AI & machine learning will transform CAE simulation and deliver usable digital twins

By Kais Bouchiba (Hexagon | MSC Software) and Kambiz Kayvantash (CADLM)

A major change in the way manufactured products are designed and made is unfolding before us in the 21st century especially as we move deeper into Industry 4.0 and as the emergence of autonomous manufacturing looms on the horizon. The convergence of mechatronic, or cyber-physical technologies, with advances in data management, artificial intelligence (AI), machine-learning (ML), and communications via the Internet of Things (IoT) is already challenging traditional industrial product manufacturing processes. Add to this the impact of the COVID-19 pandemic this year and we see the consequence of an acceleration of digital transformation happening across all industry sectors. Manufacturers today need to begin to implement rigorous systems-design processes that accommodate the complexities of developing multi-disciplinary systems, with high-fidelity virtual prototypes, or 'digital twins', at the core of their development process. This will not be achieved without challenges for sure, but we believe that the tools exist today to overcome these obstacles and connect 'digital threads' with feed-forwards and feed-backwards loops between real time design & computer-aided engineering (CAE) simulations and measured manufactured and lifetime product data that will yield cost savings, higher quality products and high levels of productivity and innovation, yet retaining simulation accuracy.

Introduction

In a recent report by PricewaterhouseCoopers, they observed that *"AI could contribute up to $15.7 trillion to the global economy by 2030, more than the current output of China and India combined"* (ref. [1]). The same report estimated that in 2018 alone, AI contributed $2 trillion to global GDP. However, further research from IDC has found 'that half of AI projects fail for one in four companies on average,' and 'the two leading reasons for an AI project failing are a lack of required skills and unrealistic expectations' (ref. [2]). In addition, an MIT Sloan Management Review from Boston Consulting Group's (BCG) 'Artificial Intelligence Global Executive Study and Research Report' validates the sobering statistics from IDC (ref. [3]). Seven out of ten companies surveyed in the BCG report showed minimal or no impact from AI so far! And of the 90% of companies that have made some investment in AI, fewer than 40% report business gains in the past three years. However, increasing revenues and diminishing costs are prizes awarded to companies capable of succeeding with AI. Many executive teams aiming to balance the demands of multiple priorities can lose focus on this fact and miss their track for digital transformation. It therefore goes without saying that many companies are trying to figure out how to avoid this fate.

Today, integrating CAE with computer aided design (CAD) simulation software has become a fundamental ingredient in the practice of engineering simulation. Finite element methods (FEM) and finite volume methods (FVM) have emerged as the principal tools for physics simulation in many fields of product design and manufacturing. Following CAE's success in addressing many design challenges, the complexity of the problems FE/FV simulation have been facing over time has grown rapidly resulting in increased model sizes and computing effort allowing for a higher fidelity in representing real-life engineering problems. Because of increased simulation efforts, the computing and human resources necessary for FE and FV simulation have grown dramatically in the last 20 years, now appearing as a significant cost component in the design process. Hence, the cost of computational resources (computer hardware, software, engineer time and computing time) has become a major obstacle for improving the design process further. Fig. 1 shows a timeline for CAE, AI and the so-called 'Digital Twin' associated with Industry 4.0. It can be seen that the last 20 years has seen an acceleration of AI and ML advances. We view a Digital Twin as a virtual model of a process, product or service and this pairing of the virtual and physical worlds allows analysis of data and monitoring of systems to head off problems before they even occur, preventing downtime, developing new opportunities and even planning for the future by using simulations.
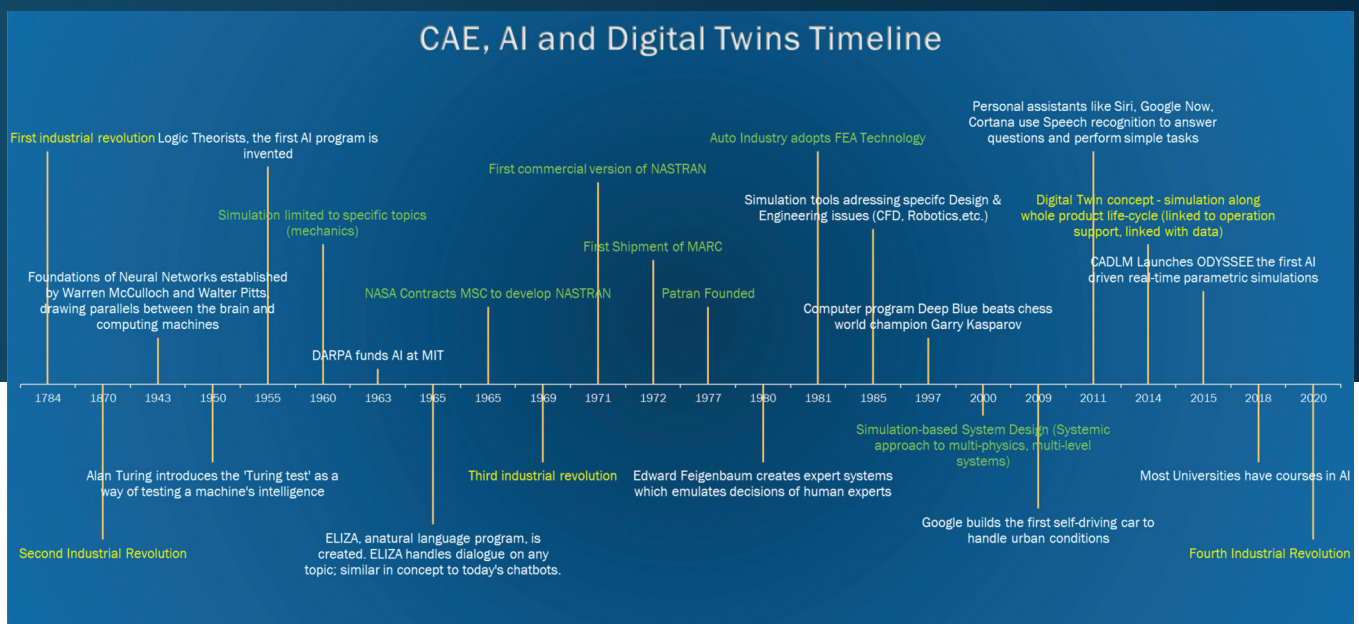

Figure 1: A timeline for the evolution of AI, Digital Twins and CAE Simulation

In parallel to the growth of CAE, AI/ML has been advancing very quickly in the last two decades and inventing itself new methods that address the complexity of the same design problems as those tackled by CAE and they have been applied to areas like finance and marketing mobile applications. In the last couple of years in particular, ML methods based on deep artificial neural networks (so called 'deep learning') have achieved tremendous success in many engineering applications as well. Additionally, advances in image and signal recognition as well as robotics using deep learning and the implementation of these methods using specially designed platforms running on GPU-based clusters are allowing ML models to significantly reduce the CAE simulation process by summarizing the results of engineering simulations. Hence, ML models are also now capable of capturing the know-how gained from multiple CAE simulation runs in DOE (Design of Experiment) loops thus enabling the democratization of complex engineering simulation tools and opening routes to new business models. Because of the above arguments, we are moving today from the traditional CAE paradigm that's existed for over 40 years to a new one showing tremendous productivity gains for manufacturers as described in Fig. 2.

We believe that ML will not only enable a new wave of process automation for CAE but it will also speed up the development of CAE simulation tools that allow non-experts to use sophisticated simulation capabilities - what amounts to 'CAE democratization' - with associated new benefits. Manufacturers should be employing AI and ML as a serious CAE option today in order to significantly impact product development processes and product life cycle development.
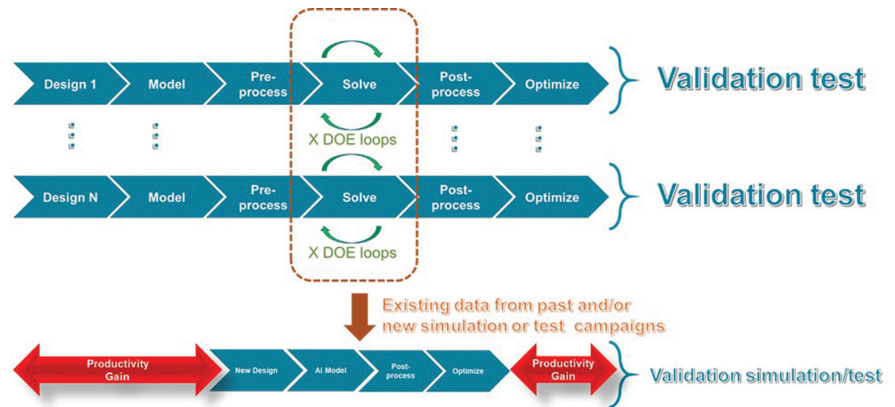

Figure 2: From simulation validated by physical test to Design of Experiment (DOE) fed AI models validated by CAE simulation or physical test

CAE will be challenged by Industry 4.0 & Digital Twins

Manufacturers can no longer afford the 'build it and tweak it' approach that has long characterized many design projects as we move towards zero prototyping. As companies face mounting pressure from global competitors and a need for constant product innovation, engineering criteria have become essential to competitively differentiate products in many markets. Based on a recent survey of 195 companies published by Tech-Clarity (ref. [3]), 80% of respondents believed that product quality is the most important product attribute to keep products competitive. Reliability and cost came next. This indicates that their customers have high expectations for quality and durability but do not want to overpay. Hence, to be successful, companies have to balance these criteria. Requirements for quality, reliability, and cost often conflict with each other and balancing them remains a challenge. Unfortunately, ever increasing product complexity these days makes it hard for engineers to know the full impact of each design decision immediately. Indeed, 76% of survey respondents rated design decisions that affect product competitiveness as 'somewhat hard' to 'extremely difficult.' This leads many engineers to 'over engineer' their products, which unfortunately drives up cost and weight. Add to this the ever-shrinking timelines in modern manufacturing processes, it therefore means that increasing trade-off decisions can lead to many well-known unfortunate consequences we see on the daily news with famous brand product recalls and adverse public and media perceptions of those brands.

In today's competitive global market, relying on CAE simulation as a design tool to optimize designs and to provide guidance for product development means that engineers have to look for a CAE simulation solution that can offer instant ("real-time" or near real time) results but also be accurate and reliable! Therefore, the traditional CAE industry needs to evolve with growing customer expectations against the reality that:

- In some instances, it takes too long to get a CAE result
- Not all CAE data is available
- 95% of CAE data is deemed invaluable
- CAE can be perceived as being too expensive to utilise
- Engineering judgement can be difficult with increasing number of disciplines involved, and
- Limited numbers of predictive adaptive CAE models are available for integrating future complexity.

All the above factors lead to abandoned (or unexplored) new product designs or variants thus limiting innovation and affecting final product quality. They also highlight the discrepancies between the digital and real worlds and explain why fusing the two worlds is a strategic and growing challenge for most enterprises. Additionally, progress in digitalization of the global economy and industries with higher expectations and standards is making this challenge even greater. The state-of-the-art in product manufacturing today is often referred to as Industry 4.0. In effect, it comprises intelligent machines, equipment and products that independently exchange information, initiate actions, and individually control or influence each other. The ultimate aim however should be to fundamentally improve industrial processes along the entire product lifecycle and manage the increasing complexity of manufactured products yet handling the development of data driven systems for knowledge capture and industrial good judgement. Industry 4.0 and digitalization therefore provides countless subject areas that are continually evolving. Hence, new technologies such as Big Data Analytics, Cyber-Physical Systems, Cloud Computing and the Internet of Things (IoT), are being developed rapidly for Industry 4.0.

Ideally, digital engineering design tools (such as CAE) should integrate into the real-world of controlling a production facility or the product itself through an end-to-end 'digital thread' of data. The challenge therefore becomes building PLM systems and approaches that will help not only during the conceptualization, prototyping, testing and design optimization phases, but also during the operation phase with the ultimate aim to use them throughout the whole product life cycle and beyond to retirement or recycling in an increasingly circular economy. While in the first phase of R&D, the importance of numerical CAE simulation tools and tests/experiments is undeniable today. In the operational phase, however, the potential for real-time availability of accurate CAE data will inevitably open up new avenues for monitoring and improving operations throughout the life cycle of a product. This has huge cost saving and quality implications for any product. Fig. 3 is a schematic illustration of this Industry 4.0 challenge for AI and ML.

It is not sufficient to talk about product development processes and product life cycle improvements without bringing into the picture, along with AI, the concept of digital twins. They are one of the leading concepts closely linked to the IoT and Industry 4.0. A digital twin can be seen to be a digital representation that simulates virtually a real-life object, process or system. Digital twins consist of 3 main components:

1. The physical object in the real world
2. A virtual object in the digital world
3. A connection between the real and virtual objects via data and information.

Digital twins should include the correct laws of physics, correct material properties, virtualized sensors & causality. Engineers can build digital twins of complex physical assets using design, manufacturing, inspection, sensor and operational data. Moreover, a digital twin does not stop when we proceed to production; we can use it throughout production and into the aftermarket. A digital twin can therefore be used not only for the maintenance of a product, but also for predictive CAE analysis since it can also contain measurement data from internal sensors in feedback loops. Ultimately, we cannot really understand data without context and intent, and the accuracy of the digital twin increases over time as more data refines the AI model. Finally, Machine Learning and the digital twin should interact and improve one another. Since the value of digital twins became clear over the last decade, they are gaining more and more interest and importance in many companies and industries. The digital twin has been placed in the top 10

strategic trends for the year 2019 by Gartner, and they estimate that by 2021, half of all significant industrial groups would use digital twins, increasing their effectiveness up to 10% ([ref 4]). As such, digital twin technology is becoming an integral part of the simulation, testing and operation of different manufactured products. Since an effective digital twin must account for change and have representations that make it possible to take long-term historical data and experiences into account, physics-based CAE simulation has a fundamental role to play in plugging data gaps throughout a product lifecycle. Machine Learning therefore helps correlate and automate these data sources, but the digital twin is only possible using physics-based simulation data, machine learning, and physical measurement together. We also believe that from a completeness and cost perspective, physics-based CAE simulation data remains essential.

High-fidelity virtual prototypes, or digital twins, should become the core of a product development process and there is a need to connect the dots between CAE, AI and digital twins. Besides applications of digital twins in Smart Cities, Transportation, Meteorology, Healthcare, and Education, some of the more advanced deployment of digital twins today are currently found in
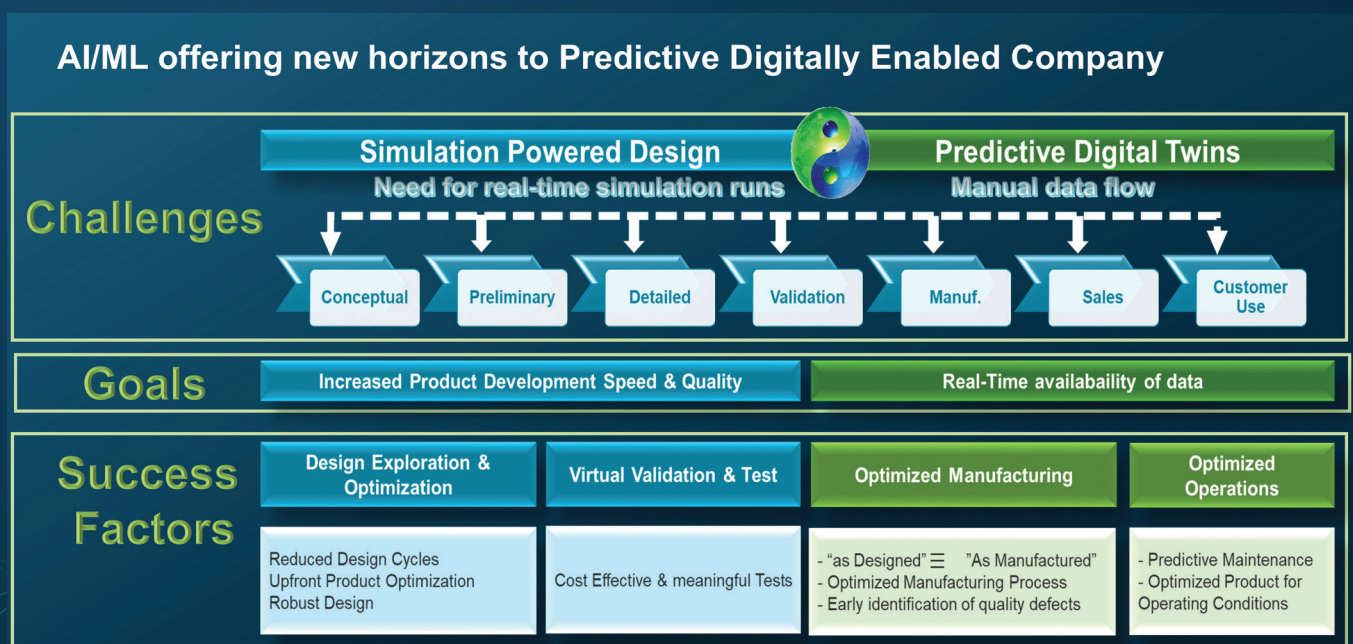

Figure 3: AI/ML offering new horizons to a Predictive Digitally Enabled Company

the manufacturing sector, with many factories already using twins to simulate production processes. The first benefit of a digital twin is the ability to produce simulated data. If we look at the Automotive industry for instance, the physical system of designing cars usually covers millions of testing miles, whereas the digital twin of the car needs to cover billions of virtual miles to robustly enhance its radar and image recognition, and vehicle-to-vehicle communication capabilities; in fact, a virtual testing environment can potentially go through an infinite number of repetitions and scenarios. The simulated data produced can then be used to train the AI model. This way the AI system can be taught potential real-world conditions that might otherwise be rare or still in the testing phase. Hexagon | MSC Software's VTD (Virtual test Drive) software does this today. The second benefit of a digital twin is the ability to plan and test new features. A digital twin therefore should represent reality today, but it can also produce a view into the future. Designers can then virtually create tomorrow's cases for their product and test scenarios. These tests can be tweaked and performed as many times as they like thus finding the most optimal solution that they can then take and make.

Although in their current state digital twins can improve design & engineering simulation because they are a great and valuable source of data to feed AI models, challenges however remain. Fig. 4 summarizes the interaction between value generation challenges and enabling technology. We can identify the following as being specific to CAE:

- For Design Engineers, reliable data would only be available *a posteriori...* In fact, we can create a virtual representation of the physical world by bringing in real-time data from some system and monitor it so that we can anticipate a problem before it occurs. What differentiates this from simulation is that, because it is based on the flow of real-time data, the answer it gives you today will likely be different to the answer it will give you in a week from now.
- Sharing data comes at a high cost and with great tension, as digital knowledge, practices and culture are not yet converging across the built environment. For instance Indeed, digital twins require real-time solver runs. This is where AI and ML techniques applied to CAE brings all the needed and missing value.

From a purely technical perspective (beside data security, data quality improvements and latency), real-time CAE simulations, large scale data fusion and assimilation, intelligent data analytics, predictive capacity, transparency and generalization of technologies across diverse application areas are considered the main challenges in developing digital twin technologies today.

We believe that the need for two digital twin enabling technologies addressing the above challenges becoming a "must" today:

I. **Physics-based Modeling:** This consists of observing a physical phenomenon of interest, developing a partial understanding of it, formulating that understanding in the form of mathematical equations and ultimately solving them. High-fidelity CAE solutions add physical realism to any digital twin while various discretization techniques over time have been developed for this. These have been extensively used in many open-source and commercial multi-physics simulation packages (like MSC Nastran, Marc, Adams, Actran, Romax, Cradle CFD, Simufact etc). Despite the immense success of using high-fidelity CAE techniques they have so far been limited to the design phase by and large. Unless their computational efficiency is improved by several orders of magnitude, their full potential will remain under-utilized in a digital twin context. However, great advances in high performance CAE solvers during the last two decades thanks to Moore's Law qualify (many of) them to be denoted 'high-fidelity' models that can serve to develop so-called "Reduced Order Models" (ROM) which may be used efficiently to establish predictive digital twins.

II. **Data-driven Modeling:** While physics-based models are the workhorse of CAEsimulation at the design phase, with an increasing supply of data in a digital twin context, open-source cutting edge and easy-to-use libraries, cheap computational infrastructure (CPU, GPU and TPU) and high quality,
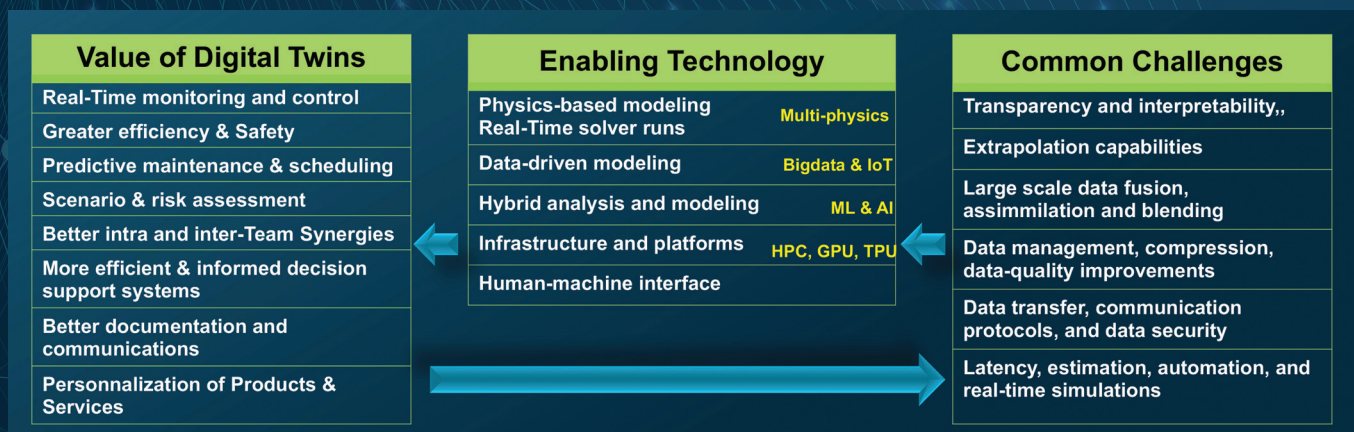
Value of Digital Twins	Enabling Technology		Common Challenges
Real-Time monitoring and control	Physics-based modeling Real-Time solver runs	Multi-physics	Transparency and interpretability,,
Greater efficiency & Safety	Data-driven modeling	Bigdata & IoT	Extrapolation capabilities
Predictive maintenance & scheduling	Hybrid analysis and modeling	ML & AI	Large scale data fusion, assimmilation and blending
Scenario & risk assessment	Infrastructure and platforms	HPC, GPU, TPU	Data management, compression, data-quality improvements
Better intra and inter-Team Synergies	Human-machine interface		Data transfer, communication protocols, and data security
More efficient & informed decision support systems			Latency, estimation, automation, and real-time simulations
Better documentation and communications			
Personnalization of Products & Services			

Figure 4: Interaction between Value Generation Challenges and Enabling Technology in Digital Twins

readily-available training resources, data-driven modeling is becoming very popular. Compared to the physics-based modeling approach, this approach is based on the assumption that since data is from both known and unknown parts of the physics in questions, by developing a data-driven model, one can account for the full physics simultaneously. Smart data analysis using ML and AI is therefore expected to play a major role in the context of digital twins. Adding machine learning to any industrial process will make the process more intelligent by getting more accurate data and predictions, accompanied by additional numerical and visual understanding of otherwise unstructured data. Another advantage of data-driven models is that they continue improving while more and more data (experiences) become available. The training part of the data-driven modeling might experience issues associated with instabilities though. However, once trained, the models will be stable and sufficient for making predictions. By adding machine learning into a CAE workflow, we don't only open up possibilities to discover previously unseen patterns in our data, but also create a single learning-system that can manage complex data.

We believe that a new approach can be developed to combine physics-based CAE modeling and data-driven modeling. The combined approach should be aimed at removing the shortfalls of pure physics-based or pure data-driven modeling approaches (see Table 1 for a summary). It should combine the interpretability, robust foundation and understanding of a physics-based modeling approach with the accuracy, efficiency, and automatic pattern-identification capabilities of advanced data-driven ML and AI algorithms.

Table 1. Physics-based modeling vs data-driven modeling

Physics-Based Modeling	Data-Driven Modeling
+ Solid foundation based on physics and reasoning + Generalizes well to new problems with similar physics	+ Takes into account long term historical data and experiences + Once the model is trained, it is very stable and fast for making predictions
− Difficult consistent engineering judgment with increasing complexity Can be too long and be too expensive − Difficult to assimilate very long-term historical data into the computational models without a Simulation Data Management System like MSC SimManager − Sensitive to numerical instability when dealing with non-linearities and ill-conditioned problems	− So far most of the advanced algorithms work like black boxes − Bias in data is reflected in the model prediction − Poor generalization on unseen problems

Figure 5: The 3 main Machine Learning techniques

Supervised Learning
✓ Pre-labeled data
✓ Ijnput and output dataset
✓ Classification
✓ Regression

Unsupervised Learning
✓ Unlabeled data
✓ Ijnput dataset only
✓ Hidden features
✓ Clustering

Machine Learning

Agent
Environment

Reinforcement Learning
✓ Reaction to an environment
✓ Rewards system
✓ Decision process

HEXAGON

What is AI/ML and how does it work?

Any machine learning can be broadly categorized into basically 3 types of techniques - summarized in Fig. 5 – involving supervised learning, unsupervised learning and reinforcement learning. They can be defined as follows:

I.   **Supervised machine learning** builds a model that makes predictions based on evidence in the presence of uncertainty. A supervised learning algorithm takes a known set of input data and known responses to the data (outputs) and trains a model to generate reasonable predictions for a response to new data. Use supervised learning if you have known data for the output you are trying to predict. One of the shortfalls of supervised algorithms is the need for dependent variables (labeled data) which might not always be available as in the case of an anomaly. Unbalanced or skewed data rarely result in reliable prediction models. In such a situation, unsupervised algorithms have better utility.

II. **Unsupervised learning** finds hidden patterns or intrinsic structures in data. It is used to draw inferences from datasets consisting of input data without labeled responses. 'Clustering' is the most common unsupervised learning technique. It is used for exploratory data analysis to find hidden patterns or groupings in data. Applications for cluster analysis include gene sequence analysis, market research, and object recognition. Beside anomalies, another important application of unsupervised algorithms like PCA and Deep Auto encoder can be for on-the fly data compression for real-time processing, communication and control of the system under operation.
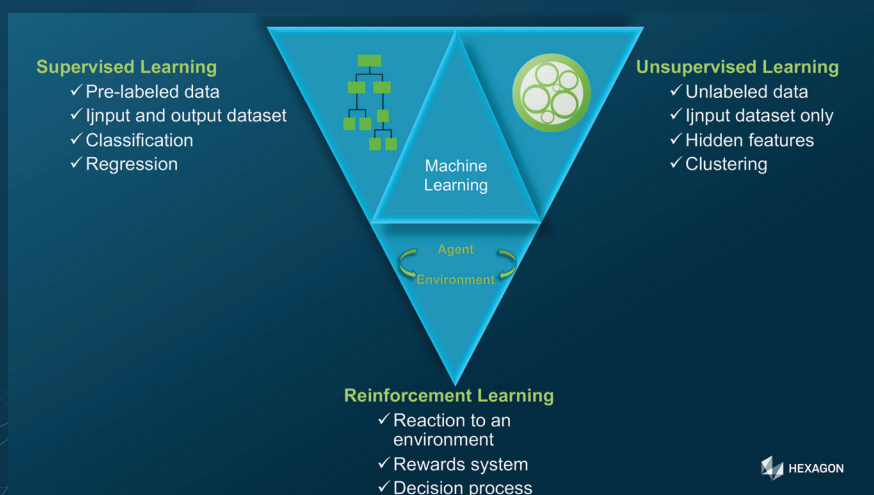
III. **Reinforcement Learning** though has the potential to help in such a data-deprived situation and is a type of learning that is based on interaction with the environment. It is rapidly growing, along with producing a huge variety of learning algorithms that can be used for various applications. To begin with, there is always a start and an end state for an agent (the AI-driven system). However, there might be different paths for reaching the end state, like a maze. This is the scenario wherein reinforcement learning is able to find a solution for a problem. Typical examples of reinforcement learning include self-navigating vacuum cleaners, driverless cars, etc.

While supervised and unsupervised learning ML algorithms have been the most-commonly employed algorithms in real applications, they are not of much use in the absence of enough data. The following table 2 gives an overview of the difference between supervised, unsupervised and reinforcement learning:

Criteria	Supervised Learning	Unsupervised Learning	Reinforcement Learning
Definition	The Machine learns by using labeled data	The Machine Is trained on unlabeled data without any guidance	An agent interacts with its environment by performing actions and learning from errors or rewards
Type of Problems	Regression and classification	Association and clustering	Reward-based
Type of Data	Labeled data	Unlabeled data	No predefined data
Training	External supervision	No supervision	No supervision
Approach	Maps the labeled inputs to the known inputs	Understands patterns and discovers the output	Follows the trial-and-error method

How does AI work with physics-based CAE simulation and how does their interaction improve the product design process?

Machine Learning techniques offer a smarter approach to product design as long as we understand that both CAE and AI can mutually deliver incremental value to each other, making the combination of both an efficient productivity improvement engine where ML offers the reduction of number of simulation runs during the design of a new, but 'almost', similar product. The mix of an AI and physics-based approach better addresses the increasingly complex problems confronting engineers today. Nevertheless, it is ironic that one of the biggest challenges when using machine learning to improve a manufacturing process is that you cannot physically create enough data! This is especially the case for internal, not easily visible physics system data (energy, stresses, strains, etc.). However, using manufacturing process simulation to generate data we can take a complex process like metal Additive manufacturing for instance, and build a large enough dataset to create predictive machine learning models. We can take the same approach in Aerospace composites, where using virtual testing of materials is the only way to augment costly coupon tests so that customers can apply machine learning. By using multi-scale materials modelling with machine learning, they can quickly understand the performance of each configuration (resin, fiber, fiber orientation, etc.) of a given material system as manufactured with each available process. In fact, data is the real 'fuel' for Machine Learning. However, on the one hand, we tend to reduce physical testing that take too long to obtain, is excessively expensive and delivers incomplete data, and on the other hand, data from the digital twins comes too late for predictive maintenance and are by essence, *a posteriori* data. Therefore, only with engineering simulation can sufficient and meaningful data be realistically, and cost effectively, generated to make AI successful in early stages of engineering (fig. 6). Then, as real-world
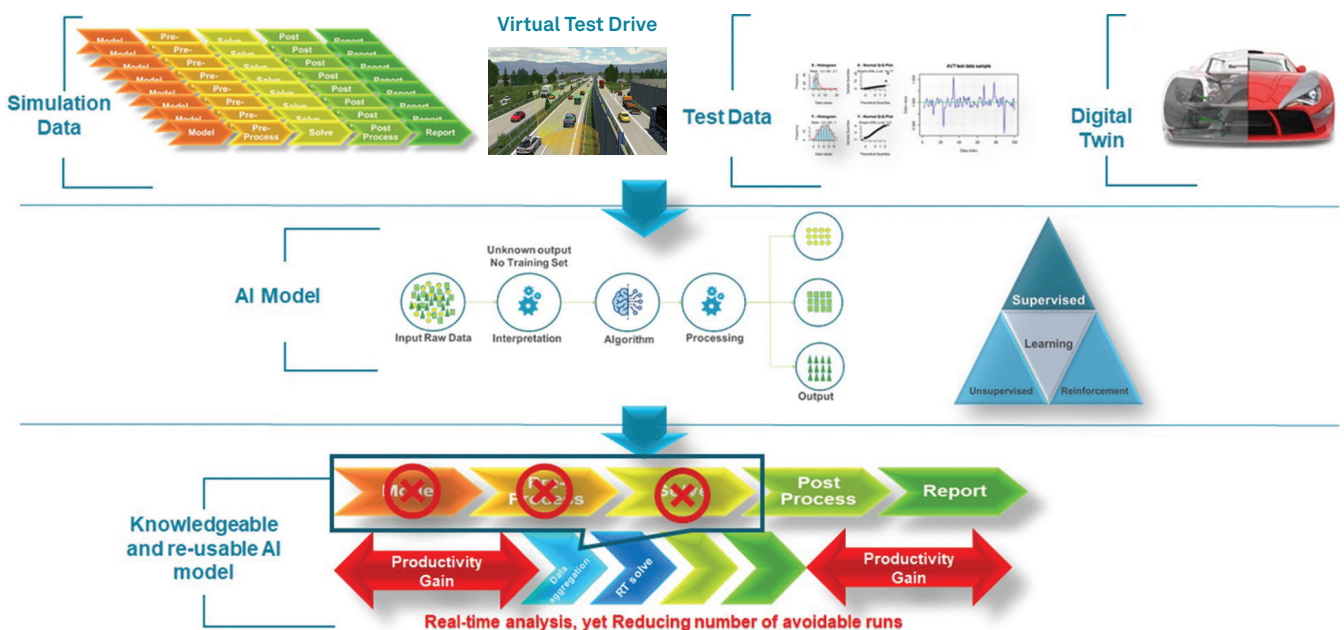

Figure 6: Simulation is the best and most realistic yet cost effective source of data – virtual test driving of car environmental simulation scenarios

data arrives over time, our models will become ever more accurate. Since we now have the data and computing capability, our strategic models and operational models can merge, with a strategic model simply being a long-running operational model.

In recent years we have seen many AI developments and several of these have been small incremental improvements one upon another. These developments can be grouped into a few major types.

I.   There has been a substantial increase in the amount of training data underlying the AI models that are interesting to many industrial applications.
II. Computing hardware is significantly more powerful today making it more realistic to train models for longer.
III. Mathematicians have discovered ways to accurately train neural networks with more than one hidden layer (deep learning).
IV. Derivatives of the deep learning approach have constructed some models that are good at specific tasks. Examples include the convolutional neural networks that are good at image recognition and 'random forests' that are good at categorizing numerical features.

Viewed like this, all four major developments have direct useful implications for CAE. The second, third and fourth developments provide the promise that CAE can be assisted by these methods. The major assistance AI/ML offers is in automation of CAE activities respectively in the reduction of the duration of these activities. The first development above is however a liability for CAE in the sense that the development requires larger datasets for learning, and these must be generated for CAE use cases before the advantages of AI can be reaped. This represents a cost, but it is unavoidable. In the product design phase, AI therefore offers opportunities to carry on simulating much larger high-fidelity models, but also increases the efficiency of the whole workflow at a reasonable cost. Machine Learning engines can leverage datasets from former simulations as well as test data sets which are

"dormant data", yet extremely valuable. Clearly, factory tooling must be CAE-aware so that the engineers can easily tune the process to their needs. In this scenario, machine learning serves as a repository of the know-how gained from running multiple simulation runs. This repository enables the democratization of complex engineering tools and opens new possibilities with respect to sharing data between companies and throughout their supply chains.

Today, Machine Learning for CAE has numerous methods at hand:

- Traditional Interpolators (RBF, Kriging, splines, …)
- Algebraic Decomposition techniques, (SVD, EV)
- FFT, Wavelets, LSE
- PCA, Kernel PCA, RDA (Redundancy Analysis), CCA (Canonical Correspondence Analysis)
- Clustering (PCoA, K-means, Tessellation)
- Support Vector Machines (with CG optimizers)
- Mixture Models (Dynamic Model Decomposition, Kalman Filters, Markov Chains)
- Forecasting (ARIMA, etc.)
- Neural Networks (MLP, etc.)
- Convolutional Neural Networks (Deep Learning)
- Entropy and complexity analysis,
- Lossy and lossless compression techniques
- Reduced Order Modelling (POD, PGD, CVT, FFT, …)

The real promise of simulating multi-physics attributes in CAE is the ability to do multidisciplinary optimization. Starting from a problem that the engineer defines, machine learning can help streamline optimization of a high number of variables. The modeling of optimization problems and multi-physics phenomena in practical engineering applications is often particularly challenging, as repeated numerical simulations are required. A remedy to this is a simplification of the physics-based CAE model but that relies on the experience and intuition of the engineers. Another avenue is Reduced-Order Modeling (ROM), a mathematical approach serving to overcome high computational costs of the simulations via decomposition techniques employing already known past responses (ref. 5 and 6). This workflow can begin by using a co-simulation CAE model to create datasets that are used to train a ROM that provides sufficiently accurate results across the physical domain required by identifying the most pertinent data from previous CAE runs to optimize the simulation's dataset before it is run. We have seen great success applying CAE-aware machine learning to creating accurate real-time CAE simulations of multi-body dynamics and testing Automotive hardware-in-the-loop. It simply was not feasible before with a full simulation model.

Reduced Order Modelling (ROM) belongs to the category of 'Fusion' or 'Dimensionality Reduction' techniques. ROMs can be considered as a simplification of a high-fidelity dynamical model that preserves essential behavior and dominant effects for the purpose of reducing solution time or storage capacity required for the more complex models. There are a great number of high-dimensional problems in the field of science (like atmospheric flows) that can be efficiently modeled based on embedded low-dimensional structures or reduced order models (ROMs). This family of models can be best described as the intersection between now classical pure data-based ML and physics-based modelling (often based on available PDE's) for high fidelity CAE simulations and data driven models. Model reduction techniques can be regarded either as algebraic reductions of the PDE's (such as 'Proper Generalized Decomposition' or PGD) or compression techniques applied to the DOE based solutions of the same equations (called 'Proper Orthogonal Decomposition', POD). While both are based on decomposition-interpolations handling of the data, their implementation differs in the sense of intrusivily in conjunction with the solver formulation itself. Both provide a reduction of the volume of a data set while preserving the most important parts of
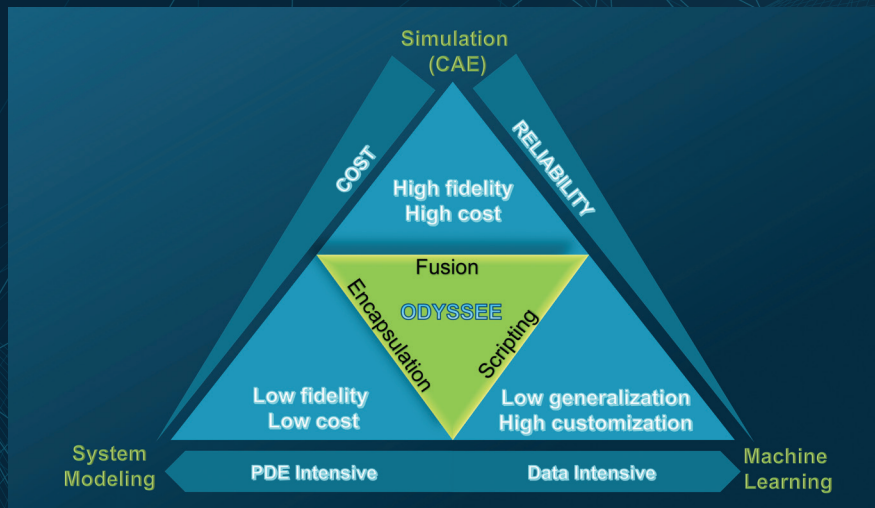
Figure 7: New and innovative AI based approach with CADLM's ODYSSEE platform

the information contained within the data (comparable to "modes" or "frequencies" of the response), necessary for retrieving all or the most essential part of the information when needed. In particular, the POD approach can be considered as a pure compression technique (and therefore a ML since the start point is data and not an equation) similar to those used in image compression and object recognition. The decomposition (or compression) can be done via matrix decomposition techniques or alternatively via clustering or any other signal processing algorithms (Fast Fourier Transforms or Clustering). We consider that the POD-like solutions are currently the most convenient and efficient implementations.

It can also be claimed that contrary to other ML techniques which are pure data based, POD-like ROMs benefit also from the fact that we are aware of the existence of an underlying physical reality, since the data are indeed issued by such existing reality, either numerically (FE) or experimentally. The uncertainty is not on whether a real model exists or not, but rather how good it can be reconstructed after fusion. Notice also that such techniques allow for creating on-board and real-time applications based on voluminous experimental or simulation results (for instance Finite Elements) with huge application potential. Whatever the affiliation, combining ROM methods and more traditional ML techniques overcome optimally the challenge to achieve accurate real-time simulation. Indeed, ROM combined with FE simulation allows for the modeling of the most complex structures, while ROM can also help optimize the use of simulation resources to make product designs more efficient without sacrificing much on accuracy. It allows a powerful tool when used by optimization algorithms since it removes the need for inaccurate and incomplete (and often costly) response surface methods, based on algebraic fitting of scalar fields.

A combination of Hexagon | MSC Software CAE solutions and CADLM's ODYSSEE platform allows for a new and innovative technology that can enhance and optimize current traditional approaches, without excessive interfacing and scripting effort. Since the primary goal of all above techniques (fig. 7) is to approximate the large-scale problem by a much smaller one, which yields somewhat less accurate results but can be solved with considerably less computational overhead, Reduced Order Models (ROMs) provide an opportunity to create a virtuous Real-Time loop between Design and Operations with real time information sharing. It also provides an opportunity for simulation software providers like Hexagon | MSC Software to truly democratize engineering simulation across the product life cycle in a scalable manner without compromising model fidelity.

This initiative to develop model reduction approaches for a variety of engineering problems while remaining agnostic to the underlying physics type allows the computer-aided engineering simulation community to tailor the level of model fidelity to the underlying simulation intent. For example, reduced-order (yet physical and not response surface based) surrogates of high-fidelity CAE models can be used to explore the design space and execute computationally intensive, vehicle reliability, and optimization tasks. Studies over a wide design space with many design, event, and manufacturing parameters will always require relatively fast-running models. The level of accuracy of ROM models used for wide design space exploration and optimization does not need to be at the level of a high-fidelity physics-based model, but rather just requires the capture of the essential behavior and relationships. While classical surrogate models are based on algebraic fitting

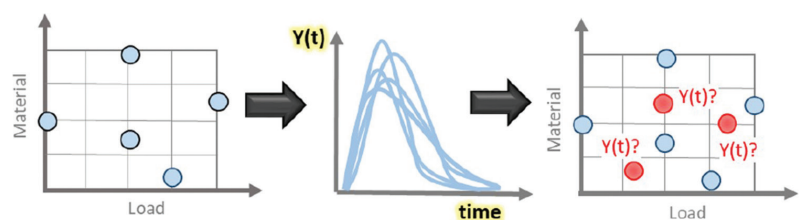Figure 8: ROM ODYSSEE Lunar analysis of an Adams Car Suspension Model

Figure 9: Data sampling through D.O.E. Design

work in some applications, they do not capture the essential behavior and relationships for many problems of current interest. The state-of-the-art of machine learning-based ROM models can do just that even for highly nonlinear and transient response across multiple physics types. ROM is applicable to multiple range of physics and has tremendous advantage over traditional "surface responses" through the ability to deliver time-dependent responses, in the same way a FEM transient analysis would (to the extent of the model reduction assumptions).

A great illustration of the ROM approach can be found in a recent proof-of-concept, combining Adams, the leading Multi-Body Dynamics simulation Software from Hexagon | MSC Software, and Lunar, the Supervised Machine Learning solution from CADLM's ODYSSEE suite, used to create Reduced Order Models (ROMs) of vehicle behavior using MSC's Adams software – see figure 8 (Ref [5] & 6]).

With such an impressive correlation between CAE and AI, it is easy to realize that combining ROMs with AI delivers the best of both words (CAE and AI), allowing Large-scale design space exploration, CAE optimization and uncertainty quantification as well as 3D-0D links. ROM combined with machine learning operates in 2 steps:

Step 1: Learning
- Decomposition of data base
- Compression (reduction) of data base
- Convergence indicators

Step 2: Testing and validation
- Reconstruction
- Testing: "leave-one-out" approach
- Prediction
- Quality indicators

Design of Experiments (DOE) are needed to feed the machine learning with data when considering well balanced (space filling) samples of "X" and computed variables "Y" (see fig. 9).

With this combined AI and CAE approach we then can achieve:

- Real-time computing – almost zero computing effort for parametric CAE studies and optimization
- Reduced computing effort – few but wisely selected sampling points and adaptive learning (improves as you learn)
- Precision and completeness – full time-history output (not only scalars!) and physical domain decomposition, not fitting (this is NOT a response surface method!)
- Production of 3D animations – no interpolations but reconstructions, and
- On-board applications (no *a-priori* knowledge).

The benefits of ROM become obvious in term of solution time (from hours to seconds!) as well as storage capacity, while preserving essential behavior and dominant physics effects. The beauty of ROMs goes beyond 3D simulations to speedup simulation time. Used for systems, ROMs provide a way to reuse modeling assets from 3-D analyses and can be integrated with other system level components for building virtual system prototypes; allowing engineers to run real-time scenarios like virtual car simulators. And ultimately thinking of them in embedded controls, ROM provide a way to introduce "virtual" sensors that can be used for controls and open the way to fusing the real world and the virtual world making real the concept of truly interconnected digital twins with in-use manufactured parts. As such, ROMs and system simulation can be used while an asset is operating and connected to an IoT platform for the purpose of enhanced monitoring, asset optimization, diagnostics and predictive maintenance. But ROM can also be a great means for CAE simulation democratization when created for non-expert users to explore the design space and perform analyses because they simulate quickly... and they deploy easily.

If we go back to the challenges faced by CAE based engineering listed before, fig. 10 complements the list of assets offered by ML to address those challenges when it comes to making data available. And, besides this fact, the approach produces answers in seconds taking advantage of data which was so far deemed invaluable and quite often deleted, with a consistent engineering judgment and reliably predictive models. Hexagon | MSC Software and CADLM's partnership has produced our I-CAE SCALE initiative in order to bring to market the most advanced ROM and ML technologies at the service of CAE. This allows both companies to merge their long time pioneering positions in CAE (Hexagon | MSC Software) and innovation capacity and highly performing solutions Machine Learning (CADLM) to a community which is eager but hesitant to exploit fully the potential of the expected digital twin adoption.
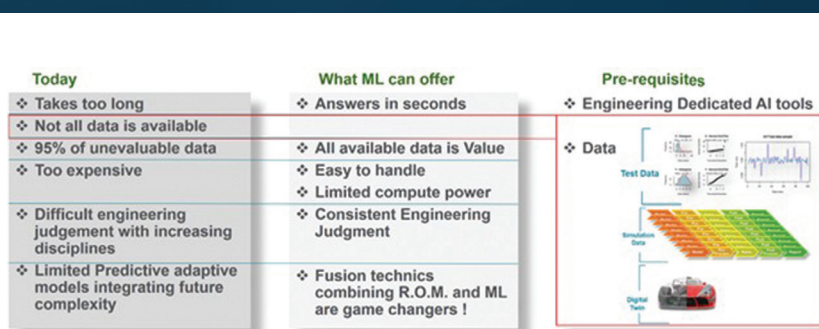

Figure 10: AI/ML addressing challenges Physics-based Engineering is facing today

In particular, CADLM provides the following elements into the collaboration:

- An Open and solver agnostic ML/ROM solution for nearly all CAE fields such as FEM, FVM, System modelling and Data Mining. This also involves interfacing ODYSSEE technology, CADLM's unified platform for analysis and development, with all Hexagon | MSC Software solvers and pre-post solutions.
- Accelerated product design and development via real-time parametric simulations with optimization, machine learning and AI tools. By real time, it is based on the understanding that design and optimization may no more be considered as two separate domains and need to be merged into one interactive environment allowing for a fast and efficient question and answer platform.

This fusion of technologies creates a comprehensive series of dedicated applications with proven solutions (use cases) for various CAE domains ranging from mechanical to crash, from CFD to vibro-acoustics and including multi-physics and multi-disciplinary optimization.

Summary: We are at the beginning of an exciting and valuable AI in CAE journey with all its challenges

It is of course reasonable to say that anything that can be done well with physics-based CAE simulation, should be. A great advantage of any physics-based modelling approach is that it is generally less biased than data-driven models since it is governed by the laws of nature. However, in some cases, AI is the only way forward and in others it is essential to use in order to reduce the time and resources required to solve complex engineering problems and finally embed a digital twin throughout a product's full lifecycle. Leading manufacturers recognize that they can no longer afford the "build it and tweak it" approach that has long characterized many design projects. They have implemented rigorous systems-oriented design processes that harness the complexities of multi-disciplinary product design. The CAE industry therefore needs to evolve with the growing expectation across its many disciplines because systems-oriented approaches make conventional 'engineering judgement' less feasible and less scalable. CAE simulation cannot be too expensive, or take too long, or the whole opportunity can't come together as it should. A really great example of addressing CAE complexity is Integrated Computational Materials Engineering (ICME). Today, we can use calibrated multi-scale modelling in Digimat and MaterialCenter to predict how a new composite will behave as manufactured. This means new customer projects are using machine learning to decide which material system should be used to make a specific part with a specific process. Hence, in Additive manufacturing (AM) for instance, we are today optimizing toolpaths for quality and weight – automatically laying fibers with the best possible alignment.

AI in CAE doesn't come without challenges: awareness and understanding are important in equal measure. AI will only be used if the CAE simulation user is satisfied by the result. That means understanding where and how it is beneficial and how to combine it with physics-based simulation. For an engineer to trust a data-driven model that uses machine learning, they must have a basic understanding of how the algorithms work. Today, most of the advanced machine learning algorithms work like black boxes – clearly someone directly
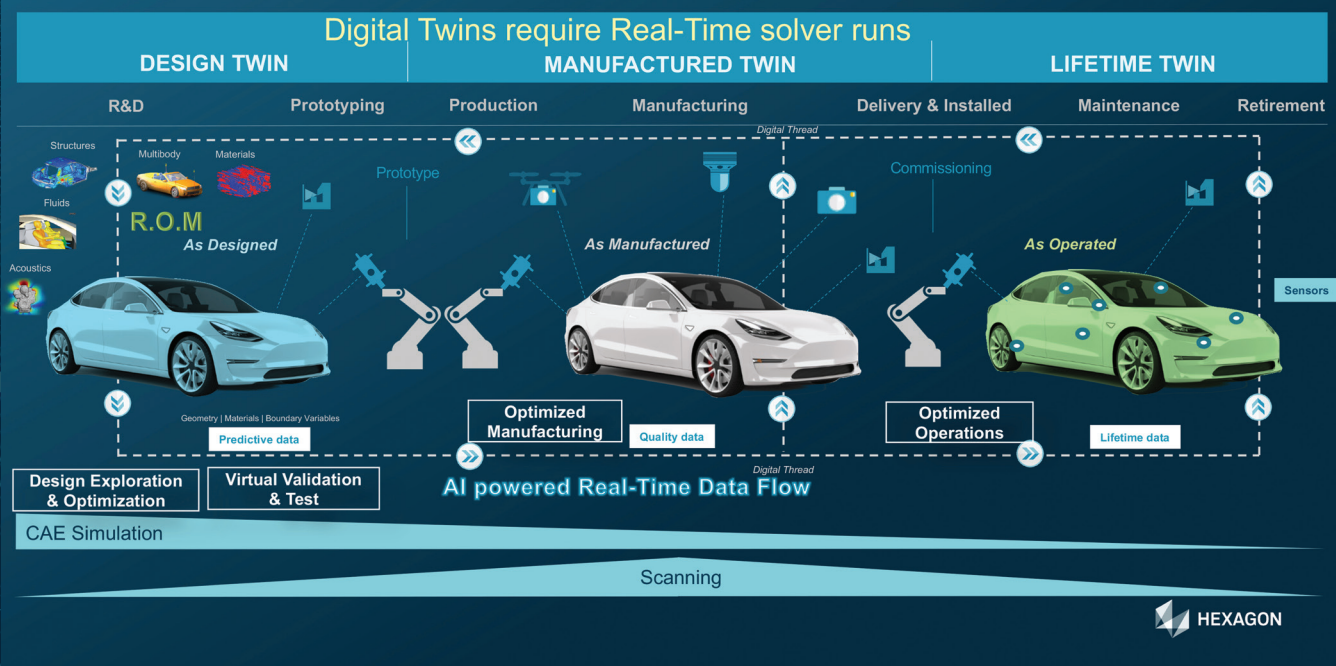

Figure 11: It requires real-time solver runs to move from Digital models towards Digital Twins

involved in the engineering workflow or tool setup must understand enough to ensure they are solving the right problem in the right way. Once the AI model is properly trained, however, it invariably is very useful for making predictions and inferences. However, if tools are used blindly by a designer, then they must be validated and only used in the context they are set up to do. If they are not carefully monitored, predictive models can also become biased over time depending on the data they are "fed" with.

Something else to be emphasized is that a model will only know what you teach it – AI can't generalize like a human if it encounters a problem that is unforeseen. This of course is a serious issue for the AI drivers in autonomous vehicles. However, missing data or noise can also become an issue for less critical predictive models, for example, 'is the humidity on the factory floor causing porosity in the metal AM process?' If prescriptive models are used to drive decisions or automate processes, then the stakes get higher. This is one of the reasons why it's important to ensure tools are refined with domain knowledge and there is continuous physical validation like environmental sensors or inline CT Scanning. Machine learning can be used to simplify a high-fidelity dynamic model into a ROM that preserves the essential behavior and dominant effects but reduces the solution time or storage capacity required. ML is applicable to multiple range of structural physics, whether linear or non-linear and has tremendous advantage over traditional surface responses because it has the ability to deliver time-dependent responses in the same way a FEM transient analysis would (to the extent of the model reduction assumptions).

To date, the use of high-fidelity CAE simulation has been most effective in the design phase. The application of machine learning will enable digital twins to greatly enhance the entire end-to-end product development process. For example, understanding how the properties of a material transform through the manufacturing cycle can decrease the development time of a product and the amount of material used. Physics-based CAE simulation alone lacks robust mechanisms to assimilate long term historical data and unless the computational efficiency of simulation is improved by several orders of magnitude, the potential of digital twins will remain under-exploited throughout product development lifecycles. In addition to setting up and performing simulations by themselves, CAE analysts with their unique engineering judgements and know-how that they have developed over decades, will soon assume additional responsibilities we believe. They will create, manage and supervise highly automated, AI-powered workflows using CAE tools. In this context, numerical analysts have to acquire the necessary skillset for these kinds of tasks. This in particular includes a working knowledge of machine learning and deep neural networks. ML And AI leads to the right assembled puzzle towards making the Digital Twin concept real and finally fusing both worlds, the digital and the real one (fig. 11).

Conclusion

Computer-aided engineering has been in existence for over half a century and it is mature engineering simulation technology. However, it is largely still used mostly in the early research and design phase of product development with limited synergies between design & engineering, production, manufacturing, deployment, maintenance and retirement/recycling. Moreover, ML and AI in the context of virtual

manufacturing and CAE can shorten the simulation lifecycle dramatically across all industries. It can be the connector between all of the data silos in the virtual and real worlds of modern product design, production and manufacturing. Although we expect to see a rapid growth in the application of these methodologies in the next few years, some key success factors need to be taken care of before AI/ML can be democratized for usage by all design engineers using CAE. The usage of physics-based simulations will continue to increase nominally, but the growth of AI-based methods will increase even more rapidly. It is clear that AI will allow us to move from the traditional paradigm to a brand new one where CAE simulation is used for DOE (Design of Experiments) to feed AI models with data that will then be re-used for much faster runs, improving productivity and allowing for more optimization of products and innovation. This is a paradigm shift from simulation validated by test to DOE-fed AI models validated by simulation and test.

References
1. "Sizing the prize - What's the real value of AI for your business and how can you capitalize?" 2019:
2. https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf
3. 'Winning with AI', MIT Sloan Management Review **2019:**
4. https://sloanreview.mit.edu/projects/winning-with-ai/
5. Tech-Clarity **'Empowering Design Engineer infographic' 2019:**
6. https://tech-clarity.com/design-engineers-infographic/7877
7. Gartner 2019 AI Council Report:
8. https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact
9. "Enabling Accurate Design Decisions While Compressing Engineering Timelines with CADLM Technology", Kambiz Kayvantash, Hemanth Kolera-Gokula, Fabio Scannavino, Manuel Chene & Raoul Spote, 2019, Hexagon|MSC Software:
10. https://www.mscsoftware.com/sites/default/files/enabling-accurate-design-decisions-while-compressing-engineering-timelines-with-cadlm-technology.pdf
11. **"Model Order Reduction Techniques for Real-time Parametric Crash and Safety Simulations",** Kambiz Kayvantash, CARHS Biomechanics conference, Dublin, 2018

Learn more about AI, ML and ROM for CAE simulation, download our free White Paper: www.mscsoftware.com/AI-Whitepaper