Journal of Intelligent Manufacturing
https://doi.org/10.1007/s10845-023-02236-8

Check for updates

The asset administration shell as enabler for predictive maintenance: a review

Jhonny Rodriguez Rahal[1,2] · **Alexander Schwarz**[1,2] · **Benjamín Sahelices**[1] · **Ronny Weis**[2] · **Simon Duque Antón**[2]

Received: 9 March 2023 / Accepted: 6 October 2023
© The Author(s) 2023

Abstract

The emergence of the Internet of Things and the interconnection of systems and machines enables the idea of Industry 4.0, a new industrial paradigm with a strong focus on interaction and communication between physical and digital entities, leading to the creation of cyber-physical systems. The digital twin and the standard for the Asset Administration Shell are concepts derived from Industry 4.0 that exploit the advantages of connecting the physical and virtual domains, improving the management and display of the collected data. Furthermore, the increasing availability of data has enabled the implementation of data-driven approaches, such as machine and deep learning models, for predictive maintenance in industrial and automotive applications. This paper provides a two-dimensional review of the Asset Administration Shell and data-driven methods for predictive maintenance, including fault diagnosis and prognostics. Additionally, a digital twin architecture combining the Asset Administration Shell, predictive maintenance and data-driven methods is proposed within the context of the WaVe project.

Keywords Asset administration shell · Predictive maintenance · Digital twin · Machine learning · Industry 4.0 · WaVe

Introduction

The continuous evolution of technology has made possible the development of fast and reliable communication systems, which nowadays seem to be ubiquitous (Carretero & García, 2013). Since the introduction of the Internet of Things (IoT) as a paradigm, large and complex networks of devices have been deployed, shaping the technological landscape. The application of IoT devices in industrial environments brings forth the fourth industrial revolution. Industry 4.0 (I4.0) is currently defining the standards for smart manufacturing and cyber-physical systems, aiming for improved connectivity and self-monitoring systems (Thoben et al., 2017).

✉ Jhonny Rodriguez Rahal
jhonny.rodriguez@estudiantes.uva.es

Alexander Schwarz
alexander.schwarz@comlet.de

Benjamín Sahelices
benjamin.sahelices@uva.es

Ronny Weis
ronny.weis@comlet.de

Simon Duque Antón
simon.duque-anton@comlet.de

1 GCME Research Group, Informatics Department, Universidad de Valladolid, Plaza de Santa Cruz, 8, Valladolid 47002, Spain

2 Comlet Verteilte Systeme GmbH, Amerikastraße 27, Zweibrücken 66482, Germany

The Asset Administration Shell

One of the main concepts used frequently within the context of I4.0 is the Digital Twin (DT), a digital replica of its physical asset in the real world. A DT can describe its physical counterpart accurately, assisting the design, testing, and manufacturing phases of the physical system, reducing time and expenses, and improving user safety (Grieves & Vickers, 2016).

In recent years, a new concept with the name Asset Administration Shell (AAS) was introduced within the I4.0-paradigm, with the objective of providing a new framework for the realization of DTs. The AAS is a standardized approach to implement DTs by creating a digital represen-

Published online: 08 November 2023

tation of an asset unambiguously. With the AAS, an asset can be represented as an entity composed of an arbitrary number of submodels, which are data blocks that allow structuring all the information related to the asset. At its core, the AAS is defined as the cornerstone of interoperability between applications that manage manufacturing systems. A detailed description of the AAS information model can be found in the specification (Bader et al., 2022).

Introducing predictive maintenance

In an industrial environment where maintaining constant operation and minimum downtime is key, the Industrial Internet of Things (IIoT) provides means to embed computers and sensor arrays into machines and industrial equipment to measure and exchange data. This allows for the acquisition of information regarding the underlying processes, especially in brown field applications with legacy industrial devices. Through the analysis of data in combination with the knowledge about the model of the systems involved, indicators can be established to determine whether such systems are performing under suitable conditions. However, ensuring continuous operation can pose a challenge because any device is subject to degradation and eventual failure due to wear and the difficulty of predicting the exact moment of failure. Executing corrective actions in time to help extend the durability of an asset is of vital importance under operational and safety points of view (Pech et al., 2021; Fioravanti et al., 2020).

Predictive Maintenance (PdM) has the aim of diagnosing and triggering maintenance in a timely manner, minimizing the risk of faults and maximizing operation time. This research field is of high current interest due to its economic and safety advantages, and lends itself readily to many application areas, namely automotive, manufacturing, aerospace and others (Zhang et al., 2021; Theissler et al., 2021; Li et al., 2022). Moreover, the use of data-driven methods is starting to become the preferred approach for deploying PdM strategies because of its ability to operate accurately on large volumes of complex operational data, which facilitates equipment prognostics (Theissler et al., 2021; Zonta et al., 2020; Zhang et al., 2019a).

Machine learning basics

Machine Learning (ML) can be defined as a set of methods capable of automatically detecting patterns in data and using these patterns for prediction or decision-making under uncertainty (Murphy, 2012). Deep learning is a subset of ML and includes multi-stage methods capable of learning different representations of the input in each stage to ultimately generate an output.

According to Murphy, there are three types of ML: supervised, unsupervised, and reinforcement learning. In the former, the model is trained by using a set of input and output samples, known as features and labels respectively. Throughout the learning process, the model generates an output for the provided features, which subsequently gets compared with the expected labels. The resulting difference between the output and the labels produces an error that is then utilized to adjust the model. Usually, the variable type that the label represents indicates the problem to be solved. For instance, if the labels represent a categorical variable, then the problem is called classification. In contrast, if the label is a continuous value, the problem is known as regression.

In unsupervised learning, the model is only exposed to the inputs. As there are no labels, the task is to find useful patterns in the data, also known as knowledge discovery. The third type, reinforcement learning, contemplates a system of positive and negative reward signals that lead the model to decision-making. However, this study will only consider the first two types of learning, as they contain the established methods for PdM.

In the context of PdM, a classification problem can correspond to either anomaly detection or fault classification. In the first case, for example, the ML model can use sensor measurements to learn to distinguish normal operating behavior from anomalous or unknown conditions. In the second case, given the features and labels for each known fault, the algorithm can determine what type of fault has occurred. Alternatively, time series forecasting and prognostics can be considered as regression problems. Furthermore, anomaly detection can be tackled using a clustering approach. In the absence of labeled data, the model groups the inliers (data during normal operation) and excludes the outliers (anomalous data) that tend to be different from the other samples. An example of this is illustrated in Fig. 1.

Paper contributions and outline

The focus of this research is the presentation of a solution architecture for implementing data-driven PdM through the AAS in the automotive industry, targeting specifically medium-duty hydrogen trucks such as those introduced by the WaVe project. The WaVe project (from the German "Wasserstoff - Verbrennungsmotor") aims to develop, test and optimize a hydrogen internal combustion engine for medium-duty trucks. In order to achieve this objective, a collection of data-driven methods for PdM and the current works regarding the AAS are surveyed to determine which approaches would be a better fit to the requirements of the project.

The remainder of this work is structured as follows. Section 2 introduces the search methodology for the related work and Sect. 3 discusses the state-of-the-art research results. Sec-

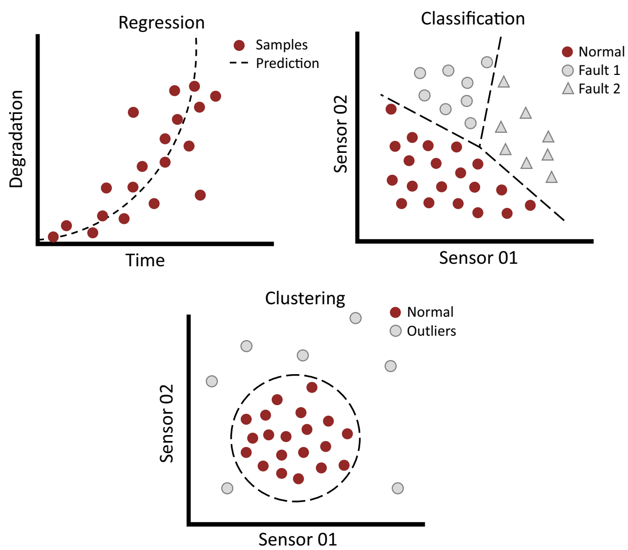

Fig. 1 Different PdM tasks according to the ML strategy

tion 4 presents the ambitions and requirements of the WaVe research project. The solution architecture is provided in Sect. 5, taking into consideration the requirements from the previous section and the results from Sect. 3. Finally, the conclusions are presented in Sect. 6. An overview of this structure is shown in a flowchart in Fig. 2.

The contributions of this work lie in:

- An Analysis of data driven-methods for PdM scenarios,
- the presentation of research project WaVe, and
- the introduction of a solution architecture for PdM based on data driven-methods and the AAS.

Search methodology for related work

The primary objective of this section is to collect and analyze the works of data-driven PdM and AAS for automotive applications. This review aims to answer the following questions:

1. RQ1: Which data-driven PdM approaches are being applied in the automotive industry?
2. RQ2: In which use cases is the AAS being implemented?
3. RQ3: How is the AAS being integrated to PdM?

The methodology used to collect the findings during this stage is based on the *PRISMA* protocol for systematic reviews proposed by Page et al. (2021). The collection of publications presented in this section is the result of a four-stage process: selection, screening, inclusion and discussion.

The selection process begins by addressing the research questions, to do so, four search queries are designed during the selection and applied to the ACM, Scopus and Web of Science (WoS) databases. All queries are limited to a time period between January 2017 and August 2022, the latter being the year in which this study is conducted. The search parameters and results are depicted in Table 1.

The first query is designed to find the publications of data-driven PdM algorithms enabled by the AAS. One of the early findings during the state-of-the-art research is the identifica-

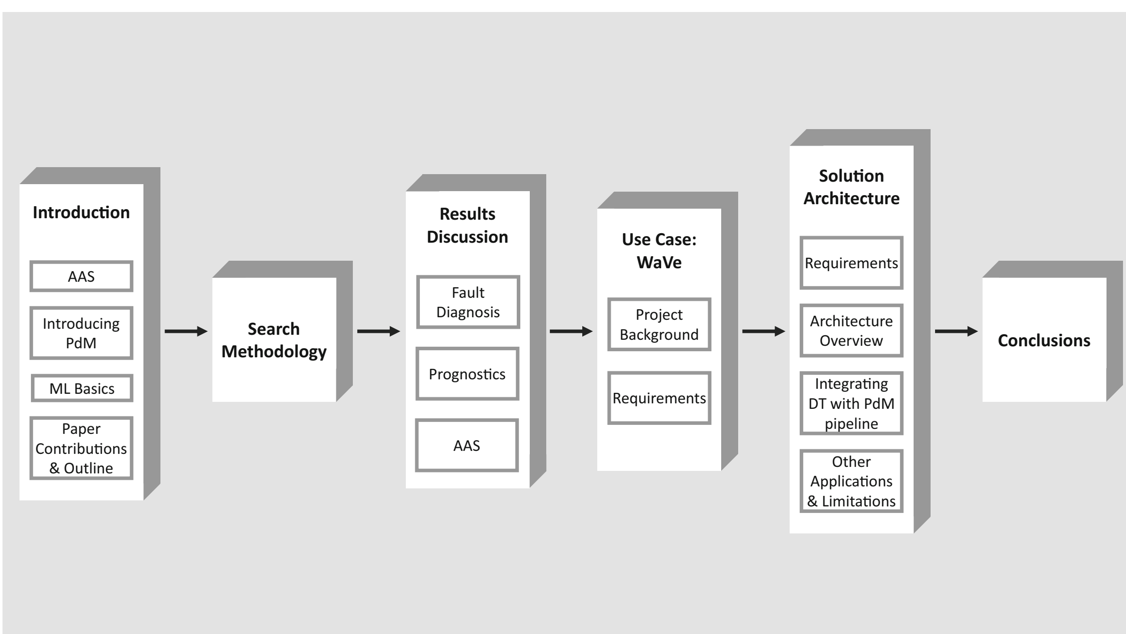

Fig. 2 Flowchart: structure of the paper

Table 1 Overview of Search Queries

#	Query	Database	Result
1	"Predictive maintenance" AND "Asset Administration Shell"	ACM	3
		Scopus	2
		WoS	4
2	"Predictive Maintenance" AND ("machine learning" OR "artificial intelligence" OR "deep learning")	ACM	218
		Scopus	1440
		WoS	940
3	"Asset Administration Shell" AND ("machine learning" OR "artificial intelligence" OR "deep learning")	ACM	1
		Scopus	8
		WoS	17
4	"Asset Administration Shell"	ACM	9
		Scopus	170
		WoS	122

Publication date: not older than 5 years

tion of a research gap in studies that contain both the AAS and PdM. The query used gave a total of 4 results from all the databases, after removing duplicate results. Due to the limited number of publications found, the search was separated into two topics: PdM using data-driven methods (such as ML and deep learning) and the AAS. The second query aims to find the works of data-driven PdM regardless of the application area leaving a total of 1958 publications. The third query aims to find publications related to the AAS including keywords related to data-driven applications. A total number of 21 papers resulted from this search query. The fourth query is utilized to expand the results of the third query by broadening the scope and considering exclusively AAS papers. This query returned 214 matches in total. The end result of merging all PdM papers and all AAS works led to 1958 and 214 publications respectively.

The next step in the methodology is the screening phase, where criteria are introduced to fine-tune the results obtained from the selection phase. Here a set of general exclusion criteria is applied to the PdM- and AAS-related publications. Additional criteria are also introduced for the PdM papers only, firstly, to narrow down the number of publications, since the list of results is extensive. For that reason, the PdM papers are filtered by application area, e.g. automotive and DT, which are key topics in this research. Secondly, to account for specific requirements related to data-driven applications such as dataset description and the use of performance metrics. On the other hand, papers addressing the AAS consider industrial use cases in general, since applying more specific fields like "automotive" reduces the scope of the search drastically. The following list includes the exclusion criteria used in this review.

General exclusion criteria for both topics:

- Only publications written in English.
- Limited to industrial application area.
- Must be a peer-reviewed article or paper.
- Must provide a structured and detailed methodology, facilitating its replication.

Additional exclusion criteria for papers addressing PdM:

- Use cases are automotive and DT.
- Must provide a description of the dataset.
- Must provide performance metrics and compare with reference or ground-truth values.

After applying the criteria to the AAS papers, a collection of 11 publications is obtained. In this case, the works with similar methodology and results were not included and the articles selected were found representative of the areas of interest of this research. On the other hand, the screening of the PdM publications reduced the resulting number to 266 papers by including only papers related to automotive and/or industrial applications. Finally, applying the remaining conditions of the exclusion criteria yielded a total of 28 publications.

For the inclusion phase, the results of the screening are organized and listed in Tables 2 and 3. Table 2 entries are sorted first by objective, in this case, fault diagnosis and prognostics and secondly by application area. In Table 3 the summarized contributions of the AAS papers are displayed. An additional column indicates whether the publication had any reference to PdM applications. The results discussion can be seen in Sect. 3.

Results and discussion

In this section, the publications listed in Tables 2 and 3 are discussed in more detail, reviewing the approaches and conditions of each scientific work to answer the research questions defined in Sect. 2. RQ1 is covered in Sects. 3.1 and 3.2, while RQ2 is answered in Sect. 3.3. RQ3 is discussed throughout Sect. 3. Therefore, following the same structure as the results tables, the focus of this section is distributed into three topics:

- Fault Diagnosis
- Prognostics
- Asset Administration Shell

Table 2 Overview Data-driven PdM Works

Area	Use Case	Approach	Ref
		Fault diagnosis	
Automotive	Transmission system	DCT	Wang et al. (2021)
	Transit bus	Gradient boosted trees	Mckinley et al. (2020)
	Hydraulic brake system	FURIA	Manghai and Jegadeeshwaran (2019)
	Internal combustion engine	MLP	Wang et al. (2020)
	Vehicle monitoring system	SVM	Shafi et al. (2018)
	Internal combustion engine	OSVM	Jung (2020)
	Vehicle's chassis	HMM	Soltanipour et al. (2020)
	Turbocharged petrol engine	Gaussian process, RF and others	Tessaro et al. (2020)
	Engine's oxygen sensor	MLP and DCT	Giobergia et al. (2018)
	Vehicle's safety components	MLP	Lee et al. (2019)
	Truck's air pressure system	MLP, BiLSTM, BiGRU and CNN	Rengasamy et al. (2020)
	Turbocharged petrol engine	CNN+LSTM	Wolf et al. (2018)
	Motorsport	LSTM + VAE	von Schleinitz et al. (2021)
	Commercial vehicle	GAN	Sun et al. (2019)
Other applications	Gearbox and roll bearings	LFGRU	Zhao et al. (2018)
	Production line	SSAE	Xu et al. (2019)
	Industrial machinery	Siamese AE	Castellani et al. (2021)
		Prognostics	
Automotive	EGR cooling system	RF	Sass et al. (2020)
	Electrical battery	ELM	Pan et al. (2018)
	High-pressure fuel system	SVM + MLP	Giordano et al. (2021)
	Internal combustion engine	Naives-Bayes classifier	Nixon et al. (2018)
	Autonomous vehicle	Dynamic Bayesian Network	Gomes and Wolf (2020)
	Battery and ball bearing	BDL	Zhu et al. (2022)
	Commercial vehicle	LSTM + MLP	Chen et al. (2021)
	Gearbox	GAN	Booyse et al. (2020)
Other applications	Bearing degradation	LSTM	Zhang et al. (2019)
	gas turbine engine	MLP, BiLSTM, BiGRU and CNN	Rengasamy et al. (2020)
	Tool wear	LFGRU	Zhao et al. (2018)
	Milling tool	RF	Wu et al. (2017)
	CNC machine	SAE + MLP	Luo et al. (2018)

Table 3 Overview AAS Publications

Contribution	PdM	Ref
Definition of an architecture of a robot platform with the AAS	✗	Ye et al. (2021)
Data conversion between physical asset and enterprise software using the AAS	✗	Ye et al. (2022)
Introduction of an administration shell to facilitate data exchange for data-driven applications	✗	Löcklin et al. (2021)
Methodology for the deployment of the AAS on embedded systems	✗	Pribiš et al. (2021)
Implementation of the AAS in an existing industrial plant using PLCs	✗	Schäfer et al. (2021)
Review of the state of the art of the AAS and implementation of AAS model in a production line use case	✗	Ye and Hong (2019)
Introduction of an approach to map different DT information models into AAS models and vice versa	✗	Platenius-Mohr et al. (2020)
Design of an user-friendly web interface to facilitate the creation and configuration of AAS-Files	✗	Arm et al. (2021)
Share insights and reasoning behind the methods to map the AAS metamodel into OPC UA	✗	Cavalieri and Salafia (2020a)
Introduction of a submodel focused on assisting maintenance carried out by humans	✗	Lang et al. (2019)
Architecture design of a DT for PdM based on the AAS	✓	Cavalieri and Salafia (2020b)

Fault diagnosis

Some of the previous works tackle the problem of fault diagnosis using ML approaches to identify key patterns that could reveal underlying faults in different parts of a system. In these publications, ML algorithms such as s Decision Tree (DCT), Support Vector Machine (SVM) and other statistical methods like Gaussian processes and Hidden Markov Model (HMM) are used for solving classification problems. For example, a solution based on DCTs is implemented in the work of Wang et al. (2021) for fault detection of the transmission system of a vehicle. Mckinley et al. (2020) approach the problem of failure prediction of NOx (nitrogen oxides) sensors on a transit bus using XGBoost a gradient boosted tree architecture, which according to the authors can handle missing data values appropriately.

Manghai and Jegadeeshwaran (2019) introduce FURIA (Fuzzy Unordered Rule Induction Algorithm) as a rule-learning strategy to classify the faults of a hydraulic brake system. Wang et al. (2020) present a Multilayer Perceptron (MLP) architecture with a single hidden layer to diagnose the faults of an engine using sound intensity analysis. A vehicle monitoring and fault predicting algorithm is proposed by Shafi et al. (2018). The authors compare the performance of K-Nearest Neighbors (KNN), DCT, Random Forest (RF) and SVM. In this study SVM shows the best performance. Jung (2020) implements an One-class support Vector Machine (OCSVM) algorithm that uses the signal residuals to detect and classify known and unknown faults of an internal combustion engine. A more statistical approach is presented by Soltanipour et al. (2020), where a HMM-based clustering is proposed as a solution to detect tire pressure and wheel hubs faults. A comparison between several ML methods is made by Tessaro et al. (2020), where Gaussian Processes and RF yield the best results in identifying the faults of a simulated Turbocharged petrol engine.

Deep learning approaches also have a strong presence in the fault diagnosis area. According to the search results, the MLP proved to be a popular choice. In the works of Giobergia et al. (2018), Rengasamy et al. (2020) and Lee et al. (2019), MLP algorithms are implemented for diverse use cases. Giobergia et al. (2018) make a complete analysis of oxygen sensor data to evaluate the sensor clogging status. The authors define a framework named "OXYCLOG" with which the procedure for processing data and training the ML algorithms is specified. Lastly, the performance of a DCT, MLP and SVM are compared. Although the MLP provides the best performance, the authors choose the DCT as the best solution due to its superior interpretability.

Rengasamy et al. (2020) develop dynamically weighted cost functions used to train deep learning architectures for the task of detecting faults in an air pressure system of heavy trucks. Since the dataset is unbalanced, the authors design the cost functions to compensate for the instances that are poorly learned during the training process leading to an improvement in the performance. This approach is tested on a MLP, a Convolutional Neural Network (CNN) and on Recursive Neural Network (RNN) such as Bidirectional Long-Short Term Memory (BiLSTM) and Bidirectional Gated Recurrent Unit (BiGRU).

Lee et al. (2019) carry out a performance comparison between ML and deep learning architectures to detect six different faults in automotive safety components. The researchers analyze four deep learning architectures and compare them against three other ML architectures. The evaluated algorithms are a MLP with and without an Autoencoder (AE), 1-D- and 2-D CNNs, while SVM, KNN and DCT algorithms serve as performance references. The results suggest that overall the deep learning approaches give more accurate predictions, being the MLP trained with statistical features, the approach with the best prediction results and the lowest training and testing times of the studied deep learning approaches. Wolf et al. (2018) introduce a hybrid deep learning network consisting of a CNN coupled with a Long Short-term Memory (LSTM) to perform data reduction and time series processing, respectively, with the goal of identifying the faults of a turbocharged petrol engine.

A RNN-based architecture is studied by Zhao et al. (2018). The authors propose a hybrid local feature-based Gated Recurrent Unit (GRU) network that uses manually extracted features to execute incipient fault diagnosis on rolling bearings and fault diagnosis on a gearbox. The proposed architecture yields higher accuracy compared to the other network topologies such as vanilla RNN, bidirectional and vanilla GRU, SVM, MLP and others. A framework design is presented by von Schleinitz et al. (2021) with the aim to achieve robust time series prediction in the presence of anomalies for motorsport applications. The proposed algorithm is a hybrid deep learning architecture that consists of a Variational Autoencoder (VAE), to facilitate the anomaly detection, and a BiLSTM network to predict the time series output in case an anomaly is present. According to the authors the framework is as accurate as the baseline for regular conditions, but for anomalous inputs the error is significantly reduced. Sun et al. (2019) introduce a time series anomaly detection using a Generative Adversarial Network (GAN). The network is tested and validated using real world vehicle data.

On another branch of fault diagnosis, Xu et al. (2019) present a fault diagnosis scheme assisted by a DT using a deep transfer learning approach for a production line use case. The authors use a stacked sparse AE network which uses the real data samples from the asset and complement them with the ones generated by simulations on the DT for training, allowing to transfer data from the virtual space into the physical space. The results show an increase in diagnostic accuracy by incorporating the data obtained from simulations.

Another study that integrates the DT to fault diagnosis is the one conducted by Castellani et al. (2021), in which a weakly-supervised approach for anomaly detection is introduced. The authors devise a neural network architecture based on a Siamese AE designed to work with limited data samples. The Siamese AE is trained using real world values and data from the DT. The results indicate that the architecture displays a better performance with the help of the DT data as opposed to only using real world data samples. The Siamese AE is also compared with other algorithms such as OCSVM, KNN, MLP and others, giving the proposed network the best performance metrics for anomaly detection.

A summary of the used models can be found in Fig. 3. The bar plot displays the data-driven models chosen as solution architectures and the ones selected for benchmarking. For simplicity, the machine and deep learning models are grouped using distinct labels given the similarities of the models. More details are available in Table 4.

From Fig. 3 it can be concluded that the group of DCT-based models (G-DCT) is the most used solution architecture for fault diagnosis applications, representing 20% of the selected solutions. From this group, the DCT model is found the one that is used the most. The G-MLP and G-RNN groups take the second place with a representation of 15%. In the case of G-RNN, the algorithm with the most uses is the vanilla LSTM. Although the group "Other" represents 20% of the total, the models that fall into this category are isolated cases which are not used as much as the previously

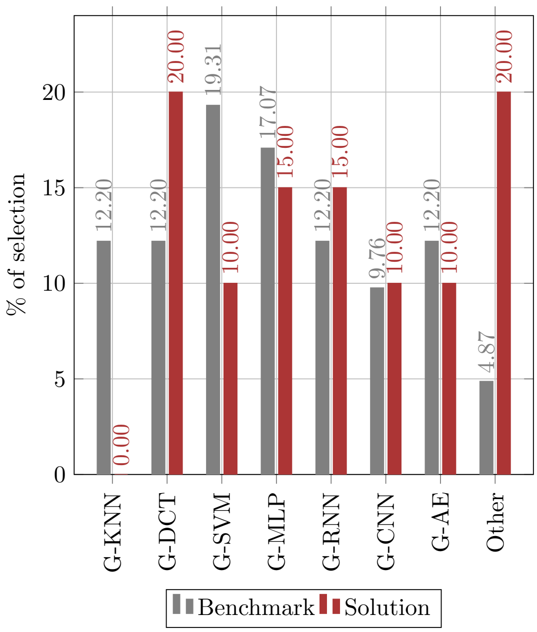

Fig. 3 Model selection for fault diagnosis applications

Table 4 Group labels for fault diagnosis models

Label	Includes
G-KNN	KNN
G-DCT	DCT, RF, Isolation Forest and Gradient Boosted Tree
G-SVM	SVM and OCSVM
G-MLP	MLP
G-RNN	RNN, (Bi)LSTM and (Bi)GRU
G-CNN	CNN, 1-D and 2-D CNNs
G-AE	AE, VAE, Sparse AE and Siamese AE
Other	FURIA, GAN, HMM and Gaussian processes

mentioned architectures. On the other hand, SVM and MLP are the most common choices for benchmarking, representing 19.31% and 17.07% respectively.

As it can be seen from the collection of studies presented, the data-driven approaches for implementing fault diagnosis are diverse. However, it seems that AAS has not yet reached this area of research. Considering that the AAS is a framework for developing DTs, it would be expected that the AAS would emerge in the works that specifically incorporate DTs into the machine learning architectures. Nevertheless, this topic is not mentioned in any of the reviewed publications.

Prognostics

On the side of prognostics and Remaining Useful Life (RUL) estimation, in which the objective is mainly model regression or multiclass classification, the applied methods are also varied. Sass et al. (2020) use several regression models to estimate the aging of the Exhaust Gas Recirculation (EGR) cooling system located in the powertrain of the vehicle. The models are trained using the datasets of three different vehicles. The authors compare a multiple linear regression model, a Bayesian linear regression and a RF regression model, being the latter the method that delivers the lowest estimation error. Wu et al. (2017) implement a RF architecture to estimate the tool wear of a CNC machine. The proposed solution is compared with MLP and SVM, with the RF being the algorithm with the best accuracy for that use case.

A State of Health (SoH) estimator for electric batteries is implemented by Pan et al. (2018). In the study, an Extreme Leaning Machine (ELM) is chosen as the ML architecture to determine the capacity degradation of the batteries. The performance of the proposed architecture is measured against a MLP, displaying less estimation error and faster prediction times. Giordano et al. (2021) introduce a prognostic pipeline for a high-pressure fuel system. The authors evaluate the performance of SVM and MLP architectures by varying different training parameters like total number of features, training size and the hyperparameters. The SVM is chosen as the final model due to its superior performance and stability in the conducted experiments.

The next group of ML approaches are based on Bayesian methods. Nixon et al. (2018) introduce a Linear Discriminant Analysis-Naives Bayes algorithm for health monitoring of diesel engines. The proposed solution demonstrates to outperform RF and SVM architectures under specific configurations of the training dataset. Gomes and Wolf (2020) develop a Dynamic Bayesian Network (DBN) to detect the faults related to the lateral and longitudinal controllers of an autonomous vehicle. The DBN allows to infer the faults not only in the current time frame but also several steps in the future.

Zhu et al. (2022) propose in their work a health prognostics strategy supported by a Bayesian Deep Neural Network (BDNN) architecture that implements active learning. The active learning strategy allows the network to start the training with fewer run-to-failure labels, add the missing labels and retrain to improve the RUL estimation. The performance of the deep learning architectures is tested and validated in two use cases: ball bearing and electric battery RUL prediction, showing good performance in both cases.

Continuing with the deep learning approaches, RNNs result to be a popular choice for the deployment of prognostics solutions (Zhang et al., 2019; Chen et al., 2021; Rengasamy et al., 2020; Zhao et al., 2018). Zhang et al. (2019) present a LSTM architecture and a novel indicator named the "waveform entropy" to estimate the degradation of bearings. Chen et al. (2021) introduce a method to estimate the RUL using statistical data taken from vehicles undergoing maintenance and complementing it with geographical information. The authors present a model named merged LSTM. The architecture consists of a LSTM network to process the sequential data and a MLP to process other types of data such as terrain data. The results show that the integration of maintenance and geographical data increases the estimation accuracy of the RUL. The proposed architecture is compared with four other ML algorithms: LSTM network, MLP, deep CNN and SVM, delivering the best accuracy but it is the second worst in training time.

Rengasamy et al. (2020) also tackle the problem of estimating RUL for gas turbine engines. Similarly to the fault detection scenario, the authors use a dynamically weighted cost function to train the models. The use of a customized cost function improved the RUL estimation considerably. Lastly, Zhao et al. (2018) validate the performance of the proposed local-feature GRU in a tool wear prediction scenario. Following the same line of thought as in the fault diagnosis use case, the authors compare the prediction error of the introduced architecture against the other ML and deep learning models, with the proposed model being the one with the best results.

Other promising approaches for prognostics-related applications are the GAN and AE-based architectures. Luo et al. (2018) design a method to estimate the health condition of a CNC machine using system identification. The implemented architecture consisted of a sparse AE and a MLP to detect valid impulse response signals from vibration measurements. On the other hand, Booyse et al. (2020) develop a generic framework for Prognostics and Health Monitoring (PHM) named "Deep Digital Twin" using GAN as deep learning architecture. According to the authors, the discriminator from the GAN is used to estimate the health indicator. The GAN architecture is also compared with a VAE network, but the former turned out to be more sensitive to variations of the healthy behaviour than the VAE, and therefore the GAN is the preferred choice.

In a similar fashion to Sect. 3.1, the statistics related to the model selection for prognostic applications are depicted in Fig. 4. Table 5 describes the group labels used in the bar plot. The plot shows the groups G-Bayes, G-DCTs and G-RNN as the most representative solutions with 21.43%. However, Bayesian methods can only be considered as a group and not individually, as the models are not used repeatedly in the selected publications. As for the DCT- and RNN-based models, RFs and LSTMs are the most used algorithms respectively. The architectures from groups G-MLP and G-RNN are the predominant choices as benchmarks, yielding 25% and 20.83%, respectively.

Moving on to the topic of the AAS and the prognostics use case, the findings for these domains, similar to the case of fault diagnosis applications, show no overlap between these two areas of study.

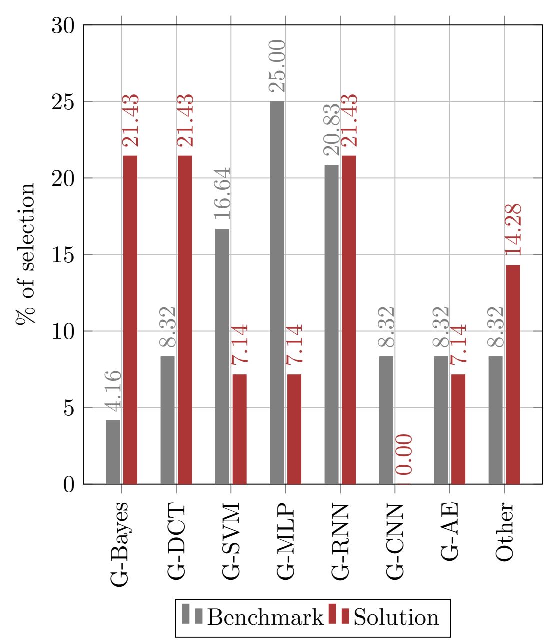

Fig. 4 Model selection for prognostics applications

Table 5 Group labels for prognostics models

Label	Includes
G-Bayes	Bayesian linear regression, dynamic Bayesian network, Bayesian neural network and LDA-Naives Bayes
G-DCT	RF and Gradient Boosted Tree
G-SVM	SVM and OCSVM
G-MLP	MLP
G-RNN	RNN, (Bi)LSTM and (Bi)GRU
G-CNN	CNN, 1-D and 2-D CNNs
G-AE	AE, VAE, Sparse AE and Siamese AE
Other	KNN, ELM and GAN

Asset Administration Shell

The collected publications regarding the AAS encompass diverse topics within the area of research, the focus of the studies spans between conversions between the AAS information models and other proprietary models, to practical implementations of the AAS in hardware and embedded platforms.

Ye and Hong (2019) provide a review of the state of the art related to the AAS and they also design an AAS template with an information model. The defined AAS model is applied to a production line to validate the approach. In another study, Ye et al. (2021) propose an AAS-enabled architecture for a robotic platform. The main contribution is not only the definition and design of the architecture but also the conversion algorithms between the AAS Exchange Format (AASX) files and the OPC Unified Architecture (OPC UA) node description. In this sense, the mapping between all the properties defined in the AASX file are linked to a different node for later access in OPC UA.

As a follow-up work, Ye et al. (2022) define a data conversion solution based on AASX for data exchange between field-level devices and enterprise applications, such as Microsoft Excel. Such solution includes OPC UA as secure communication protocol between the field-level control devices and the AASX. The AASX data format serves as an intermediate step between enterprise applications and the field device. In the final step, the AASX file would be converted to an Excel file using a web converter tool.

Löcklin et al. (2021) derive from the AAS concept the idea of the "Data Administration Shell", which consists of an administration shell dedicated to manage, describe and exchange information used exclusively for data-driven and data science approaches, i.e. datasets, training data, and others. The motivation behind the concept is to improve and facilitate the exchange of information between the companies that have access to the data, and the companies or individuals that can analyze it. A study on the integration of embedded systems into I4.0 networks is conducted by Pribiš et al. (2021) with the use of OPC UA and AAS. The authors describe the implementation of the AAS in the OPC UA information model as an approach to allow hardware with limited resources, such as microcontrollers, to act as a host of an AAS instance.

Schäfer et al. (2021) expose a methodology to deploy AAS with a Programmable Logic Controller (PLC). In the experiment conducted by the authors, the PLC is able to interface with an automatically generated AAS previously stored in a database. The main goal of the research is to convert the plant with PLCs into a reconfigurable plant design by using the AAS.

A solution to improve interoperability between DTs is introduced by Platenius-Mohr et al. (2020), presenting a method to convert the information models of different DT representations. As a use case, the authors validate the approach by enabling file and API-based information exchange between ABB Ability (a proprietary DT solution from ABB) and the AAS format in both directions. Arm et al. (2021) design a web-based, user-friendly configuration wizard to accelerate the creation of AAS standardized instances, properties and submodels. The created AAS is validated using a case study of a virtual assembly line.

Cavalieri and Salafia (2020a) provided extensive considerations to define OPC UA information models that would be used to expose the contents of the AAS and other information relevant to I4.0 specific domains of interest. The goal of the research is to improve the interoperability of the AAS by facilitating the information exchange between industrial applications using OPC UA. Lang et al. (2019) introduce the concept of a maintenance model in the AAS, which aims to support human operators during the maintenance process. The submodel includes the detailed procedure for carrying out maintenance and a list of the necessary requirements. i.e. materials, equipment, safety concerns and others. Lastly, Cavalieri and Salafia (2020b) introduce an architecture of a DT for PdM based on the AAS. The paper also references the use of ML algorithms for PdM without specifying the ML model used. In general, the main contribution of this paper is to describe an architecture that uses the AAS and enables the PdM.

The integration of the AAS with PdM solutions, an important topic of this research, remains however underdeveloped. As it is observed early during the implementation of the search queries in Sect. 2, the topic of PdM enabled by the AAS is not covered extensively in the search results. From the collected publications, only the study from Cavalieri and Salafia (2020b) addresses this particular topic. Other works for example Castellani et al. (2021) and Xu et al. (2019) integrate data-driven methods and DT for fault diagnosis applications, but none of these approaches implement the AAS.

Use case: the WaVe project

The following section aims to explain the conditions and requirements of WaVe, the project that motivates this research study. The content of this section is twofold, the first part serves as an introduction to the project and its motivations, and the second part explains the general requirements and challenges.

Project background

WaVe is a joint project promoted by the German Federal Ministry for Economics Affairs and Climate Action (BMWK), whose objective is to develop and test environmentally friendly hydrogen-based drive systems for commercial vehicles in the medium-duty segment. The development and manufacturing of electric vehicles is increasing rapidly with each passing year. Germany's annual reports from 2021 and 2022 reveal that approximately 25% of the newly registered vehicles for personal use were electric (Kraftfahrt-Bundesamt, 2022). Additionally, Germany is currently investing on researching H_2-driven engines to reduce the carbon footprint, since this technology shares similarities with the typical internal combustion engine, a mature, well-studied and robust technology.

The stages of WaVe cover the entire design cycle of the internal combustion engine, spanning from the mechanical design, testing and optimization of the H_2 engine to the research and development of the vehicle's hydrogen supply. The initiative gathers the expertise of 19 industrial partners to cover each aspect of the project and its requirements. The ultimate goal of WaVe is to develop and test two medium-duty vehicle prototypes with hydrogen-based drive systems: the Mercedes-Benz Unimog U400l and an unmanned tracked vehicle (Commercial Vehicle Cluster-Nutzfahrzeug GmbH, 2021). *Comlet verteilte Systeme GmbH* is one of the companies taking part in the WaVe project and is currently developing a DT solution to support both of the above mentioned prototypes.

Requirements

The WaVe's prototypes will be using hydrogen as fuel, which dissipates effortlessly in thin air, making fuel leaks a primary concern to be prevented. It is therefore essential to monitor with special care the variables that describe and influence the behaviour of the system during operation, in order to determine whether the vehicles are performing according to the manufacturer's specifications and, more importantly, to reduce any risk to the users.

Data-driven PdM offers the possibility of extracting complex patterns from high-dimensional data to evaluate the overall state of the asset, detect anomalies and indicate when faults will occur. However, these approaches require large volumes of data for model training and ultimately to diagnose the vehicle's condition and evaluate its operation. Ensuring data availability is one of the challenges when implementing PdM or machine learning models. Therefore, it is key to guarantee large datasets to better tune the accuracy of predictive maintenance algorithms. The DT will serve as an information hub, providing the available information of the asset and its underlying subsystems at a high level of detail using secure data channels.

In theory, the information contained in the DT is not limited, as any type of data can be integrated into the DT. The information in a DT can be divided into two groups: static and dynamic data. Static data includes documentation files, instruction manuals, datasheets, serial numbers, CAD models, and others. As the name suggests, this data tends to remain unchanged. In contrast, dynamic data refers to the DT's data that are changing or updated regularly, such as sensor readings and software updates. Since the data availability can be guaranteed using a DT solution, it is reasonable to consider using the DT and PdM as an integrated package.

Solution architecture

The following section describes a proposal for an architecture of a DT integrated with a PdM workflow. The considerations adopted for this proposal are motivated by the conclusions of the results discussed in Sect. 3 and adapted to the conditions and goals of the WaVe project in Sect. 4. Since there was insufficient data from the WaVe project prototypes at the time of this research, a solution architecture for a local test platform is designed first, considering that the features of this architecture can be transferred and adapted to a new one in a future work, as the project progresses and data from the prototypes become available.

Requirements

The results of the review in Sect. 3 revealed a research gap in the integration of PdM functionalities into AAS-frameworks. As a further development of this topic, an integration of the AAS into a typical machine learning pipeline for PdM is introduced.

The following requirements constitute the key elements of the solution architecture:

1. Incorporation of the AAS as standardized implementation of the DT
2. Support of the data pipeline
3. Binding of points 1) and 2) to guarantee automatic data flow between AAS and PdM pipeline

4. Support for additional services to display the asset's information in different formats.

Architecture overview

As a starting point, the proposed architecture follows the five-dimension DT model introduced by Tao et al. (2018) for health management of wind turbines. An overview of both architectures can be seen in Figs. 5 and 6.

The elements shown in Fig. 5 represent the five dimensions of this framework, where *PE* stands for the physical entity, *VE* for the virtual entity, *CN* represent the connection between each element, *SS* stands for the services, and in the middle is the data element represented by the letters *DD*. The modified 5D architecture shown in Fig. 6 is oriented towards the services layer, which serves as an interface between the DT and the physical asset as well as the outside world. In this sense, the virtual elements remain encapsulated within the DT, simplifying the communication model with the physical asset by limiting the communication points. Regarding security, this approach allows better monitoring of the data traffic and the access management of the data model.

Following this approach, the services are also redefined as standalone applications that work on top of the stored data. For example, in Fig. 6, the preprocessing service can filter the raw data, obtain the frequency spectrum and other statistical features for ML purposes. The PdM service receives raw or preprocessed data as input and outputs an estimation of the state of the asset followed by a warning message. The simulation service executes a model of the system according to the stored parameters and delivers the simulated data as output. It is worth noting that the total number of services may vary depending on the nature of the data and the asset itself.

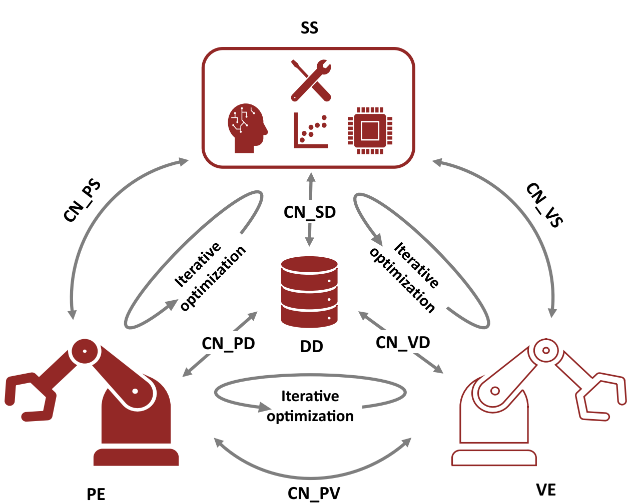

Fig. 5 5D architecture based on model from Tao et al. (2018)

Integrating the DT with the PdM pipeline

An illustration of the PdM pipeline and the DT is depicted in Fig. 7. As the figure suggests, the physical asset acts as a separate entity, which sends the sensor data to the AAS using the communication interfaces supported by the AAS standard, e.g. OPC UA and REST. Next, the received data is stored in a database for further use. The AAS acts as the central element of this scheme, serving as a data hub for PdM and other related applications. The submodel instances are used to store the properties of the asset and the application accesses the submodel to provide the service related to it. In this sense, the predefined properties inside the submodels are accessed by the services to carry out their respective tasks. Additionally, the unique identifier that every submodel has makes every property of the asset indexable, allowing to find the correct information source for the corresponding service in spite of the presence of a complex structure of submodel instances.

The PdM pipeline can benefit from this structure by accessing the submodels to automate the data flow. In Fig. 7 the PdM pipeline is illustrated as a 4-stage process: feature extraction, training, prediction and maintenance. In this example, the first two stages encompass the iterative process of training and tuning the prediction model. The feature extraction can use the simulation submodel to trigger a simulation and obtain the output data for training purposes. The historical data submodel can be accessed in a similar fashion to query a batch of data previously stored in the database of the DT. Additionally, a preprocessing submodel can be used to extract information about the necessary filtering, normalization and other necessary precalculations that are desirable to apply to the data prior to training. An additional submodel could be added to specify training parameters such as the learning rate and the number of epochs, leaving more room for the customization of the pipeline through the AAS.

Once the model is trained, the live data submodel and its application program are used to feed the machine learning model with new data to perform fault diagnosis and/or RUL estimation. Finally, the maintenance block contains the decision logic to generate the corresponding warning messages and send them back to the test platform, depending on the outcome of the previous phase.

Regarding the implementation aspect of the architecture, there is an extensive collection of libraries and frameworks for developing data-driven pipelines. The paper by Nguyen et al. (2019) provides a survey of the available frameworks for machine and deep learning and compares them in terms of their advantages and disadvantages. According to this study, the Python programming language is a popular choice for ML applications, as many tools are developed in Python or have support for its interfaces. Abadi et al. (2015) and Paszke et al. (2019) are both examples of powerful open source libraries

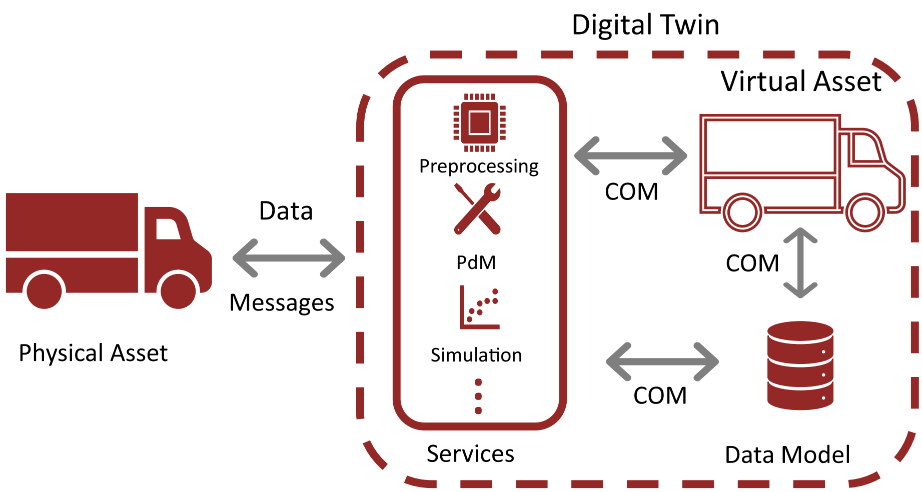

Fig. 6 Proposed 5D DT Model

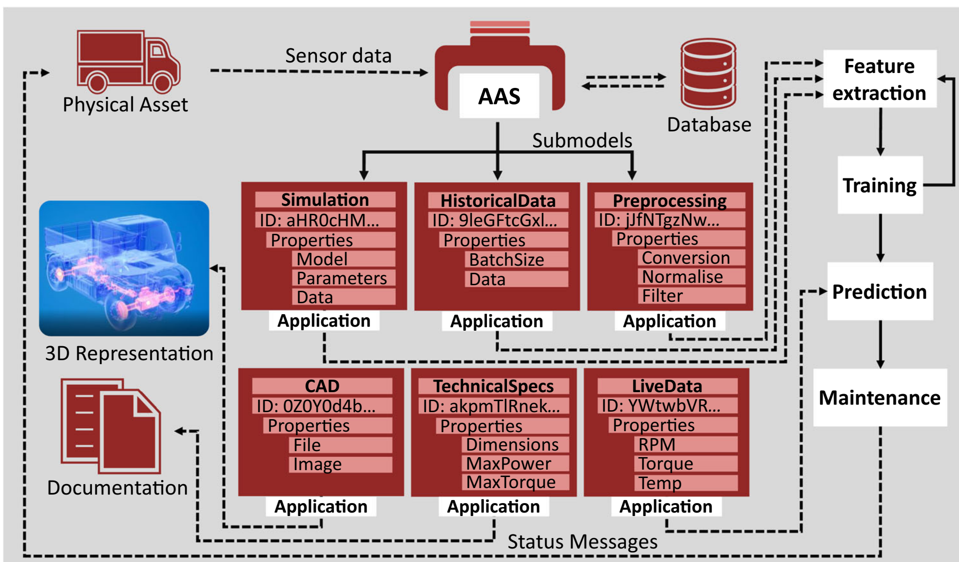

Fig. 7 Proposed DT model for the WaVe project

used for developing and training ML models that are also available in this programming language.

Other Python libraries such as McKinney (2010) and Harris et al. (2020) are essential libraries for the data analysis and computation that can be combined with the above mentioned modules to cover every stage of the PdM pipeline. In addition, the implementations of OPC UA and REST interfaces are also supported in Python through their corresponding open source libraries. Therefore, Python can be considered as a versatile software toolkit for the implementation and deployment of the proposed architecture.

Ultimately, it is expected that the data generated by the DT will consist of multiple sensor measurements represented as time series, which in this case highlights the relevance of using RNN-based models. During the literature review it was found that RNNs and its derived models are widely used for time series processing and sequence prediction. Therefore, models such as LSTM and GRU can be a key element in this architecture for the prediction of faults and RUL. Hybrid architectures that combine other models such as CNNs and MLPs can also be considered. An example use case would be to place the CNN in the input stage for feature reduction and the MLP in the output stage to perform more complex predictions from a set of predicted signals.

Other applications and limitations

As seen in the previous section, the submodels can be used for a wide range of use cases. Submodels containing static data can also serve for visualization and documentation to assist users and operators. CAD and technical specification submodels, as shown in Fig. 7, can include 3D models and documentation files, which can be displayed with the help of the corresponding application programs for monitoring and evaluation purposes.

Although this architecture is initially designed to support the asset during its operation, it is not limited to this single application. This solution architecture can also be implemented to assist the asset during other phases of its lifecycle, such as development and manufacturing, thanks to the flexibility offered by the AAS. For example, in the area of manufacturing, this solution architecture can be implemented to monitor a fabrication process, where the physical asset could be a production line or manufacturing equipment and the PdM would be implemented to monitor the production equipment and/or the product being manufactured.

It is worth highlighting that the AAS-based architecture for PdM described in this work relies fundamentally on two core elements: communication and the information model of the AAS. On the one hand, reliable and robust communication channels are needed to maintain the data flow and the stability of the services, otherwise the PdM pipeline cannot sustain operation. On the other hand, it is essential to support the latest features of the AAS standard while maintaining compatibility with legacy versions of it, since the AAS standard is still in development. In this manner, the compatibility and interpretability of submodels and other data structures of the AAS can be ensured by this architecture.

Conclusions

Nowadays PdM and the AAS have emerged as topics of interest in academia and industry due to their enabling capabilities. This paper explores the scientific works regarding both topics, with the aim to evaluate the state of the art of data-driven methods for PdM in industrial and automotive use cases and the research direction of AAS-related applications.

The results of the review reveal that out of a selection of 28 publications related to PdM, DCT-based models were used the most in fault diagnosis applications. In the prognostics area, the predominant approaches were based on Bayesian, RNN and DCT models. On the side of the AAS, a collection of 11 publications is reviewed. The topics related to the AAS prove to be diverse, showing promising research areas from the very implementation of AAS on hardware and embedded platforms to new definitions and insights to improve the standard. Therefore, a major achievement of this review is the identification of a research gap of a PdM strategy integrated with the AAS. This discovery motivated the introduction of a solution architecture that supports PdM for the monitoring of the hydrogen-driven vehicle and the AAS for developing DTs.

The implementation of the proposed architecture will be the next step of this line of work. The results and the insights gained in this review revealed that RNN-based architectures are relevant for time series forecasting and therefore were found to be suitable for prognostics and fault diagnosis.

Acknowledgements This work was supported by the German Federal Ministry for Economic Affairs and Climate Action (BMWK), under Grant No. (19I21028R), research project WaVe. The authors alone are responsible for the content of the paper.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data availability Data sharing not applicable to this article as no datasets were generated during the current study.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., & Zheng. X. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. https://doi.org/10.5281/zenodo.4724125

Arm, J., Benesl, T., Marcon, P., Bradac, Z., Schröder, T., Belyaev, A., et al. (2021). Automated design and integration of asset administration shells in components of industry 4.0. *Sensors, 21*(6), 2004. https://doi.org/10.3390/s21062004

Bader, S., Barnstedt, E., Bedenbender, H., Berres, B., Billmann, M., & Ristin, M. (2022, March). Details of the asset administration shell - part 1: The exchange of information between partners in the value chain of industrie 4.0 (version 3.0rc02). *Technical report, Federal Ministry for Economic Affairs and Climate Action* (BMWK). https://doi.org/10.21256/zhaw-27075

Booyse, W., Wilke, D. N., & Heyns, S. (2020). Deep digital twins for detection, diagnostics and prognostics. *Mechanical Systems and Signal Processing, 140*, 106612. https://doi.org/10.1016/j.ymssp.2019.106612

Carretero, J., & García, J. D. (2013). The internet of things: Connecting the world. *Personal and Ubiquitous Computing, 18*(2), 445–447. https://doi.org/10.1007/s00779-013-0665-z

Castellani, A., Schmitt, S., & Squartini, S. (2021). Real-world anomaly detection by using digital twin systems and weakly supervised learning. *IEEE Transactions on Industrial Informatics, 17*(7), 4733–4742. https://doi.org/10.1109/tii.2020.3019788

Cavalieri, S., & Salafia, M. G. (2020). Insights into mapping solutions based on OPC UA information model applied to the industry 4.0 asset administration shell. *Computers, 9*(2), 28. https://doi.org/10.3390/computers9020028

Cavalieri, S., & Salafia, M. G. (2020). A model for predictive maintenance based on asset administration shell. *Sensors, 20*(21), 6028. https://doi.org/10.3390/s20216028

Chen, C., Liu, Y., Sun, X., Cairano-Gilfedder, C. D., & Titmus, S. (2021). An integrated deep learning-based approach for automobile maintenance prediction with GIS data. *Reliability Engineering and System Safety, 216*, 107919. https://doi.org/10.1016/j.ress.2021.107919

Commercial Vehicle Cluster-Nutzfahrzeug GmbH. (2021). Verbundvorhaben "WaVe" - Entwicklung und Prototypische Erprobung von Wasserstoffverbrennungsmotoren. Retrieved January 11, 2023 from https://wave.cvc-suedwest.com/wp-content/uploads/2021/08/CVC_News_1_2021_Verbundvorhaben_WaVe.pdf

Fioravanti, R., Kumar, K., Nakata, S., Chalamala, B., & Preger, Y. (2020). Predictive-maintenance practices: For operational safety of battery energy storage systems. *IEEE Power and Energy Magazine, 18*(6), 86–97. https://doi.org/10.1109/mpe.2020.3014542

Giobergia, F., Baralis, E., Camuglia, M., Cerquitelli, T., Mellia, M., Neri, A., et al. (2018), October. Mining sensor data for predictive maintenance in the automotive industry. In *2018 IEEE 5th international conference on data science and advanced analytics* (DSAA). IEEE. https://doi.org/10.1109/dsaa.2018.00046.

Giordano, D., Pastor, E., Giobergia, F., Cerquitelli, T., Baralis, E., Mellia, M., et al. (2021). Dissecting a data-driven prognostic pipeline: A powertrain use case. *Expert Systems with Applications, 180*, 115109. https://doi.org/10.1016/j.eswa.2021.115109

Gomes, I. P., & Wolf, D. F. (2020). Health monitoring system for autonomous vehicles using dynamic bayesian networks for diagnosis and prognosis. *Journal of Intelligent and Robotic Systems, 101*(1), 19. https://doi.org/10.1007/s10846-020-01293-y

Grieves, M., & Vickers, J. (2016, August). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary perspectives on complex systems* (pp. 85–113). Springer

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., Ã... Oliphant, T. E. (2020). September. Array programming with NumPy. *Nature, 585*(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Jung, D. (2020). Data-driven open-set fault classification of residual data using Bayesian filtering. *IEEE Transactions on Control Systems Technology, 28*(5), 2045–2052. https://doi.org/10.1109/tcst.2020.2997648

Kraftfahrt-Bundesamt (KBA). (2022). Neuzulassungen von personenkraftwagen (pkw) im jahresverlauf 2022 nach marken und alternativen antrieben. Retrieved October 07, 2022, from https://www.kba.de/SharedDocs/Downloads/DE/Pressemitteilungen/DE/2022/pm_36_2022_alternative_Antriebe_08_2022.pdf?__blob=publicationFile&v=5

Lang, D., Grunau, S., Wisniewski, L., & Jasperneite, J. (2019, July). Utilization of the asset administration shell to support humans during the maintenance process. In *2019 IEEE 17th international conference on industrial informatics* (INDIN). https://doi.org/10.1109/indin41052.2019.8972236

Lee, N., Azarian, M.H., Pecht, M., Kim, J., & Im, J. (2019). A comparative study of deep learning-based diagnostics for automotive safety components using a raspberry pi. In *2019 IEEE international conference on prognostics and health management* (ICPHM). https://doi.org/10.1109/ICPHM.2019.8819436

Li, L., Aslam, S., Wileman, A., & Perinpanayagam, S. (2022). Digital twin in aerospace industry: A gentle introduction. *IEEE Access, 10*, 9543–9562. https://doi.org/10.1109/access.2021.3136458

Luo, B., Wang, H., Liu, H., Li, B., & Peng, F. (2018). Early fault detection of machine tools based on deep learning and dynamic identification. *IEEE Transactions on Industrial Electronics, 66*(1), 509–518. https://doi.org/10.1109/TIE.2018.2807414

Löcklin, A., Vietz, H., White, D., Ruppert, T., Jazdi, N., & Weyrich, M. (2021). Data administration shell for data-science-driven development. *Procedia CIRP, 100*, 115–120. https://doi.org/10.1016/j.procir.2021.05.019

Manghai, A., & Jegadeeshwaran, R. (2019). Application of FURIA for finding the faults in a hydraulic brake system using a vibration analysis through a machine learning approach. *International Journal of Prognostics and Health Management, 10*(1), 3. https://doi.org/10.36001/ijphm.2019.v10i1.2748

Mckinley, T., Somwanshi, M., Bhave, D., & Verma, S. (2020). Identifying nox sensor failure for predictive maintenance of diesel engines using explainable AI. In *PHM Society European conference*, Vol. 5. https://doi.org/10.36001/phme.2020.v5i1.1267

McKinney, W. (2010). Data structures for statistical computing in Python. In *Proceedings of the 9th Python in science conference*, pp. 56 – 61. https://doi.org/10.25080/Majora-92bf1922-00a

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., García, Á. L., Heredia, I., Malík, P., & Hluchý, L. (2019). Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey. *Artificial Intelligence Review, 52*(1), 77–124. https://doi.org/10.1007/s10462-018-09679-z

Nixon, S., Weichel, R., Reichard, K., & Kozlowski, J. (2018). A machine learning approach to diesel engine health prognostics using engine controller data. In *Proceedings of the annual conference of the Prognostics and Health Management Society*, PHM, Vol. 10. https://doi.org/10.36001/phmconf.2018.v10i1.587

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D. et al. (2021, March). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ: n71. https://doi.org/10.1136/bmj.n71

Pan, H., Lü, Z., Wang, H., Wei, H., & Chen, L. (2018). October. Novel battery state-of-health online estimation method using multiple health indicators and an extreme learning machine. *Energy, 160*, 466–477. https://doi.org/10.1016/j.energy.2018.06.220

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J.,& Chintala, S. (2019). Pytorch: An imperative style, high-performance deep learning library. https://doi.org/10.48550/arXiv.1912.01703.

Pech, M., Vrchota, J., & Bednář, J. (2021). Predictive maintenance and intelligent sensors in smart factory. *Sensors, 21*(4), 1470. https://doi.org/10.3390/s21041470

Platenius-Mohr, M., Malakuti, S., Grüner, S., Schmitt, J., & Goldschmidt, T. (2020). December. File- and API-based interoperability of digital twins by model transformation: An IIoT case study

using asset administration shell. *Future Generation Computer Systems, 113*, 94–105. https://doi.org/10.1016/j.future.2020.07.004
Pribiš, R., Beňo, L., & Drahoš, P. (2021). Asset administration shell design methodology using embedded OPC unified architecture server. *Electronics, 10*(20), 2520. https://doi.org/10.3390/electronics10202520
Rengasamy, D., Jafari, M., Rothwell, B., Chen, X., & Figueredo, G. P. (2020). Deep learning with dynamically weighted loss function for sensor-based prognostics and health management. *Sensors, 20*(3), 723. https://doi.org/10.3390/s20030723
Sass, A. U., Esatbeyoğlu, E., & Iwwerks, T. (2020). Data-driven powertrain component aging prediction using in-vehicle signals. In *SOFSEM*.
Schäfer, S., Schöttke, D., Kämpfe, T., Ralinovski, K., Tauber, B., & Lehmann, R. (2021). Design and deployment of digital twins for programmable logic controllers in existing plants. In *Proceedings of the 2nd international conference on innovative intelligent industrial production and logistics*, Vol. 2022, pp. 145–150. SCITEPRESS - Science and Technology Publications. https://doi.org/10.5220/0010711000003062
Shafi, U., Safi, A., Shahid, A. R., Ziauddin, S., & Saleem, M. Q. (2018). Vehicle remote health monitoring and prognostic maintenance system. *Journal of Advanced Transportation, 2018*, 1–10. https://doi.org/10.1155/2018/8061514
Soltanipour, N., Rahrovani, S., Martinsson, J., & Westlund, R. (2020, September). Chassis hardware fault diagnostics with hidden markov model based clustering. In *2020 IEEE 23rd international conference on intelligent transportation systems* (ITSC), pp. 1–6. IEEE. https://doi.org/10.1109/itsc45102.2020.9294468
Sun, Y., Yu, W., Chen, Y., & Kadam, A. (2019), October. Time series anomaly detection based on GAN. In *2019 sixth international conference on social networks analysis, management and security* (SNAMS), pp. 375–382. IEEE. https://doi.org/10.1109/snams.2019.8931714
Tao, F., Zhang, M., Liu, Y., & Nee, A. (2018). Digital twin driven prognostics and health management for complex equipment. *CIRP Annals, 67*(1), 169–172. https://doi.org/10.1016/j.cirp.2018.04.055
Tessaro, I., Mariani, V. C., & dos Santos Coelho, L. (2020, November). Machine learning models applied to predictive maintenance in automotive engine components. In *Proceedings of 1st international electronic conference on actuator technology: Materials, devices and applications*, Vol. 64. MDPI. https://doi.org/10.3390/iecat2020-08508
Theissler, A., Pérez-Velázquez, J., Kettelgerdes, M., & Elger, G. (2021). Predictive maintenance enabled by machine learning: Use cases and challenges in the automotive industry. *Reliability Engineering & System Safety, 215*, 107864. https://doi.org/10.1016/j.ress.2021.107864
Thoben, K. D., Wiesner, S., & Wuest, T. (2017). "Industrie 4.0" and smart manufacturing-a review of research issues and application examples. *International Journal of Automation Technology, 11*(1), 4–16. https://doi.org/10.20965/ijat.2017.p0004
von Schleinitz, J., Graf, M., Trutschnig, W., & Schröder, A. (2021). VASP: An autoencoder-based approach for multivariate anomaly detection and robust time series prediction with application in motorsport. *Engineering Applications of Artificial Intelligence, 104*, 104354. https://doi.org/10.1016/j.engappai.2021.104354
Wang, Y., Cao, Y., & Wang, F. Y. (2021). Anomaly detection in digital twin model. In *2021 IEEE 1st international conference on digital twins and parallel intelligence* (DTPI) (pp. 208–211). https://doi.org/10.1109/DTPI52967.2021.9540116
Wang, Y., Liu, N., Guo, H., & Wang, X. (2020). An engine-fault-diagnosis system based on sound intensity analysis and wavelet packet pre-processing neural network. *Engineering Applications of Artificial Intelligence, 94*, 103765. https://doi.org/10.1016/j.engappai.2020.103765
Wolf, P., Mrowca, A., Nguyen, T. T., Baker, B., & Gunnemann, S. (2018, November). Pre-ignition detection using deep neural networks: A step towards data-driven automotive diagnostics. In *2018 21st international conference on intelligent transportation systems* (ITSC), pp. 176–183. IEEE. https://doi.org/10.1109/itsc.2018.8569908
Wu, D., Jennings, C., Terpenny, J., Gao, R. X., & Kumara, S. (2017). A comparative study on machine learning algorithms for smart manufacturing: Tool wear prediction using random forests. *Journal of Manufacturing Science and Engineering, 139*(7), 071018. https://doi.org/10.1115/1.4036350
Xu, Y., Sun, Y., Liu, X., & Zheng, Y. (2019). A digital-twin-assisted fault diagnosis using deep transfer learning. *IEEE Access, 7*, 19990–19999. https://doi.org/10.1109/access.2018.2890566
Ye, X., & Hong, S. H. (2019). Toward industry 4.0 components: Insights into and implementation of asset administration shells. *IEEE Industrial Electronics Magazine, 13*(1), 13–25. https://doi.org/10.1109/mie.2019.2893397
Ye, X., Hong, S. H., Song, W. S., Kim, Y. C., & Zhang, X. (2021). An industry 4.0 asset administration shell-enabled digital solution for robot-based manufacturing systems. *IEEE Access, 9*, 154448–154459. https://doi.org/10.1109/ACCESS.2021.3128580
Ye, X., Song, W. S., Hong, S. H., Kim, Y. C., & Yoo, N. H. (2022). Toward data interoperability of enterprise and control applications via the industry 4.0 asset administration shell. *IEEE Access, 10*, 35795–35803. https://doi.org/10.1109/access.2022.3163738
Zhang, B., Zhang, S., & Li, W. (2019). April. Bearing performance degradation assessment using long short-term memory recurrent network. *Computers in Industry, 106*, 14–29. https://doi.org/10.1016/j.compind.2018.12.016
Zhang, S., Dong, H., Maschek, U., & Song, H. (2021). A digital-twin-assisted fault diagnosis of railway point machine. In *2021 IEEE 1st international conference on digital twins and parallel intelligence* (DTPI) (pp. 430–433).
Zhang, W., Yang, D., & Wang, H. (2019). Data-driven methods for predictive maintenance of industrial equipment: A survey. *IEEE Systems Journal, 13*, 2213–2227. https://doi.org/10.1109/JSYST.2019.2905565
Zhao, R., Wang, D., Yan, R., Mao, K., Shen, F., & Wang, J. (2018). Machine health monitoring using local feature-based gated recurrent unit networks. *IEEE Transactions on Industrial Electronics, 65*(2), 1539–1548. https://doi.org/10.1109/tie.2017.2733438
Zhu, R., Chen, Y., Peng, W., & Ye, Z. S. (2022). Bayesian deep-learning for RUL prediction: An active learning perspective. *Reliability Engineering and System Safety, 228*, 108758. https://doi.org/10.1016/j.ress.2022.108758
Zonta, T., da Costa, C. A., da Rosa Righi, R., de Lima, M. J., da Trindade, E. S., & Li, G. P. (2020). Predictive maintenance in the industry 4.0: A systematic literature review. *Computers and Industrial Engineering, 150*, 106889. https://doi.org/10.1016/j.cie.2020.106889

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.