected in its counterpart, creating a closed feedback loop (that it's not true in the simulation mode execution) Segovia and Garcia-Alfaro [56], Ferko et al. [26] and Eramo et al. [21].

Further steps are the identification of *technologies* needed for DT development and specification of *input and output parameters* of the PT and so of the DT, according to each use

Fig. 3 Proposed guidelines to implement a DT in a predictive maintenance example, based on the references Ariansyah et al. [3], Du et al. [20], Liu et al. [40], Segovia and Garcia-Alfaro [56], Psarommatis and May [50], Schroeder et al. [55] and Barricelli et al. [7]
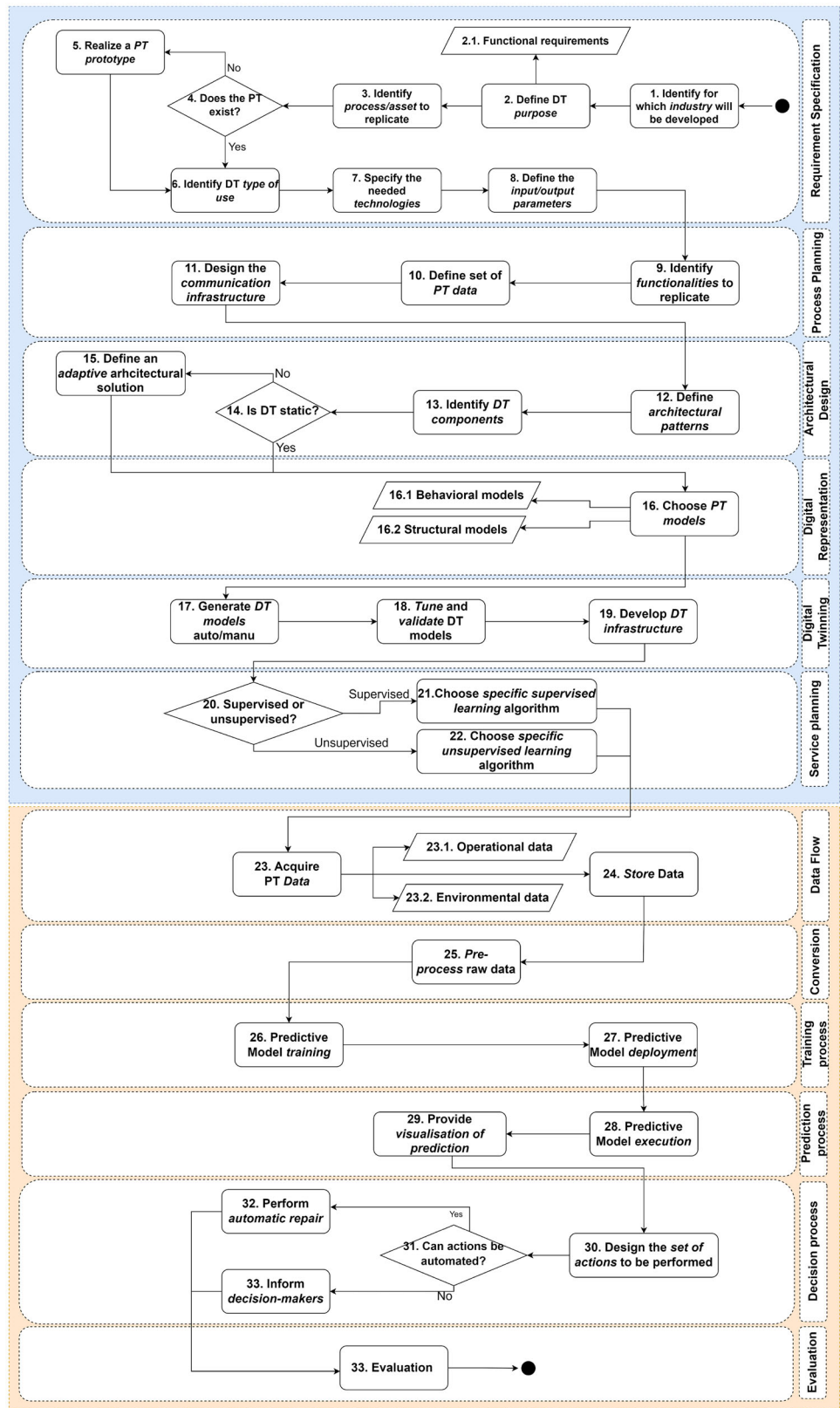
case. This will increase the DT understandability by a broader audience and the re-usability of DT in other scenarios [50].

5.2 Process planning

As the DT is typically a virtual replica of only a part of the physical process/asset [7, 38], the purpose of process planning is to determine the *functionalities* and properties of the physical twin to be modelled in the DT. In some cases, the DT may be an identical copy of the PT, but this one-on-one replication could most probably bring more complexity and redundancy [56].

In this task, the other two objectives are: i) the definition of the comprehensive *set of data* that suitably describes the physical twin [56] (e.g., environmental, historical, real-time or near real-time operational data) and that are used to tune, validate and eventually update the DT; and ii) the design of the DT and PT *communication infrastructure*. Depending on requirement specifications, for example, a 5G, B5G, or 6G communication technology could be taken into account.

5.3 Architectural design

As pointed out in Sect. 4, there is no consensus about DT components and corresponding architecture, but software architectures are cornerstones for the set-up and engineering of software complex systems such as DTs [15, 26]. For this reason, it is crucial to perform the architectural design task by focusing on both the *architectural pattern choice* and *DT components identification* steps.

Although there is a lack of a widely accepted architectural solution for DTs, according to Ferko et al. [26], the majority of the academic and industrial proposals are built using the layered architectural pattern or the combination of layered and service-oriented patterns. Furthermore, regardless of DT-based application, it is possible to identify a set of strictly needed DT components [55]: i) the *Models* that digitally represent the physical process/asset; ii) an *Even Source* whose objective is to send information and/or commands to the PT; iii) a set of *algorithms* (e.g., AI) provides the reasoning capability that aims to extract useful information which feeds DT models and the event source element.

Finally, if DT nature is dynamic, the architecture is expected to evolve over time to integrate system changes (e.g., new components), to *adapt* the system behaviour, and also to support the interconnection with other DTs [18, 56].

5.4 Digital representation

This task is entirely dedicated to the *modelling* activity since PT models are the core elements of the virtual replica. It is mandatory to reproduce the physical properties, geometry, behaviours, and rules of the physical system as faithfully as possible. According to Segovia and Garcia-Alfaro [56] and Schroeder et al. [55], it is possible to distinguish two modelling techniques:

1. **Behavioral modelling**, which expects to characterise the system behaviour through mathematical or computational descriptions of how the variables of interest are related to each other. This category includes:
   - *control models* that, based on control theory, apply physics laws and compare simulated results with known information, i.e., mathematical models;
   - *data-dependent models*, which can leverage AI and use data structures that store all the variables describing reality at a predefined abstraction level;
   - *hybrid control-data models*, which combine control and data-dependent models to obtain advantages from both of them;

2. **Structural modelling**, which is oriented at identifying the relationships between the parties undertaking certain activities in a structured description. This class comprehends:
   - *physical models*, representing physical properties and phenomena (e.g., deformation);
   - *geometrical models*, mirroring the geometry, shapes, sizes, positions, logic, and interfaces of the real system.

5.5 Digital twinning

The objective of this task is the *set-up of the whole DT infrastructure*, starting with the generation, tuning and validation of PT models [50, 56]. Indeed, in order to obtain a comprehensive view of the physical system, DTs are characterized by different kinds of models previously listed. Thus, in this task, it is required to integrate hierarchically or collaboratively models and other DT components (which and how many depend on the use case). It is worth underlining that deciding whether to automate the building of the DT or rather to construct it manually are both feasible choices: for example, Ariyachandra and Brilakis [4] generates a geometric DT (gDT) from its Point Cloud Data.

5.6 Service planning

This task aims to design a service that leverages DT technology. Thus, after DT design, in our work, we suggest defining the *predictive maintenance model* and choosing whether to perform it with supervised or unsupervised learning. Notable implementations relying on supervised learning algorithms exploit Neural Networks [71], Support Vector Machines or Naive-Bayes [14]. Similarly, unsupervised learning (such as

the K-means algorithm) has been used for anomaly detection [14]. Regardless of the chosen learning algorithms, this task will lay the basis for the implementation of the DT-based predictive maintenance service.

5.7 Data flow

Once the set-up of DT is completed, the virtual replica has to be connected to the physical twin through the seamless bidirectional connection that describes the DT itself, with respect to the communication infrastructure previously designed. Thus, the sixth task aims to create the *data flow*, i.e., the stream of environmental and operational data coming from the physical world combined with historical data. For example, in a railway context, operational data are the position of the railway [20], the temperature of the motor [11], or the vibrations [17], environmental data are the different weather conditions or humidity. All data will be collected and stored using one or more storage solutions, such as cloud services or a local server with Internet access.

5.8 Conversion

The objective of the conversion task is to *process*, *clean* and eventually *fuse and reduce data* due to their multi-temporal, multi-dimensional, multi-source and heterogeneous nature [7, 15, 39].

5.9 Training process

After collecting data from the physical system for the ML-based problem (i.e., predictive maintenance) and properly treating them with AI techniques, it is possible to build the *predictive model training* and *deploying* it. At the time of writing this paper, the most used DT platforms for predictive maintenance applications are OpenModelica and MathWorks Simulink [61].

5.10 Prediction process

This task operates when *ML-based predictive maintenance* model has been deployed. The algorithm executes using real-time data continuously coming from the physical twin and a *visualisation* of the prediction.

5.11 Decision process

Once the insight is obtained and a fault is predicted, the decision process task designs the set of actions to be performed on the physical twin. The DT can be used to support human decisions, perform automatic repair actions, or assign tasks to the right personnel [40].

5.12 Evaluation

The last task is the evaluation which can be done through the generation of a proof-of-concept. Indeed, although model validation metrics for AI-based models are overall accuracy, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and so on, different works [7, 26] have demonstrated that evaluation of DT-based solutions must almost always be carried out through examples. This is due to the lack of DT-specific evaluation measures because of the heterogeneity of its components and its dependency on the physical system.

6 An high-level architectural model for AI-assisted DTs

As discussed in Sect. 4, one of the main challenges related to AI-assisted DTs in the railway context is the lack of a standardized and widely accepted framework for their implementation. In this paper, we present a high-level architectural model that can support the architectural design task of our guidelines described in Sect. 5. The model, depicted in Fig. 4, leverages and extends the conceptual architecture proposed in De Benedictis et al. [15] by detailing the capabilities needed to implement a predictive maintenance service.

As already said before, one of the most recurring architectural patterns is the layered one, as it enables better addressing of the complexity of software-intensive systems such as DTs [26]. Thus, recalling Tao's dimensions [51], our architectural model is structured into four layers, i.e., the Physical, Data, Virtual and Services layers. Please, note that in the architecture proposed by De Benedictis et al. [15] there are other two levels, depicted vertically in relation to the other ones, and they are called Security and Connectivity Layer(s): the first one is taken into account in order to cope with DT security issues (e.g., protection of data on which digital models are constructed and/or that feed real-time model), while the second one indicates the two kinds of connection that can be set up in a DT environment. An in-depth analysis of these aspects is beyond the scope of this paper, as it is strictly related to the *connectivity* and *data privacy and security* issues identified in Sect. 4. Therefore, the reference architecture discussed below simply includes the necessary components that an AI-assisted DT should involve, without taking into account external threats or connections (besides those with the physical asset).

In addition to that, as in De Benedictis et al. [15], the proposed architecture is quite general. Indeed, each layer groups a subset of homogeneous DT core capabilities, which can be implemented by one or more components. However, as our focus is on DT usage in the railway context and in particular for predictive maintenance services, we specify the service layer and eventually modify the other levels with respect to
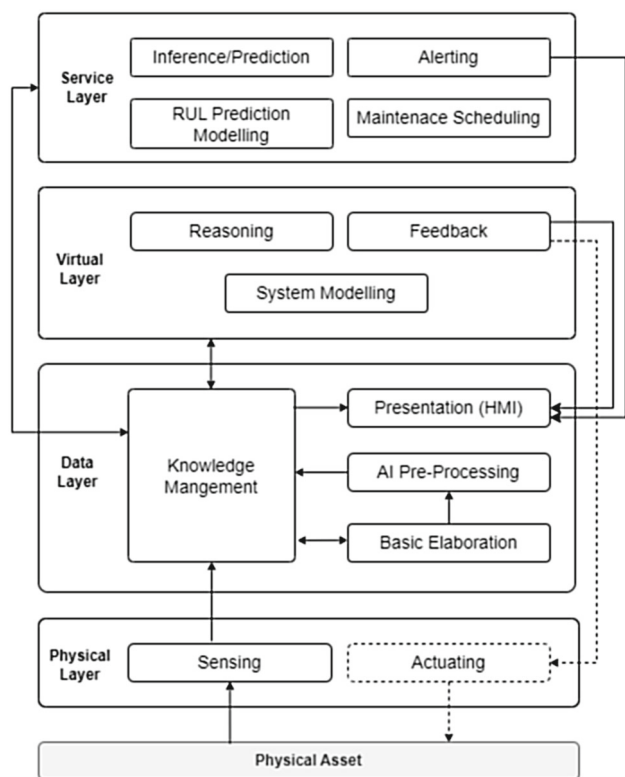
Fig. 4 The proposed high-level architecture for AI-assisted DTs for predictive maintenance of railway assets

functionalities needed for PdM, e.g., the necessity for data pre-processing in lower levels due to criticality and complexity of railway systems. Indeed, as far as we try to generalize the architectural model presented in this work, there are some aspects and needed capabilities that are strictly related to the application context and the kind of service implemented through DT technology [21, 51, 56]. In the following paragraphs, we will explain each layer of Fig. 4.

6.1 Physical layer (PL)

This layer interconnects the physical asset and its virtual counterpart (the DTs) by means of suitable **sensing** and **actuating** capabilities. The (IoT) sensors monitoring the physical asset collect different kinds of operational data which are then sent to the Data Layer for preliminary elaborations. Actuators, on the other hand, physically execute the decisions received from the upper layers.

6.2 Data layer (DL)

The Data Layer is in charge of managing both the row data collected through sensors at the PL, and the information retrieved through the elaborations performed at the upper layers, i.e., the Virtual Layer (VL) and Service Layer (SL).

This is achieved through a fundamental **knowledge management** capability that leverages shared knowledge. At the PL, depending on the asset and the choices made within the Requirement Specification phase of the guidelines proposed in Fig. 3, sensors may collect many heterogeneous data (e.g., vibration, acceleration, video, audio). We can subdivide these data into *Primary Data (PD)* and *Secondary Data (SD)*. By PD we mean those data that, after some basic elaboration that does not involve any AI application (e.g., cleaning, transformation), can be directly adopted at the upper layers to model the DT and perform the necessary analyses. Differently, by SD, we mean data that need to undergo an "advanced" pre-processing to be used at the upper layers. As an example, data captured by a camera may be processed by suitable AI approaches to extract motion features (i.e., new data) that can be used in the VL and SL for modelling and/or analysis. Capabilities that enable to extract both PD and SD are referred to as **basic elaboration** and **AI pre-processing**, respectively. It is worth noting that the integration of AI pre-processing capabilities is very helpful in a context such as a railway system. Indeed, most railway assets are safety-critical systems which must be compliant with strict standards and regulations to be deployed; intrusive modifications (e.g., the physical installation of new sensors on the deployed system) may lead to time-consuming and expensive (re)certification processes [16]. New assets, recently deployed, may already include (by design) all the sensors needed to produce their digitalized version (i.e., a DT); however, issues may arise in the case of legacy assets for which sensorization and digitalization might not be viable. The possibility to adopt AI pre-processing capabilities to elaborate data from non-intrusive sensors (e.g., cameras, microphones, LiDARs) and use these data to potentially digitalize any kind of railway asset would avoid the aforementioned issues. To conclude, the DL includes a **presentation** capability related to the Human Machine Interface (HMI), devoted to displaying information to human operators.

6.3 Virtual layer (VL)

This layer includes the core capabilities required to create and manage the virtual representation of a physical asset. The core capability within the VL is **modelling**, which typically encompasses the set-up and management of different models, characterising the structure and/or behaviour of the different parts of the asset. Besides "traditional models" (e.g., geometric, mathematical), AI can be adopted to build a behavioural model leveraging learning processes capable of extracting hidden information from data. In particular, system models may be built from historical data; for example, taking into account an ML approach, a model may be trained, validated, and tested on historical data and then used (without any modification) for simulations or predictions. An alternative

approach may adopt, for instance, online machine learning [32] to allow ML models to evolve in real-time with the on-line stream of data. The latter approach, despite being potentially more appropriate to keep the digital model constantly updated, is quite more challenging especially when it comes to Deep Learning applications most of which, notwithstanding their potential, are *opaque* (i.e., it is not always possible to understand their functioning and how the mapping between input and output has been performed) and *unstable* (i.e., to slight variations of the input may correspond heavy variations on the output). It is worth noting that it is already extremely challenging to demonstrate the dependability and trustworthiness of Deep Learning models [28] built starting from a specific and, if possible, reliable set of historical data. On-line learning would introduce additional complexity that would make this demonstration process even more complicated to address. A suitable compromise would be to periodically update the AI models (manually or automatically) so as to have the possibility to properly integrate new data and re-validate and re-test the new versions of the models.

Another core capability of the VL is that related to **feedback**, namely to the the propagation of decisions taken in the upper Service layer toward the Physical layer. However, given i) the safety-critical nature of most of the railway assets, ii) the current maturity level of AI approaches, and iii) the strict (yet necessary) standards regulating the rail sectors, it is quite unlikely that these actions can be autonomously performed by the DT itself. This is the motivation behind the fact that the arrows representing the connection (by means of the transmission of suitable control signals) between the feedback and the actuating capabilities provided by the PL are dashed; actually, in our conceptual architecture, the whole actuation capability is dashed to indicate the same limitation. At this stage, it is reasonable to think that "self-healing" features can be directly applied only if produced by certifiable functionalities. It is worth noting that, in any case, the feedback capability can still be integrated to support human operators (with proper messages generated for the HMI) or to perform simulations.

Finally, the **reasoning** capability involves the processing of data in order to extract useful information to feed DT models and generate feedback.

6.4 Service layer (SL)

This layer implements the forecasting and decision-making processes that can be attributed to DTs. In our proposal, it includes some capabilities that are specific to predictive maintenance, while additional capabilities may be considered depending on the specific application (e.g., simulation, monitoring, anomaly detection, etc.). The **inference/prediction** capability represents the forecasting logic

based on the selected prediction models. In our conceptual architecture, we have explicitly considered a **RUL Prediction Modelling** capability to set-up the models for the estimation of the remaining useful life (RUL) of an asset, as it is a key indicator identified by the research community to determine the maintenance time [54]. Based on the run-time prediction and other information stored in the shared knowledge, the **maintenance scheduling** capability is responsible for optimizing the maintenance operation scheduling. Finally, **alerting** capabilities must be typically considered to support maintenance.

7 Conclusions

In this manuscript, we provided an overview of the approaches related to the usage of digital twins in railway applications, with a focus on the role of AI and ML as enabling technologies together with the IoT paradigm. We highlighted some promising applications, design guidelines, as well as challenges to be tackled and future opportunities. Among the identified main open challenges, we focused on the lack of reference guidelines and architectural frameworks for the development of digital twins. Thus, starting from the review findings combined with the results of general DT literature analysis, we provided a set of guidelines to support the design and implementation of AI-Assisted DT. In conclusion, to further support researchers and practitioners, we realised a first architecture intended to serve as a skeleton for realising AI-assisted DT matching the introduced and described design guidelines in the context of predictive maintenance in railways.

Future studies will be oriented at analysing the remaining challenges, with a particular focus on security and data privacy issues, with the aim of providing proper considerations about (i) how the Security and the Connectivity Layers [15] can be integrated within the proposed reference architecture and (ii) the role AI could play to improve/support their functionalities. We will also further investigate and enlarge the proposed methodology, with the aim of fulfilling real-world requirements in terms of interoperability and trustworthiness. The aim is to realize a tool intended to support several practical applications, especially for predictive maintenance, operational optimisation, and proactive safety.

Availability of data and materials This declaration is "not applicable" as the reported research does not involve the use of any dataset.

Declarations

Conflict of interest All the authors have no competing interests or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Ethical approval This declaration is "not applicable" as the reported research does not involve any data from humans or animals.

References

1. Alderisio F (2019) Manutenzione predittiva con matlab. MATLAB Expo 2019. https://it.mathworks.com/content/dam/mathworks/ebook/gated/predictive-maintenance-ebook-all-chapters.pdf

2. Ambra T, Macharis C (2020) Agent-based digital twins (ABM-DT) in synchromodal transport and logistics: the fusion of virtual and pysical spaces. In: 2020 Winter simulation conference (WSC). p 159–169

3. Ariansyah D et al (2020) Digital twin development: a step by step guideline. In: 9th International conference on through-life engineering services

4. Ariyachandra M, Brilakis I (2020) Detection of railway masts in airborne lidar data. J Constr Eng Manag 146(9):04020105

5. Avizzano CA, Scivoletto G, Tripicchio P (2021) Robust image stitching and reconstruction of rolling stocks using a novel Kalman filter with a multiple-hypothesis measurement model. IEEE Access 9:154011–154021

6. Bao L, Wang Q, Jiang Y (2021) Review of digital twin for intelligent transportation system. In: International conference on information control, electrical engineering and rail transit (ICEERT). p 309–315

7. Barricelli BR, Casiraghi E, Fogli D (2019) A survey on digital twin: definitions, characteristics, applications, and design implications. IEEE Access 7:167653–167671. https://doi.org/10.1109/ACCESS.2019.2953499

8. Bešinović N, De Donato L et al (2021) Artificial intelligence in railway transport: taxonomy, regulations and applications. IEEE Trans Intell Transp Syst 23(9): 1–14

9. Boockmeyer A et al (2021) From CCS-planning to testautomation: the digital testfield of deutsche bahn in scheibenberg—a case study. In: 2021 IEEE international conference on cloud engineering (IC2E). p 258–263

10. Borjigin AO, Sresakoolchai J, Kaewunruen S et al (2022) Digital twin aided sustainability assessment of modern light rail infrastructures. Front Built Environ. https://doi.org/10.3389/fbuil.2022.796388

11. Boscaglia L, Boglietti A, Nategh S et al (2021) Numerically based reduced-order thermal modeling of traction motors. IEEE Trans Ind Appl 57(4):4118–4129

12. Bustos A, Rubio H, Soriano-Heras E et al (2021) Methodology for the integration of a high-speed train in maintenance 4.0. J Comput Des Eng 8:1605–1621

13. Chen S, Wu X, Zhou M et al (2022) Digital twin based train delay prediction system: design and realization. In: 2022 41st Chinese control conference (CCC). p 305–310. https://doi.org/10.23919/CCC55666.2022.9901773

14. Consilvio A, Solis-Hernandez J, Jimenez-Redondo N et al (2020) On applying machine learning and simulative approaches to railway asset management: the earthworks and track circuits case studies. Sustainability 12:2544

15. De Benedictis A, Mazzocca N, Somma A et al (2022) Digital twins in healthcare: an architectural proposal and its application in a social distancing case study. IEEE J Biomed Health Inf. https://doi.org/10.1109/JBHI.2022.3205506

16. De Donato L, Flammini F, Marrone S et al (2022) A survey on audio–video based defect detection through deep learning in railway maintenance. IEEE Access 10:65376–65400. https://doi.org/10.1109/ACCESS.2022.3183102

17. Dimitrova E, Tomov S (2021) Digital twins: an advanced technology for railways maintenance transformation. In: 2021 13th electrical engineering faculty conference (BulEF). p 1–5

18. Dirnfeld R (2021) Reuse in self-adaptive software systems: a literature review. Master's thesis, Linnaeus University, Faculty of Technology, Department of computer science and media technology (CM)

19. Dirnfeld R, De Donato L, Flammini F et al (2022) Railway digital twins and artificial intelligence: challenges and design guidelines. In: European dependable computing conference. Springer, p 102–113

20. Du W, Zhang T, Zhang G et al (2021) A digital twin framework and an implementation method for urban rail transit. In: 2021 Global reliability and prognostics and health management (PHM-Nanjing). The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (ISPRS Archives), Copernicus.org, Virtual event, pp 1–4

21. Eramo R, Bordeleau F, Combemale B et al (2022) Conceptualizing digital twins. IEEE Softw 39(2):39–46. https://doi.org/10.1109/MS.2021.3130755

22. Errandonea I, Beltrán S, Arrizabalaga S (2020) Digital twin for maintenance: a literature review. Comput Ind. https://doi.org/10.1016/j.compind.2020.103316

23. Errandonea I, Goya J, Alvarado U et al (2021) Iot approach for intelligent data acquisition for enabling digital twins in the railway sector. In: 2021 International symposium on computer science and intelligent controls (ISCSIC). p 164–168

24. Febrianto E, Butler L, Girolami M et al (2022) Digital twinning of self-sensing structures using the statistical finite element method. Data Centr Eng 3:e31. https://doi.org/10.1017/dce.2022.28

25. Ferdousi R, Laamarti F, Yang C et al (2022) Railtwin: a digital twin framework for railway. In: 2022 IEEE 18th international conference on automation science and engineering (CASE). p 1767–1772. https://doi.org/10.1109/CASE49997.2022.9926529

26. Ferko E, Bucaioni A, Behnam M (2022) Architecting digital twins. IEEE Access 10:50335–50350. https://doi.org/10.1109/ACCESS.2022.3172964

27. Flammini F (2021) Digital twins as run-time predictive models for the resilience of cyber-physical systems: a conceptual framework. Philos Trans R Soc A 379:20200369

28. Flammini F, De Donato L, Fantechi A et al (2022) A vision of intelligent train control. In: Collart-Dutilleul S, Haxthausen AE, Lecomte T (eds) Reliability, safety, and security of railway systems. Modelling, analysis, verification, and certification. Springer International Publishing, Cham, pp 192–208

29. Futai MM, Bittencourt TN, Carvalho H et al (2022) Challenges in the application of digital transformation to inspection and maintenance of bridges. Struct Infrastruct Eng 18(10–11):1581–1600. https://doi.org/10.1080/15732479.2022.2063908

30. Gan TH, Kanfoud J, Nedunuri H et al (2021) Industry 4.0: why machine learning matters. Springer, Cham, pp 397–404

31. Gürdür Broo D, Bravo-Haro M, Schooling J (2022) Design and implementation of a smart infrastructure digital twin. Autom Constr 136:104171. https://doi.org/10.1016/j.autcon.2022.104171

32. Hoi SC, Sahoo D, Lu J et al (2021) Online learning: a comprehensive survey. Neurocomputing 459:249–289. https://doi.org/10.1016/j.neucom.2021.04.112

33. Ikeda M (2020) Recent research and development activities in maintenance technologies for electric railway power supply systems. Q Rep RTRI 61:6–10

34. Jiang R, Wang W, Xie Y et al (2022) Research and design of infrastructure monitoring platform of intelligent high speed railway. In: 2022 IEEE 6th information technology and mechatronics engineering conference (ITOEC). p 2096–2099

35. Kaewunruen S, Sresakoolchai J, Lin Y (2021) Digital twins for managing railway maintenance and resilience. Open Res Eur 1(91):91

36. Krishna VV, Wu Q, Hossein-Nia S et al (2022) Long freight trains & long-term rail surface damage—a systems perspective. Veh Syst Dyn. https://doi.org/10.1080/00423114.2022.2085584

37. Laamarti F, Badawi HF, Ding Y et al (2020) An ISO/IEEE 11073 standardized digital twin framework for health and well-being in smart cities. IEEE Access 8:105950–105961

38. Liu M, Fang S, Dong H et al (2021) Review of digital twin about concepts, technologies, and industrial applications. J Manuf Syst 58:346–361

39. Liu Z, Meyendorf N, Mrad N (2018) The role of data fusion in predictive maintenance using digital twin. AIP Conf Proc 1949(1):020023. https://doi.org/10.1063/1.5031520

40. Liu Z et al (2018b) Industrial ai enabled prognostics for high-speed railway systems. In: 2018 IEEE international conference on prognostics and health management (ICPHM). p 1–8

41. Lu J, Zheng X, Schweiger L et al (2021) A cognitive approach to manage the complexity of digital twin systems. In: West S, Meierhofer J, Ganz C (eds) Smart services summit. Springer International Publishing, Cham, pp 105–115. https://link.springer.com/book/10.1007/978-3-030-72090-2

42. Lumban-Gaol YA et al (2021) A comparative study of point clouds semantic segmentation using three different neural networks on the railway station dataset. Int Arch Photogram Remote Sens Spat Inf Sci XLIII–B3–2021:223–228

43. Magnanini M, Tolio T (2021) A model-based digital twin to support responsive manufacturing systems. CIRP Ann 70(353):356

44. Malek NG, Tayefeh M, Bender D et al (2021) Live digital twin for smart maintenance in structural systems. IFAC-PapersOnLine 54:1047–1052

45. Milosevic M (2021) Towards model-based condition monitoring of railway switches and crossings. Ph.D. thesis, Chalmers Tekniska Hogskola

46. Orive D, Iriondo N, Burgos A et al (2019) Fault injection in digital twin as a means to test the response to process faults at virtual commissioning. In: 2019 24th IEEE international conference on emerging technologies and factory automation (ETFA). p 1230–1234

47. Patwardhan A, Thaduri A, Karim R et al (2022) Federated learning for enablement of digital twin. IFAC-PapersOnLine 55(2):114–119. https://doi.org/10.1016/j.ifacol.2022.04.179. (**14th IFAC workshop on intelligent manufacturing systems IMS 2022**)

48. Peng F, Zheng L (2022) Fuzzy rule-based neural network for high-speed train manufacturing system scheduling problem. Neural Comput Appl. https://doi.org/10.1007/s00521-022-07190-9

49. Perno M, Hvam L, Haug A (2022) Implementation of digital twins in the process industry: a systematic literature review of enablers and barriers. Comput Ind 134(103):558

50. Psarommatis F, May G (2022) A literature review and design methodology for digital twins in the era of zero defect manufacturing. Int J Prod Res. https://doi.org/10.1080/00207543.2022.2101960

51. Qi Q, Tao F, Hu T et al (2021) Enabling technologies and tools for digital twin. J Manuf Syst 58:3–21. https://doi.org/10.1016/j.jmsy.2019.10.001. (**Digital twin towards smart manufacturing and industry 4.0**)

52. Ricondo I, Porto A, Ugarte M (2021) A digital twin framework for the simulation and optimization of production systems. Procedia CIRP 104:762–767

53. Sahal R, Alsamhi S, Brown K et al (2021) Blockchain-empowered digital twins collaboration: smart transportation use case. Machines 9(9):193

54. Sang GM, Xu L, de Vrieze P (2021) A predictive maintenance model for flexible manufacturing in the context of industry 4.0. Front Big Data. https://doi.org/10.3389/fdata.2021.663466

55. Schroeder GN, Steinmetz C, Rodrigues RN et al (2021) A methodology for digital twin modeling and deployment for industry 4.0. Proc IEEE 109(4):556–567. https://doi.org/10.1109/JPROC.2020.3032444

56. Segovia M, Garcia-Alfaro J (2022) Design, modeling and implementation of digital twins. Sensors. https://doi.org/10.3390/s22145396

57. Semeraro C, Lezoche M, Panetto H et al (2021) Digital twin paradigm: a systematic literature review. Comput Ind 130:103469

58. Tang R, De Donato L, Bešinović N et al (2022) A literature review of artificial intelligence applications in railway systems. Transp Res Part C Emerg Technol 140(103):679

59. Tao F, Zhang H, Liu A et al (2019) Digital twin in industry: state-of-the-art. IEEE Trans Ind Inf 15(4):2405–2415. https://doi.org/10.1109/TII.2018.2873186

60. Torzoni M, Manzoni A, Mariani S (2022) Structural health monitoring of civil structures: a diagnostic framework powered by deep metric learning. Comput Struct 271(106):858. https://doi.org/10.1016/j.compstruc.2022.106858

61. van Dinter R, Tekinerdogan B, Catal C (2022) Predictive maintenance using digital twins: a systematic literature review. Inf Softw Technol 151:107008. https://doi.org/10.1016/j.infsof.2022.107008

62. Wang X, Song H, Zha W et al (2021) Digital twin based validation platform for smart metro scenarios. In: 2021 IEEE 1st international conference on digital twins and parallel intelligence (DTPI). p 386–389

63. Wu C, Zhou Y, Pereia Pessôa MV et al (2021) Conceptual digital twin modeling based on an integrated five-dimensional framework and triz function model. J Manuf Syst 58:79–93. https://doi.org/

10.1016/j.jmsy.2020.07.006. (**Digital twin towards smart manufacturing and industry 4.0**)

64. Wu X, Lian W, Zhou M et al (2022) A digital twin based fault diagnosis framework for bogies of high-speed trains. IEEE J Radio Freq Identif. https://doi.org/10.1109/JRFID.2022.3216331

65. Yang C, Ferdousi R, El Saddik A et al (2022) Lifetime learning-enabled modelling framework for digital twin. In: 2022 IEEE 18th international conference on automation science and engineering (CASE). p 1761–1766. https://doi.org/10.1109/CASE49997.2022.9926693

66. Yang J, Sun Y, Cao Y et al (2021) Predictive maintenance for switch machine based on digital twins. Information 12:485

67. Yokoyama A (2015) Innovative changes for maintenance of railway by using ict-to achieve smart maintenance. Procedia CIRP 38:24–29. https://doi.org/10.1016/j.procir.2015.07.074. (**Proceedings of the 4th international conference on through-life engineering services**)

68. Zhang S, Dong H, Maschek U et al (2021) A digital-twin-assisted fault diagnosis of railway point machine. In: 2021 IEEE 1st international conference on digital twins and parallel intelligence (DTPI). p 430–433

69. Zhang T, Du W, Zhang G et al (2021) Phm of rail vehicle based on digital twin. In: 2021 Global reliability and prognostics and health management (PHM-Nanjing). p 1–5

70. Zhang T, Ren G, Ming H et al (2022) Application exploration of digital twin in rail transit health management. In: 2022 Global reliability and prognostics and health management (PHM-Yantai). p 1–5. https://doi.org/10.1109/PHM-Yantai55411.2022.9942083

71. Zhou M, Hou Z, Liu J et al (2021) Digital twin-based automatic train regulation for integration of dispatching and control. In: 2021 IEEE 1st international conference on digital twins and parallel intelligence (DTPI). p 461–464

72. Zhou S, Dumss S, Nowak R et al (2022) A conceptual model-based digital twin platform for holistic large-scale railway infrastructure systems. Procedia CIRP 109:362–367. https://doi.org/10.1016/j.procir.2022.05.263. (**32nd CIRP design conference (CIRP Design 2022)—design in a changing world**)

73. Zonta T, da Costa CA, da Rosa Righi R et al (2020) Predictive maintenance in the industry 4.0: a systematic literature review. Comput Ind Eng 150:106889. https://doi.org/10.1016/j.cie.2020.106889