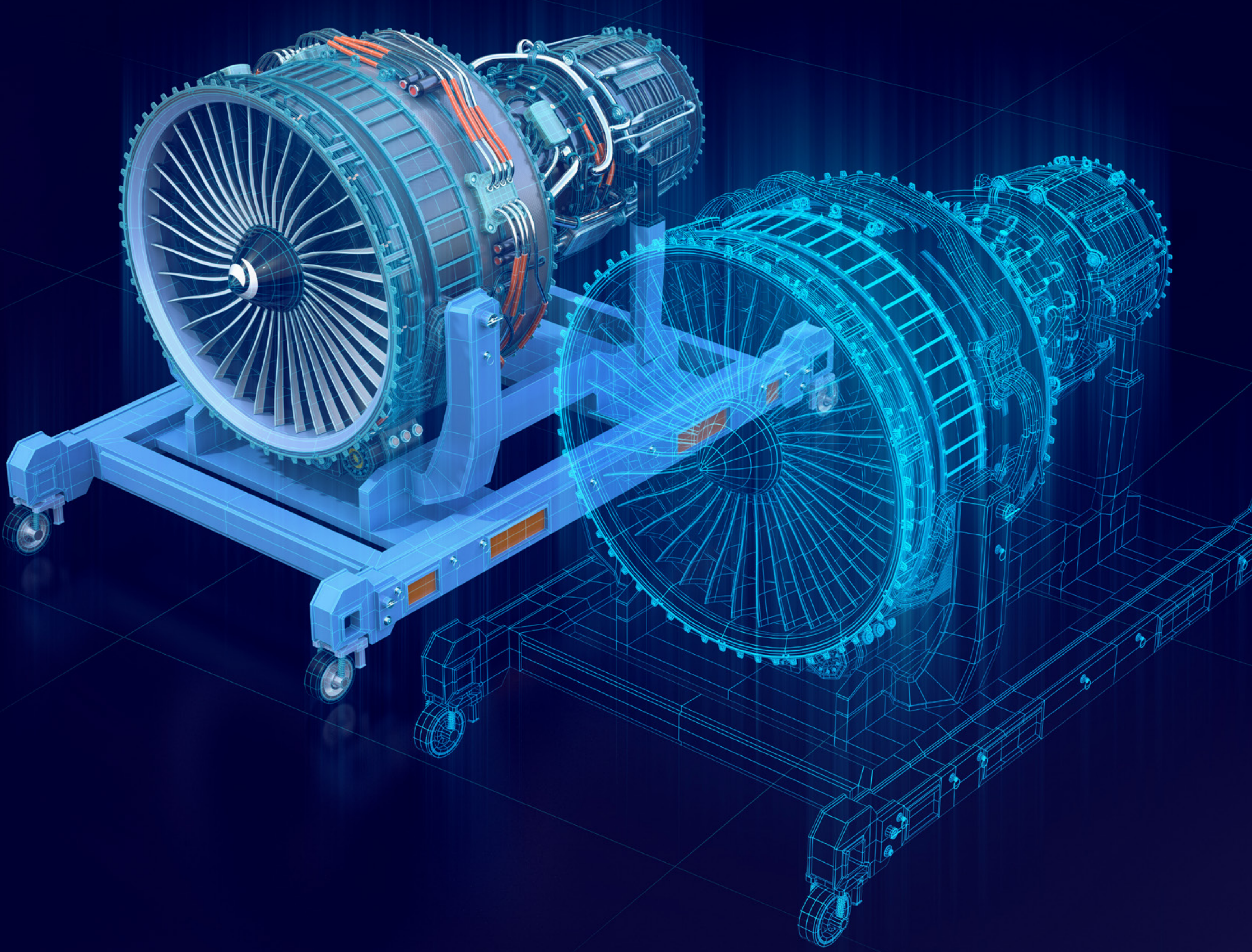
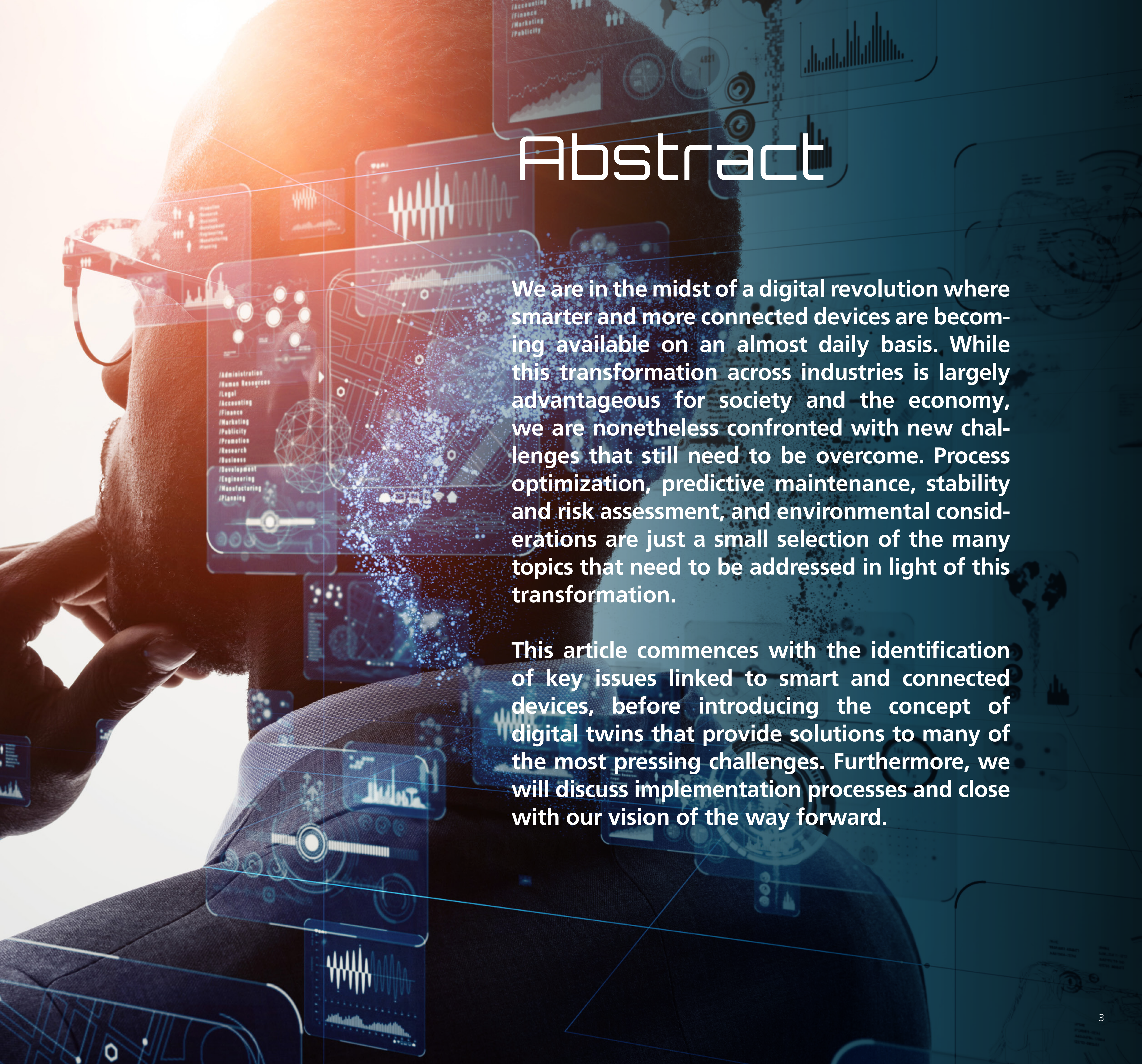


MHPWHITE PAPER

NEW PERSPECTIVES

How AI-based Digital Twins will Shape our Future

Content



Abstract

We are in the midst of a digital revolution where smarter and more connected devices are becoming available on an almost daily basis. While this transformation across industries is largely advantageous for society and the economy, we are nonetheless confronted with new challenges that still need to be overcome. Process optimization, predictive maintenance, stability and risk assessment, and environmental considerations are just a small selection of the many topics that need to be addressed in light of this transformation.

This article commences with the identification of key issues linked to smart and connected devices, before introducing the concept of digital twins that provide solutions to many of the most pressing challenges. Furthermore, we will discuss implementation processes and close with our vision of the way forward.

Introduction

The world of machines is becoming increasingly smarter and more connected. Devices continue to operate interconnected, if not already entirely autonomously, while interacting with other devices and their environment in a complex fashion. This progress is visible in day to day life when comparing a phone, laptop, or television from ten years ago with current models or when comparing past and current production processes in the manufacturing industry. The latter especially applies with regard to products such as cars or electronics. Devices have inevitably become more complex, autonomous and self-organized, which will continue to be the trend for years to come as the use of AI continues to rise.

While technological advancements open a wide range of exciting opportunities in many areas, society, as well as most businesses, significantly benefit from many – though not all – new developments in this field. However, we regularly face different and formerly unseen challenges that need to be overcome. Most of these challenges are directly linked to the large and ever-growing number of personal devices, their inherent complexity, the convolution of their interactions, and the increasing level of integration that is enabled by progressing technologies and digitalization, as the following examples illustrate.

Process optimization

The complexity of a system of interacting devices inevitably scales with the number of independently acting entities, i.e., the number of smart and connected devices in a particular system. While this may take the burden off having to tediously engineer respective processes from a top-down perspective and in much detail, such self-organization comes at the price of reduced insight into how individual tasks are executed and controlled. The system acts as a black box to some extent. Direct control of processes – and especially how economically these processes can be executed – is limited or entirely lost.

The challenge of process optimization can be illustrated by a typical example in manufacturing, as follows: Imagine a smart warehouse operating a fleet of autonomously self-driving vehicles that transport machine parts to an attached factory for assembly in sequence and in time. With growing numbers of individual items being delivered, parts may not reach the production line on time and cost-intensive delays could ensue. An obvious and straightforward way to tackle this issue is to simply increase the number of vehicles. However, while this is likely to reduce delays, it will not solve the underlying problem of the way the process is organized. By doing so, the problem is not approached with the intent to eliminate the root cause, but only to postpone it to a later date when it will have to be solved with far greater effort and, likely, at a higher cost.

Predictive maintenance

With devices becoming smarter and more autonomous on one hand and featuring greater intrinsic complexity and higher levels of integration on the other, the need for human interference and oversight is greatly reduced in a purely operational context. This can be exemplified by the above case of a smart warehouse which only requires minimal human oversight, or the example of a modern car compared to a car from one hundred years ago. Back then, a trained mechanic was required to maintain the car on a regular and frequent basis; nowadays cars do not need to be serviced as frequently, despite the multitude of functions present in modern cars compared to older models, especially when taking into account standard features such as infotainment systems, air conditioning, and so on.

While reducing human oversight, the elevated intrinsic complexity of devices and device longevity certainly comes with great advantages in many cases, but maintenance becomes more challenging and potentially cost-intensive. Firstly, with little oversight, defects may either initially go unnoticed and result in higher costs when detected at a later stage, or unnecessary costs may potentially be incurred through too-frequent check-ups. Secondly, it may be challenging or even impossible to test and service certain features of a modern integrated device. In the latter context for instance, Li-ion batteries are difficult to repair if broken, and measuring the exact state-of-health of a battery in operation remains difficult.

Stability and risk assessment

Quite rightly, highly interconnected systems give rise to stability concerns and demand comprehensive risk assessments. With greater self-organization in smart and autonomous systems, even small deviations from standard scenarios may trigger chain reactions with results that potentially have not been anticipated. Another concern is the vulnerability of a system to cyber-attacks.

In the case of our exemplary smart factory, one may ask what happens if a self-driving vehicle fails to operate? What happens if two vehicles fail simultaneously? Answering these questions is more challenging in a modern production context than it was only few years ago.

Environmental considerations

Last but certainly not least, the consideration of environmental aspects is vital, though often complex, in highly interconnected and autonomous systems that are designed in a bottom-up fashion. We have seen above that optimizing processes in regard to economic and efficiency targets may be difficult due to a lack of insight into process details. In the context of environmental targets, such optimization processes may even be more challenging as it is difficult to gage success. It is inarguably more difficult to quantify the carbon emissions of a production process in a cradle-to-grave-type fashion than to measure the respective return of an investment or speed of production.

In our example of a smart warehouse, we saw that processes can be challenging to optimize concerning economic or efficiency targets. However, an environmental analysis might include finding the environmentally optimal usage of the batteries used in self-driving vehicles, which is a significantly more demanding task compared to optimizing the operating efficiency of vehicles.

We emphasize that these are just four examples of challenges that arise from digital transformation through the use of smart and connected devices, and associated issues are more diverse.

The following section discusses the concept of a digital twin, a relatively new concept that can easily be exploited to tackle the above problems in many ways. Please note that we have briefly mentioned the concept of digital twins in the context of the quality control of complex systems elsewhere [1]. A comprehensive analysis of the opportunities that digital twins offer with regard to batteries in electric mobility can be found in prior literature [2].

The digital twin – an overview

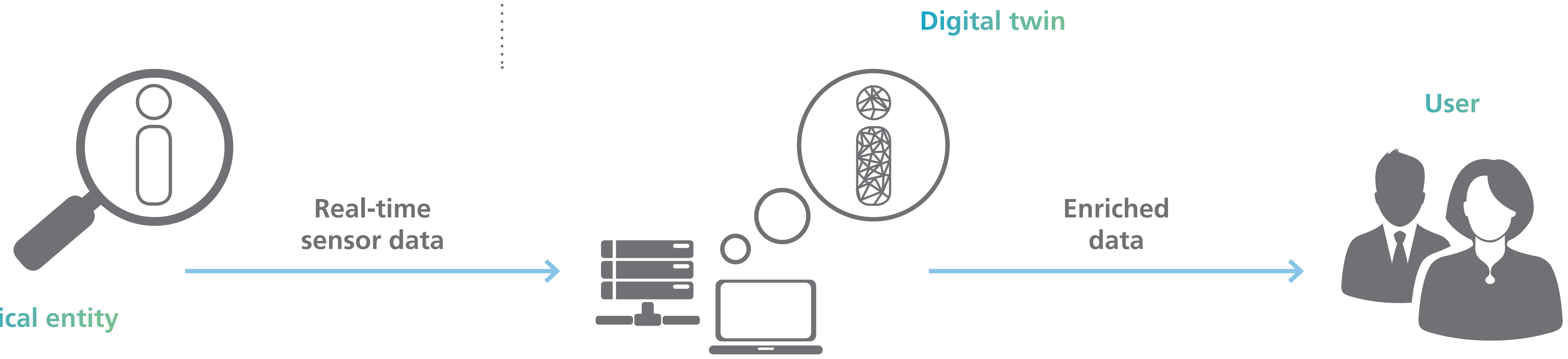


Figure 1: Basic structure of a digital twin. Please note that in application, the information exchange between the different entities can be significantly more comprehensive and complex as insight gained by the user through the digital twin may be fed back into the operation of the physical entity and the digital twin.

The idea of the digital twin was largely developed by Michael Grieves and others in the early 2000s. On the basis of real-time data, it aims to create real-time representations of physical entities that reflect features of interest without the actual physical entity having to be analyzed. Therefore, the digital twin can be understood to be a virtual counterpart of a real-life entity which mirrors its state and operation. In 2016 Grieves and Vickers [3] defined the digital twin as follows:

"The Digital Twin is a set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level. At its optimum, any information that could be obtained from inspecting a physical manufactured product can be obtained from its Digital Twin."

Grieves and Vickers then proceed to distinguish digital twin prototypes and instances, which we will briefly discuss in the narrow context below. It is important to differentiate between a digital twin according to this definition and a simple collection of data representing a physical entity. The interested reader should refer to the original study [3] for full details.

A more comprehensive literature review reveals that various other definitions exist by other authors, which are similar, and typically comprise three key elements as illustrated in Figure 1, which are outlined below.

The physical entity

In the context of the digital twin approach, a physical entity of interest is mirrored by the digital twin. This entity may feature any physical dimensions and any level of complexity, so there are no limitations regarding the nature of the modeled entity. The sole requirement is that sufficient data is gathered to describe the entity features to be reflected, i.e., the data recorded must correlate with these features in a way so that the desired information can be inferred. If a statistically relevant number of similar digital twins exists in a particular context, even future physical entities may be modeled, if desired.

The digital twin

The digital twin is a model of a physical entity that lives in the virtual world. It captures one or multiple relevant features of its counterpart and provides the means to share information of interest with the person operating it. The shared information can then be exploited in a broad range of contexts, as discussed below. A digital twin is herein always an abstraction of its real-world counterpart as its features can only be represented to a limited degree of detail. The level of abstraction is mostly dependent on two factors: the availability of data and the purpose of the digital twin. While the first inevitably sets boundaries to the degree of detail that can be achieved with the digital twin and can be adjusted by collecting more or different data, the latter must be carefully evaluated for each case. Higher degrees of abstraction are generally preferable for simplicity and cost efficiency as long as the abstraction does not restrict the intended purpose.

Information exchange

It is possible that the most vital element is the continuous exchange of information between the physical entity and the digital twin. In order to be a true representation, changes in the physical entity must be considered in the digital twin as the digital twin is continuously updated. To this end, the physical entity is equipped with suitable sensors to monitor its activity and state as well as appropriate interfaces to share this information with the digital twin. In some cases, communication may be bidirectional to facilitate the gathering of only the most vital information where, for instance, bandwidth is a limiting factor.

The above definition is broadly applicable on all scales; It covers the modeling of an entire factory, for instance, or only a single, small element in the same factory. For clarity and focus, the scope of the following discussion is therefore limited to ensembles of similar physical entities that are each modeled individually by a respective individual digital twin. For example, this could be a fleet of cars with every car being represented by a digital twin. Hence, this approach is used as a tool to understand how an individual smart and connected device of which many instances exist, functions, interacts, and evolves across its entire life cycle (see Figure 2). In this context, we can leverage domain knowledge, information gathered across the entire ensemble of similar physical entities, and historical information that may have been recorded long before an individual physical entity was even produced in order to create even more accurate digital twins. In emerging technologies such as the Internet of Things

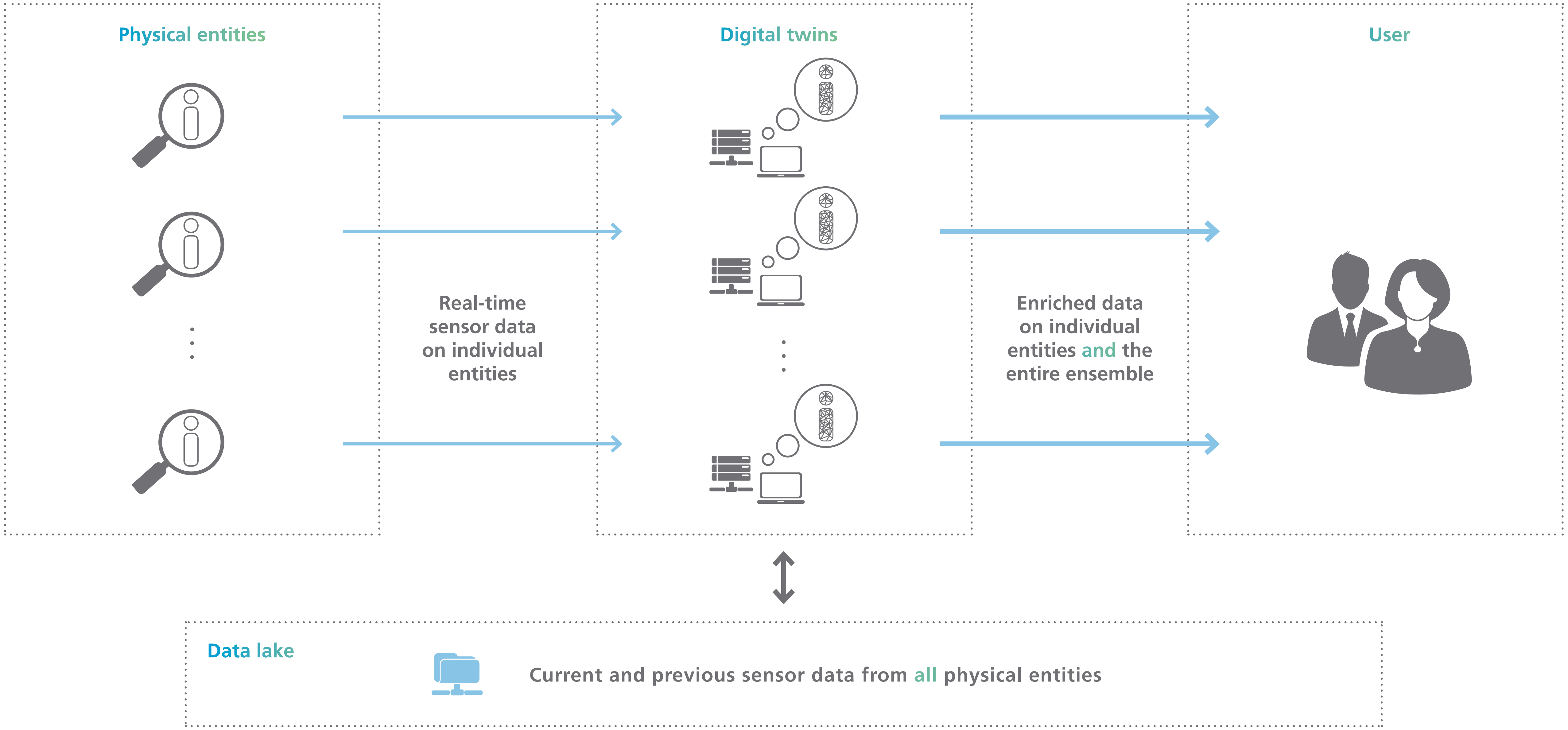


Figure 2: The structure of an ensemble of digital twins. Experience and acquired knowledge can be leveraged across the fleet by an information exchange between the individual digital twins that is enabled by a shared data lake.

(IoT) and the Industrial Internet of Things (IIoT), the concept of ensembles of digital twins has received overwhelming attention. And, despite its novelty, it is already broadly acknowledged as an important way forward for apparent reasons beyond those mentioned above. In Industry 4.0, devices are always connected, usually equipped with sensors – often smart – and mostly mass produced. The concept of a digital twin or ensembles can therefore often be readily applied without having to implement major changes, or any changes at all, to the hardware. In this context, the approach further unfolds its full potential regarding the modeling and planning of future scenarios. Once an understanding of the operation of an ensemble of entities has been achieved, the respective model can be easily extended to include future entities and may evolve into a powerful planning tool.

So far, we have defined the concept of a digital twin and ensembles of digital twins. In the following section we will take a closer look at the anatomy of the virtual model and compare and contrast different approaches.

Mathematics, artificial intelligence & digital twins

One of the greatest challenges in designing a digital twin is the development of a virtual model that matches the requirement of the given task. The model must accurately infer, or at least estimate, the desired information on the respective physical entity on the basis of the data available, however the model does not necessarily need to reflect this information in a straightforward fashion. To illustrate this challenge, let us contemplate two examples. First, let us consider an arc welding robot that assembles metal components in a factory and its digital twin. One of several purposes of the digital twin is to briefly pause the production process for recalibration if the temperature of the welding electrode falls below or exceeds certain temperature thresholds. To this end, the temperature is measured continuously at the electrode, transmitted to the digital twin, and an automated computational analysis of the digital twin triggers maintenance, if required. The digital twin receives the temperatures from the sensor and assigns its value to the internal representation of the electrode. In the second example, let us consider the same welding robot in a slightly different context. Over time, the electrode degrades and must be replaced, which the robot can do autonomously following the reception of a respective trigger. This trigger is received from the analysis tool which utilizes the digital twin of the robot. Unlike in the first example, the degradation status of the electrode cannot be measured directly but must be inferred from a large number of related measurements that are recorded by the robot. These may include measurements such as temperature, humidity, the number of joints welded, the rate of use, current densities, and others which are linked to the degradation status of the electrode in a complex fashion. In stark contrast to the first example, establishing a model is significantly more challenging

> **"Digital Twins are currently one of the most exciting technological trends – Interdisciplinary, innovative, and with the capacity for enormous business potential."**

Dr. Enno Kätelhön

as it requires a deep understanding of the physics of the welding process as well as empirical data.[1]

[1] Please note that the two examples above can be solved with classical machine learning approaches. The use of a digital twin is merely a change in perspective. Rather than "only" designing a function that interprets the robot sensor data to trigger events, a digital copy of the robot is built and interpreted. While there is no difference between the benefits of the two approaches in these simplified examples, the concept of digital twins unfolds its full potential in more complex scenarios as we will shortly see.

Various types of models can be chosen for digital twins, which all come with different advantages and disadvantages. Most notably, differences can be found with regard to model accuracy, model transparency, and the complexity of developing the model. In the following sections we compare and contrast three model types: a purely mathematical model, an artificial intelligence (AI) model, and a hybrid of the two.

Mathematical modeling

If the underlying processes are understood entirely in terms of their causality and numerical solutions can be found at a reasonable computational expense, then mathematical modeling is always the method of choice. The dependence of the quantities of interest on the available input data transmitted to the digital twin can be represented by a mathematical function or a corresponding computer algorithm. In the case of the second example of the welding robot, this would be a function that accepts the measured input quantities such as temperature, humidity, and so on and produces a binary output indicating whether the electrode should be replaced or not.

A mathematical model is generally desirable as the results are reliable, usually computationally inexpensive and – most importantly – the underlying rationale is understood and explainable. Vital insight into processes is readily available and further knowledge can be gathered through the digital twin. Knowledge can later be fed back into research and development processes and may set a basis for future management and strategy decisions. In addition, mathematical models usually only require minimal amounts of training data, if any. However, these advantages come at a price; the development of mathematical models is often significantly more time-consuming and, depending on the understanding of the underlying processes, may not be feasible. Furthermore, a relatively large effort must be taken to evaluate whether a mathematical model can be found.

AI modeling

AI models are relatively general computational models and similar models can often be applied successfully to a very broad range of problems. More importantly, the underlying mathematics or physics of the problem do not need to be understood beforehand as the model dynamically adapts to the modeled data and adjusts its internal parameters, both accordingly and automatically. By these means, an algorithm can be found that – similar to the mathematical approach – links input parameters to the unknown output parameters of interest. However, in stark contrast to the mathematical approach, this algorithm features a significant number of parameters which are determined through training processes during which the parameters are adjusted to match large amounts of training data.[2]

[2] We emphasize that AI models are technically a group of mathematical models. The term AI is often used imprecisely in the literature and media, and here, we attempt to use it along the lines of how it is most commonly interpreted: Mathematical models are based on previously established rules, while AI models analyze large amounts of data to extract rules to a specific problem.

In many, if not most, computational models being built for commercial applications today, AI plays a central role. While the reasons for this widespread use are manifold, there are several key advantages that are certainly of great importance to many users. AI models tend to be much cheaper to develop and operate than mathematical models as, within certain boundaries, they are more generally applicable without in-depth domain knowledge, are versatile in their application, and dynamically adapt to changing circumstances. Furthermore, AI models tend to increase in accuracy with large amounts of data, which is presently cheap to store and gather, and often readily available. However, most AI models tend to be "black boxes," which means that the rationale on which decisions are made often remain in the dark, and "learned" results cannot be utilized beyond the predication made by the AI.[3] This lack of transparency in how decisions are made obstructs the derivation of research and development feedback as well as input into management and strategic decisions. Depending on the application, and only in rare cases, this lack of transparency in decision making may lead to ethical issues that forbid the use of some types of AI models.

[3] It should be noted that not all machine learning and AI algorithms are black box approaches. This may be thought of in terms of a classical decision tree, for instance, which is directly readable and interpretable. Most algorithms that are used today, however, are significantly more difficult to interpret as the model complexity is high, as is the number of internal model parameters that are adjusted in the model training. Algorithms, in which the rationale of how model results are achieved is difficult or currently impossible to extract, and include many deep learning, boosting, and ensemble approaches.

Mathematics and AI hybrid modeling

The above two model types can be combined to create a mathematics and AI hybrid. First, a mathematical model is developed to capture the essence of the underlying processes without the aspiration to describe the system exactly. The aim is rather to exploit the available domain knowledge to build a mathematical model that is as accurate as possible within the limitations imposed on the development in terms of time and financial constraints. In the second step, this mathematical model is extended to include AI. There are different ways to do so, one of which is to parameterize the mathematical AI model and have the AI predict the deviation between the result of the mathematical model and the exact result [4]. Other options include using the mathematical model in advanced feature engineering. Independent of the pursued approach, a hybrid model is developed that fulfills exactly the same tasks as the two above approaches, i.e., a mathematical and AI model; the determination or estimation of the quantities of interest on the basis of the input data that is transmitted from the physical entity to the digital twin.

Hybrid models combine most of the advantages that the purely mathematical models and AI models offer. Compared to a pure AI approach, they excel in transparency as the mathematical model and its parameters are accessible and can be rationalized, often in computational performance, and in the amount of data required to train the model. Maintenance efforts are comparable to a pure AI approach, if assumptions made in the mathematical model remain valid. Initial costs, however, are slightly elevated as time and labor must be allocated to the development of the mathematical model.

An illustration of the three different approaches can be found in Figure 3.

This section elaborates on the concept of digital twins from a methodical point of view. We have looked into various approaches to its implementation and different types of models have been compared and contrasted. The following extends this discussion to include the implementation of the digital twin concept in a typical business environment.

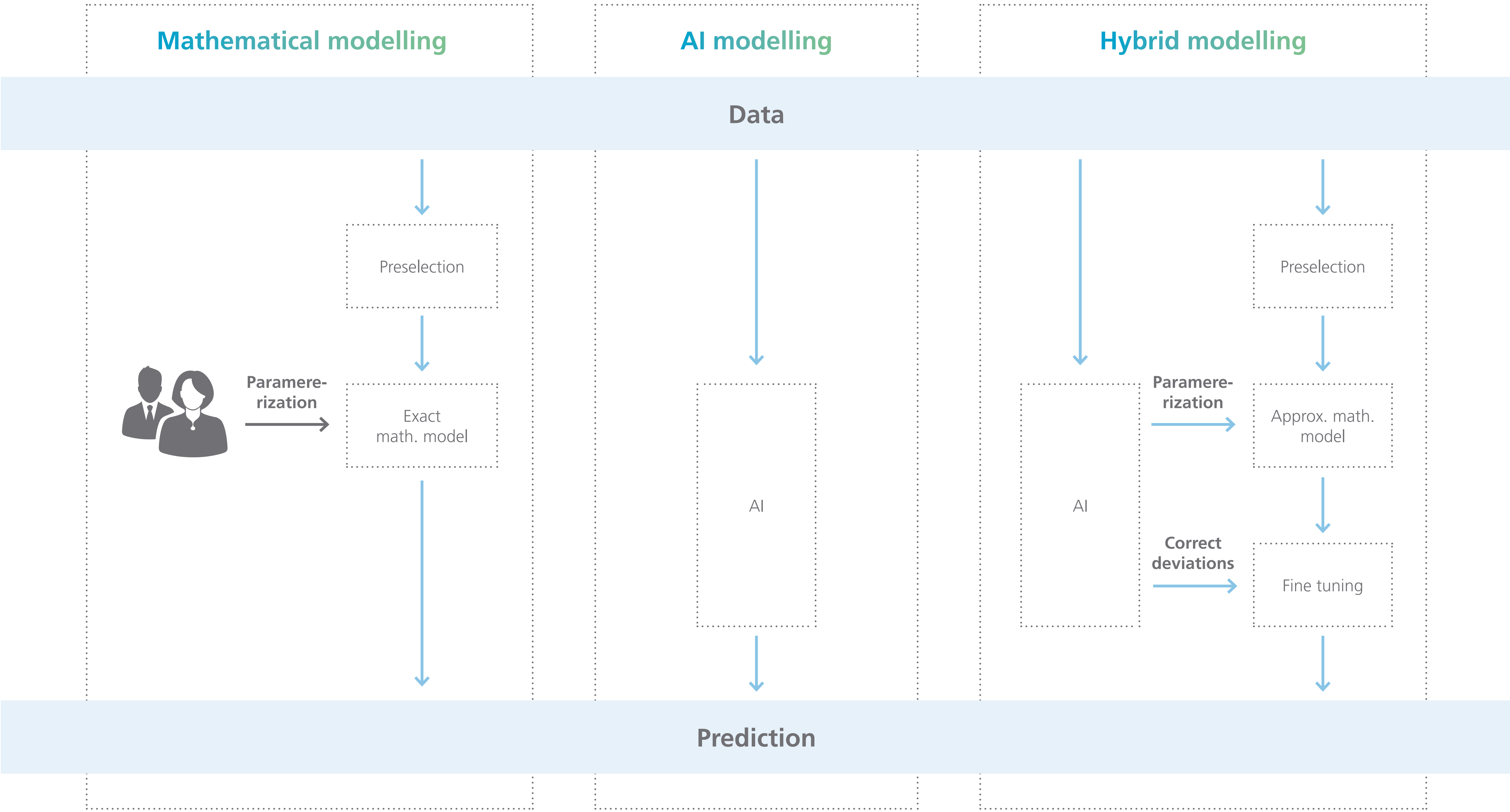


Figure 3: Comparison of a mathematical, an AI, and a hybrid modeling approach. Please note that the hybrid approach depicted is chosen as an example and there are numerous other options to build a hybrid model. Typical other approaches include using the mathematical model for advanced feature engineering.

Building successful digital twins

We have seen that the idea of digital twins is built on knowledge and data, which are both crucial to the success of any implementation of the concept. The deeper the understanding of underlying processes and the more data is available, the more accurate, useful, and eventually financially beneficial the implementation will be. While a number of factors influence the success of use cases, we firstly draw your attention toward the data that can be leveraged before we very briefly introduce a unique end-to-end implementation approach which has been developed by MHP.

An individual digital twin, i.e., a digital twin that exchanges information with a single physical entity to mirror its properties and state, can leverage a significant amount of information. The development of the model can be based on domain knowledge, empirical data of prior similar use cases, and other data made available by scientific communities or obtained from proof-of-concept projects. In addition, and as a fundamental component of any implementation of a digital twin, the model will make use of sensor data that is collected and transmitted by the physical entity to be reflected. For this reason, at MHP, AI-driven digital twins are gaining increasing significance in areas such as IIOT transformation, where they are classified as one specific use case category.

In ensembles of digital twins, opportunities to leverage knowledge and data are even greater and especially include experiences made with prior instances of digital twins. If a statistically relevant number of physical entities alongside respective digital twins are already in use and further instances are to be deployed, previously gathered knowledge can be used to precondition the digital twins of the instances that are to be added. For instance, this can be achieved by finding a statistical mean of the previously deployed models, via a clustering of the prior instances and the determination of the cluster affiliation of the future instances, or more sophisticated approaches.

To tackle the challenging implementation processes, MHP developed the unique and well-established AIDev methodology which leverages the experience MHP has gathered in numerous AI and data science projects over many years. It is characterized by its lean and agile nature and was designed to minimize costs in the identification of use cases and their evaluation while the process of algorithm deployment into a productive environment has been streamlined for maximum efficiency.

"The vast possibilities made feasible by digital twins will only become reality when we recognize and fully comprehended these potentials, so that conditions for regulated and secure data can be put in place."

Dr. William Cobbah

Opportunities & the way forward

Bearing in mind the challenges laid out in the introduction and the broad opportunities explored thereafter, there is no need to stress the potential of digital twins any further, although it is vital to note that these opportunities scale with the complexity of the tasks that need to be solved.

Let us think beyond the rather practical examples discussed above and consider a slightly more forward-looking example of complex interactions of a smart city. By leveraging the potential of ensembles of digital twins in a representation of all physical entities acting in an electrical grid, it will be possible to simulate and predict energy consumption to a stunning degree of accuracy, as large numbers of digital twins can learn from each other and increase their performance. Analogously, traffic and environmental challenges can be solved by using the same ensembles of digital twins and extensions thereof that may, at some point, evolve into a digital twin of aspects of a smart city. Similar opportunities can be identified in industry. For instance, it is foreseeable in manufacturing that integrated supply chains will use the idea of digital twin ensembles to produce models of suppliers and their interactions. At the required level of abstraction, the simulation of supply bottlenecks caused by individual specific entities will incorporate, for example, supplier risk data based on geo location, with representations of entities within the supply network.

All of the above is only feasible providing means are found to generalize, modularize and standardize digital twins and their ensembles. This enables large-scale virtual representations, which allow the simulation and analysis of future scenarios so that the physical world can be spared the adverse effects of growing complexity. This vision clearly goes far beyond the concept of ensembles of similar entities but is within the realms of possibility. The potential for abstract virtual representations for modeling purposes are truly stunning.

Visions of the above generous opportunities being used will only become reality if the benefits of digital twins become more widely recognized and understood, prompting frameworks for the regulated and safe availability of data.

References

[1] J. Janezko "Quality Management of Complex Systems," MHP AIdeaBook 2019, pages 19–22.

[2] K. Schubert, M. Schütten, M. Baumann, L. Hinrichs, A. Maheswari, „Der digitale Zwilling als Schlüssel zur ganzheitlichen Optimierung des Lademanagements in der Elektromobilität" in H. Proff (editor) "Neue Dimensionen der Mobilität – Technische und betriebswirtschaftliche Aspekte," Springer, 2020, pages 195–210.

[3] M. Grieves and J. Vickers, "Digital Twin: Mitigating unpredictable, undesirable emergent behavior in Complex Systems" in F.-J. Kahlen, S. Flumerfelt, and A. Alves (editors) "Transdisciplinary perspectives on system complexity – new findings and approaches," Springer, 2017, pages 85–114.

[4] T. Giese, S. Franz, W. Cobbah, „Automation of Production Planning Enhanced by AI", MHP AIdeaBook 2019, pages 24–29.

Publisher

MHP Management- und IT-Beratung GmbH

MHP is a leading international management and IT consultancy. We develop pioneering mobility and manufacturing solutions for multinational corporations, mid-sized companies and disruptive startups. As a premium business and technology partner, we are shaping tomorrow's digital future, today.

Our consulting approach is unique, combining holistic IT and tech knowledge with deep expertise in management. This makes MHP the ideal partner for a successful digital transformation. As digitalization experts, we deliver innovative strategies on the basis of strong analysis. These turn your change processes into sustained success.

Over 3,000 employees are driving digital progress in 20 locations worldwide, for over 300 clients. We display excellence at every level. **MHP: DRIVEN BY EXCELLENCE**

Contact persons and Authors

Dr. Enno Kätelhön
Senior Consultant
Intelligent Software
enno.kaetelhoen@mhp.com

Dr. William Cobbah
Associated Partner
Intelligent Software
william.cobbah@mhp.com

Contact persons international

USA:
Tobias Hoffmeister
CEO MHP Americas
tobias.hoffmeister@mhp.com

Greg Reynolds
Sales Director
greg.reynolds@mhp.com

UK:
Guy Williamson
CEO MHP UK
guy.williamson@mhp.com

CHINA:
Markus Müssig
CEO MHP China
markus.muessig@mhp.com

Photo credits
Cover ©shutterstock chesky // Page 2/3 ©istock metamorworks // Page 18 ©istock gorodenkoff

Layout
Freiland Design

MHP: DRIVEN BY EXCELLENCE

20 **MHP**Offices in Germany, England, USA, China,
Romania, Czech Republic, Austria, Israel, and Hungary.

International

Atlanta (USA)
Reading (England)
Cluj-Napoca (Romania)
Timișoara (Romania)
Prague (Czech Republic)
Shanghai (China)
Zell am See (Austria)
Tel Aviv (Israel)
Budapest (Hungary)

Germany

Ludwigsburg
(Headquarters)
Berlin
Essen
Frankfurt a. M.
Ingolstadt
Munich
Nuremberg
Wolfsburg

www.mhp.com