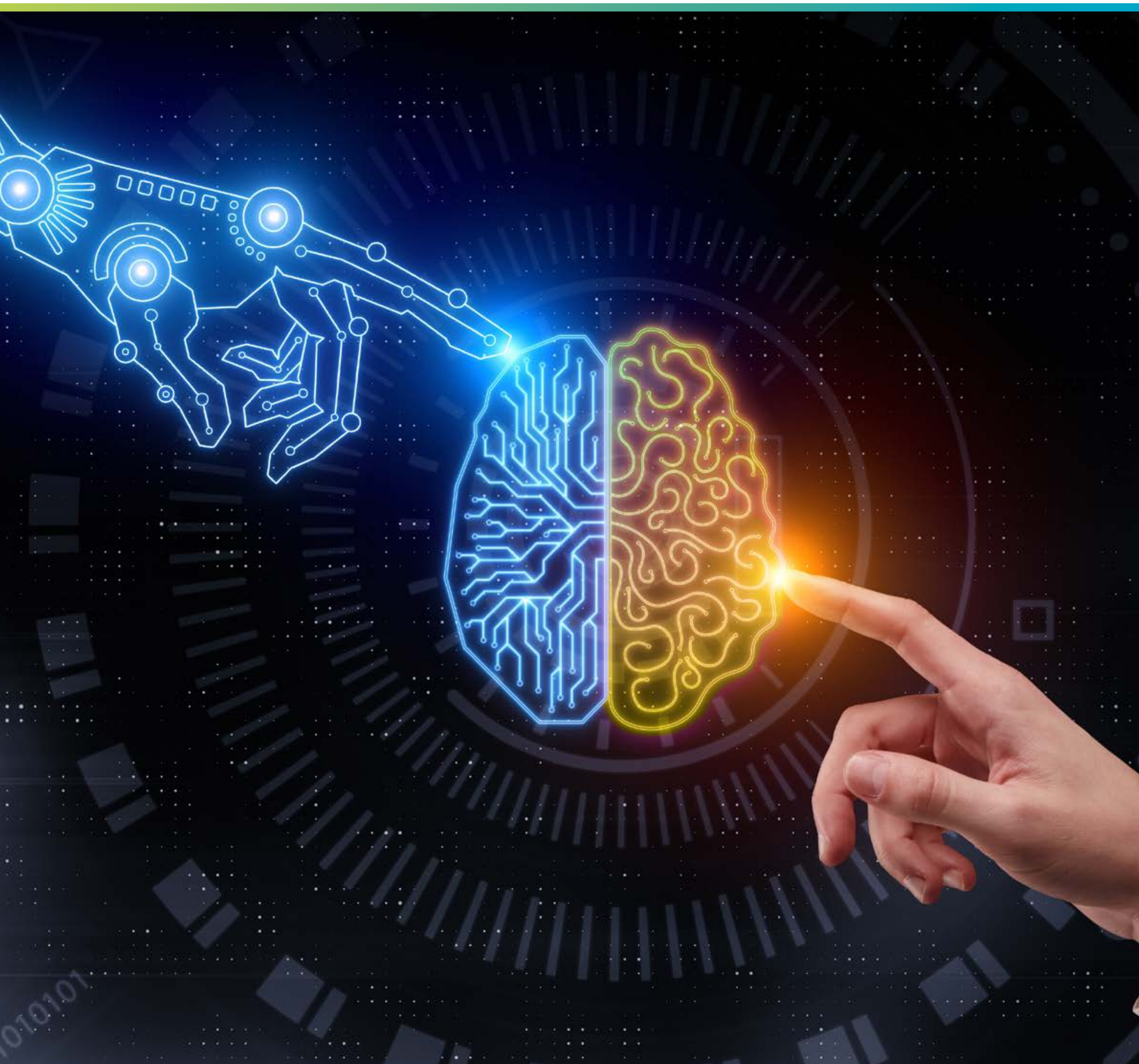
Strategic Partner

White paper

The Emergence of Artificial Intelligence & Machine Learning in CAE Simulation

Kais Bouchiba, Kambiz Kayvantash and Keith Hanna



Executive summary

There can be no doubt that a major change in the way manufactured products are designed and made is unfolding before us in the 21st century as we move deeper into Industry 4.0 with autonomous manufacturing looming 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 and the impact of COVID-19 is accelerating this digital transformation. Manufacturers 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 the development process. This will not be achieved without challenges, but we believe that the tools exist today to overcome these obstacles and connect the 'digital thread' with feed-forwards and feed-backwards loops of real time simulation and measured data that will yield cost savings, higher quality products and high levels of productivity and innovation, yet retaining accuracy.

Computer-aided engineering has been in existence for over half a century and is mature engineering simulation technology. However, it is largely still used mostly in the early design phase with limited synergies between design & engineering, production, manufacturing, deployment, maintenance and retirement/recycling. This white paper outlines what ML and AI is in the context of virtual manufacturing and CAE, and how ML models can shorten the simulation lifecycle dramatically across all industries. 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. It can also be the connector between all of the data silos in the virtual and real world of modern product design, production and manufacturing.

Table of contents

Introduction

In the last few years, machine learning (ML) methods based on deep artificial neural networks (deep learning) have achieved tremendous success in many applications in many industries. Often branded and known as Artificial Intelligence (AI), these methods are able to provide accurate, data-driven process automation. In the context of Computer Aided Engineering (CAE), AI has the potential to speed up the development of tools that allow non-experts to use sophisticated simulation capabilities (so-called 'CAE democratization') to increase employees' productivity, to optimize the computational resources required for the simulations, and to improve the product design process through new insights. As such, this powerful combination of AI and a physics-based simulation approach is well positioned to better address the increasingly complex design problems confronting design engineers today, although it doesn't come without some challenges to overcome.

In a recent report, PricewaterhouseCoopers observes 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]). Our national economies are already undergoing this transformation; the same report estimates that in 2018 alone, AI contributed $2 trillion to global GDP. However, recent 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]). The MIT Sloan Management Review plus 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 their 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. Needless to say, many companies are trying to figure out how to avoid this fate. Moreover, the COVID-19 pandemic that hit the world in early 2020 has wrought seismic changes to the manufacturing environment worldwide and we will likely see an acceleration of digitalization and digital transformation in the next few years according to a recent McKinsey report (ref. [4, 5]).

To better understand what manufacturing companies face today and, more importantly, what they are thinking about when it comes to AI initiatives moving forward, the 2019 Gartner AI Council surveyed more than 170 global enterprise customers to gather their feedback and listen to the challenges and successes companies have had with AI. An analysis of their work revealed ten major keys to AI success (ref. [6]):

Key to Success #1: Setting the tone from the top - Executive Sponsorship Is Critical.

The single biggest predictor of a company generating returns on its investment in AI is that executives commit to sponsoring AI initiatives.

Key to Success #2: Align top-down with bottom-Up.

Successful companies complement strategic direction from the top of the organization together with bottom-up knowledge and experience of people in the business.

Key to Success #3: Early adopters of AI gain a distinct advantage.

Many organizations are falling short of capturing the enormous potential of AI by only piloting a program.

Key to Success #4: Get everybody on the team invested in the process.

In order to build a successful AI program, an organization or division or department must get everyone on board and contributing. This means focusing on interdisciplinary collaboration and bringing together a diversity of perspectives.

Key to Success #5: Educate to participate.

Because exec sponsorship is critical to AI success, executives must be comfortable communicating the value of AI planned for their area of the business. Educating the executive team is essential for success.

Key to Success #6: Insist on AI you can trust.

No AI should be a black box. When using an AI's outcomes to make high-stakes decisions, it's important to know which information it did and did not take into account. Interpretability is key.

Key to Success #7: Understand that data can be like cheese —Messy or Full of Holes.

Having messy data is not a strong enough reason not to initiate an AI project. Data does not have to be perfect to be useful in AI, it just has to be predictive of the outcome.

Key to Success #8: Automation Is key for scaling AI.

By accelerating the work of creating AI, automation generates momentum which itself is of value to the business.

Key to Success #9: AI that is not integrated into the business has little value.

Integrating AI into the business means that data science thinking and modeling flows seamlessly into business processes.

Key to Success #10: Measure ROI and compare to initial estimates.

Clear and continued focus on business value is a trait common to companies succeeding with their AI initiatives.

- However, questions remain with respect to ML and AI and computer-aided engineering:
- Can ML be relevant to serious engineering simulation work and aren't AI and ML mainly about measured data and not about models?
- Isn't CAE all about models in the form of FEA, FSI, and CFD simulation, incapable of replacing physical measurements?
- Can AI enhance CAE simulations that were previously precluded by a shortage of available physical data and are there challenges that must be addressed before AI becomes a standard tool of design engineers?
- Can we talk about an ML-CAE twin and profit from both technologies?

We aim to address all these questions in this White Paper.

The Golden age of physics-based Computer-Aided Engineering

Increasingly, integrating CAE with computer aided design tools (CAD) simulation software has become a fundamental ingredient in the practice of engineering simulation. Finite element methods (FEM) and finite volume methods (FV) have emerged as the principal tools for simulation in many fields of product design and manufacturing. Following their 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 size and computing effort allowing for a higher fidelity in representing real-life engineering problems.

Because of increased efforts, the computing and energy 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 (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 Twins' associated with Industry 4.0 where the last 20 years has seen an acceleration of AI and ML advances. A Digital Twin is 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, prevent downtime, develop new opportunities and even plan for the future by using simulations.

In parallel to the growth of CAE, AI/ML has been advancing very quickly in the last two decades and inventing new methods that address the complexity of the same design problems as those tackled by CAE where they have been applied to areas like finance and marketing mobile applications. In the last couple of years, ML methods based on deep artificial neural networks (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 simulations. Consequently, more recently, ML models are also 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 tools and opening routes to new business models. Because of the above arguments, we are moving today from the traditional CAE paradigm to a new one showing tremendous productivity gains as described in Fig. 2:

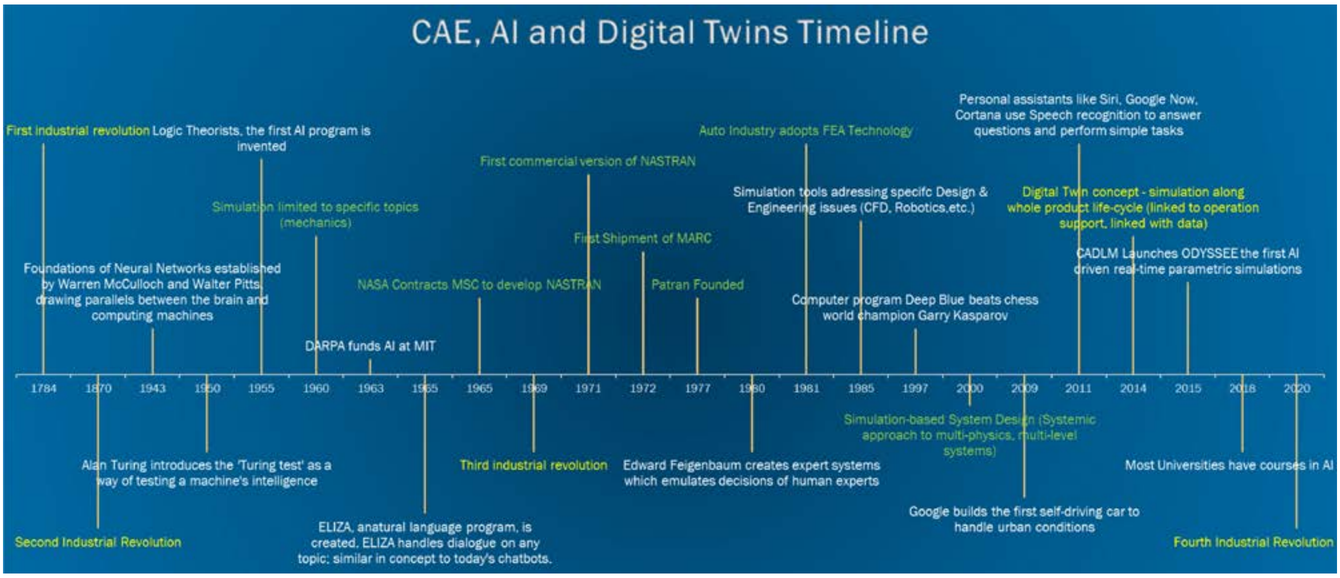


Figure 1: A timeline for the evolution of AI, Digital Twins and CAE Simulation

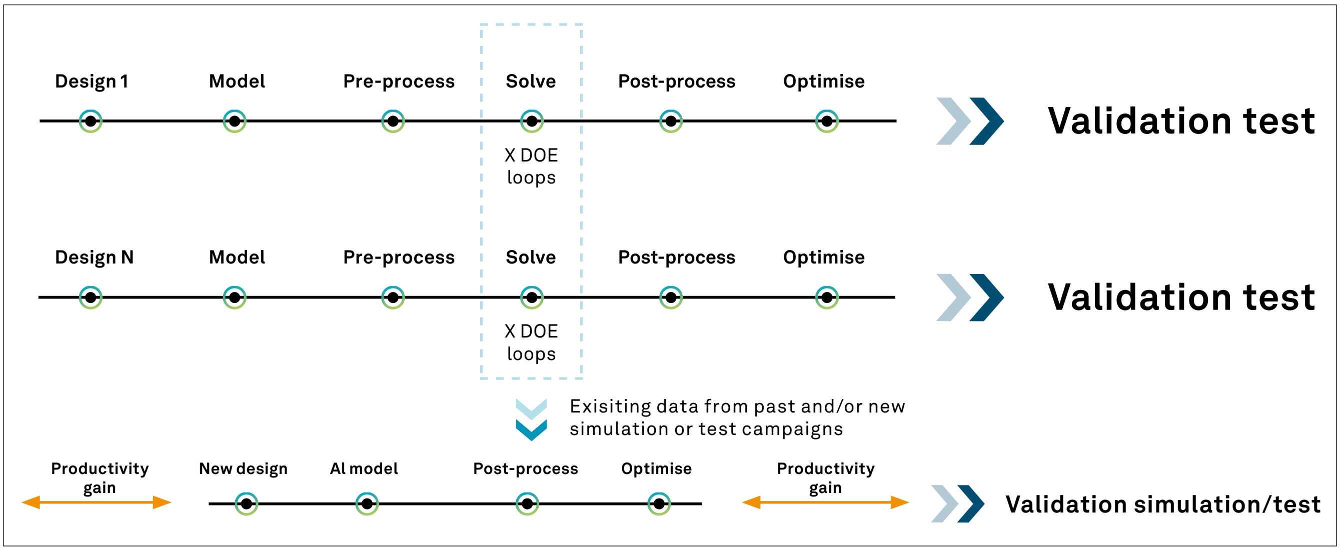


Figure 2: From simulation validated by physical test to Design of Experiment (DOE) fed AI models validated by CAE simulation or physical test

It is reasonable to acknowledge that machine learning will not only enable a new wave of process automation for CAE but it will also speed up the development of simulation tools that allow non-experts to use sophisticated simulation capabilities - what amounts to 'CAE democratization' - with associated new business models. Hence, **a legitimate consideration should be employing AI and Machine Learning as serious CAE options in order to significantly impact product development processes and product life cycle development.**

CAE Challenged by Industry 4.0 & Digital Twins – what is changing?

Manufacturers can no longer afford the 'build it and tweak it' approach that has long characterized by many design projects. As companies face mounting pressure from global competitors, 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. [7]), 80% of respondents believed that product quality is the most important product attribute to keep products competitive. Reliability and cost come next. This indicates customers have high expectations for quality and durability but do not want to overpay. 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, increased product complexity these days makes it hard for engineers to know the full impact of each design decision immediately. Indeed, 76% of survey respondents rate design decisions that affect product

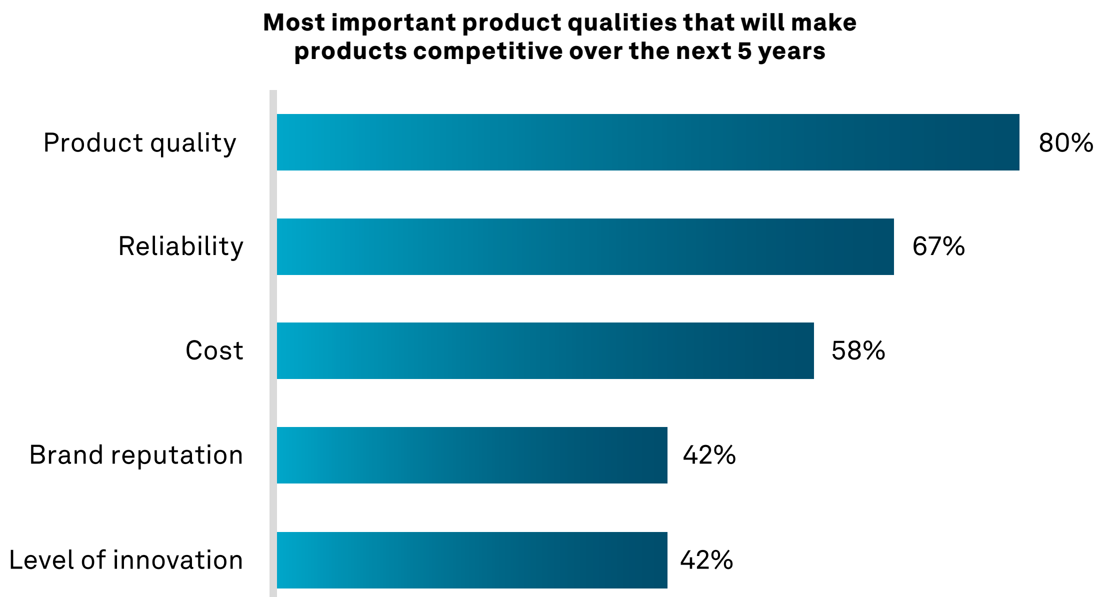


Fig. 3: The most important product qualities that will make products competitive over the next 5 years (Source: Tech-Clarity)

competitiveness as 'somewhat hard' to 'extremely difficult.' This leads many engineers to 'over engineer' their products, which unfortunately drives up cost. Add to this the ever-shrinking timelines in modern manufacturing processes, it means increasing that trade-off decisions can lead to many well-known unfortunate consequences and product recalls and adverse public and media perceptions of the brand.

In today's competitive global market, relying on simulation as a design tool to optimize designs and to provide guidance for product development, engineers have to look for a CAE simulation solution that can offer instant ("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
- Engineering judgement can be difficult with increasing number of disciplines involved, and
- Limited numbers of predictive adaptive 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 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, is to fundamentally improve industrial processes along the entire product lifecycle and manage the increasing complexity of 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.

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 the increasingly circular economy'. While in the first R&D phase, the importance of numerical simulation tools and tests/experiments is undeniable today. In the operational phase, however, the potential for real-time availability of data will open up new avenues for monitoring and improving operations throughout the life cycle of a product. This has huge cost saving and quality implications. Fig. 4 is a schematic illustration of this Industry 4.0 challenge for AI and ML.

It is not sufficient in our opinion to talk about product development processes and product life cycle improvements without bringing into the picture, along with AI, the concept of Digital Twins which have been around for 20 years but have become imperative to many businesses

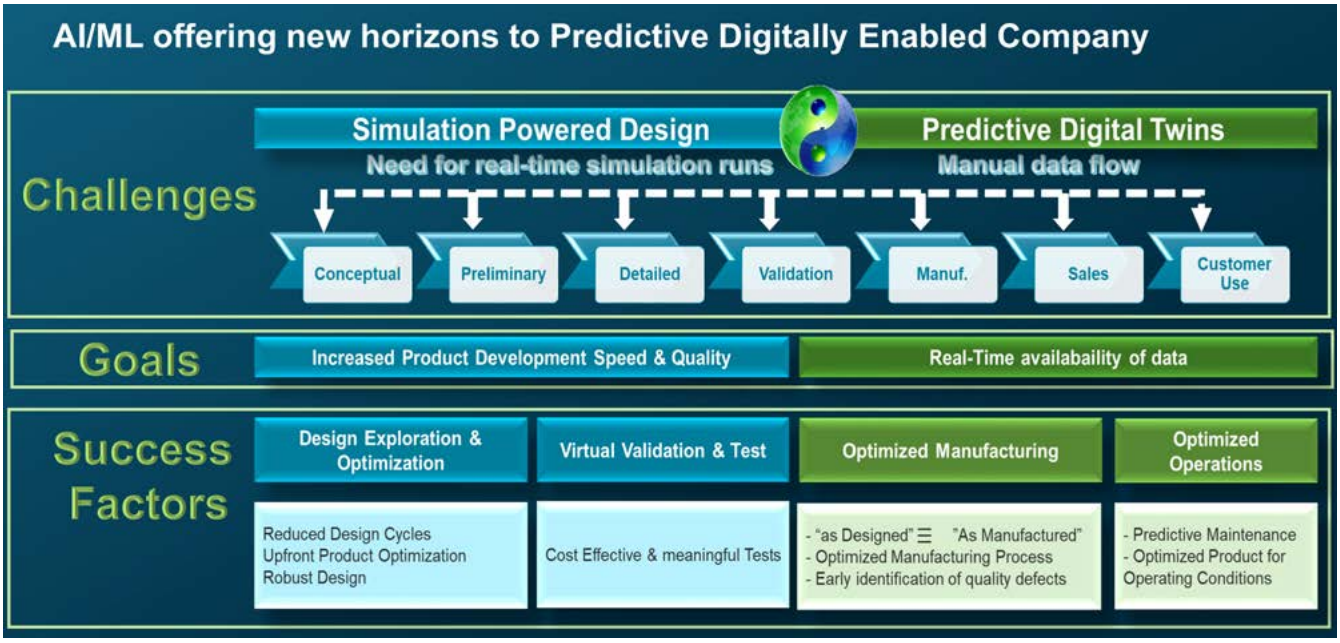


Figure 4: AI/ML offering new horizons to a Predictive Digitally Enabled Company

today. A digital twin is one of the leading terms closely linked to the IoT (Internet of Things) and Industry 4.0. A digital twin is already outlined above and 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.

And digital twins often include laws of physics, 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 use it throughout production and into the aftermarket. A digital twin can therefore be used not only for the maintenance, but also for predictive CAE analysis since it can also contain measurement data from internal sensors. Ultimately, we cannot really understand data without context and intent, and actually the accuracy of the digital twin increases over time as more data refines the AI model. ***Finally, Machine Learning and the Digital Twin interact and improve one another.***

Since the value of digital twins became clear, 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. 8]). 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 simulation data remains essential.

High-fidelity virtual prototypes, or Digital Twins, should become the core of a development process and we can now connect the dots between CAE, AI and Digital Twins.

Beside applications of Digital twins in Smart Cities, Transportation, Meteorology, Healthcare, Education, some of the more advanced deployment of digital twins today are currently found in 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. A good illustration is to compare the virtual test miles driven vs. real miles road tested (Fig. 5). While a 14 Million Km drive would take 9 years for Waymo with real world testing, 13 Million Km of virtual testing would only take one day with an extra parametric scheme making it achievable to move from Passive systems (up to 100's test cases correlating with physical tests) to Active systems (up to 100,000's but not everything is physically tested) to Autonomous

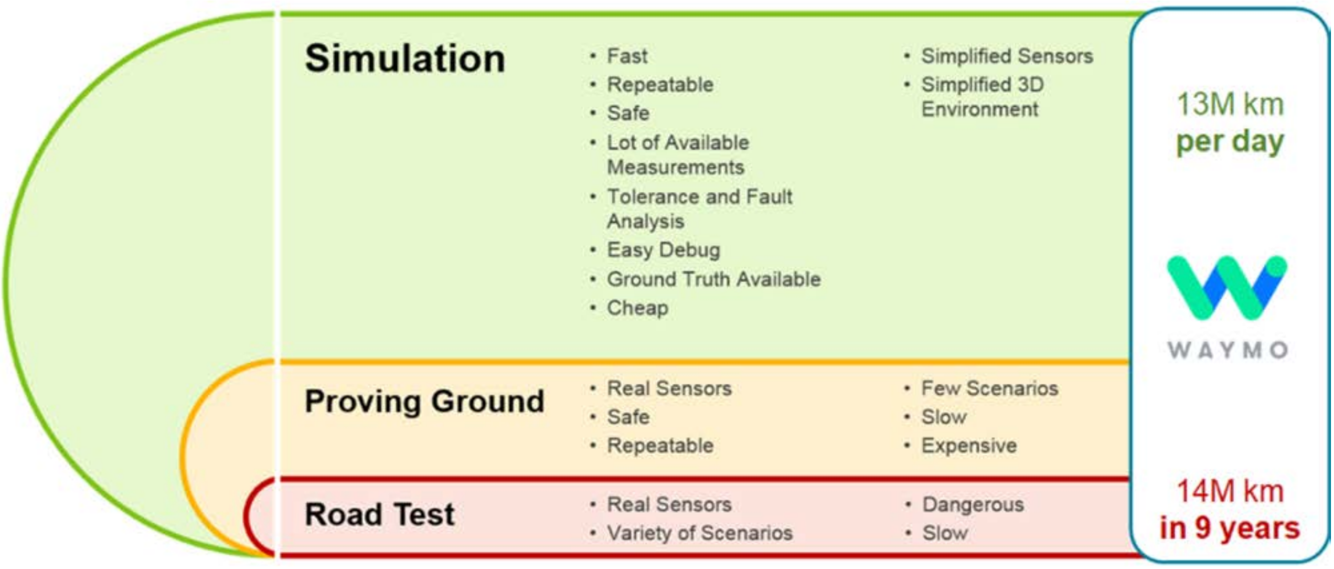


Figure 5: Simulate, simulate... and simulate more (the Waymo paradigm)!

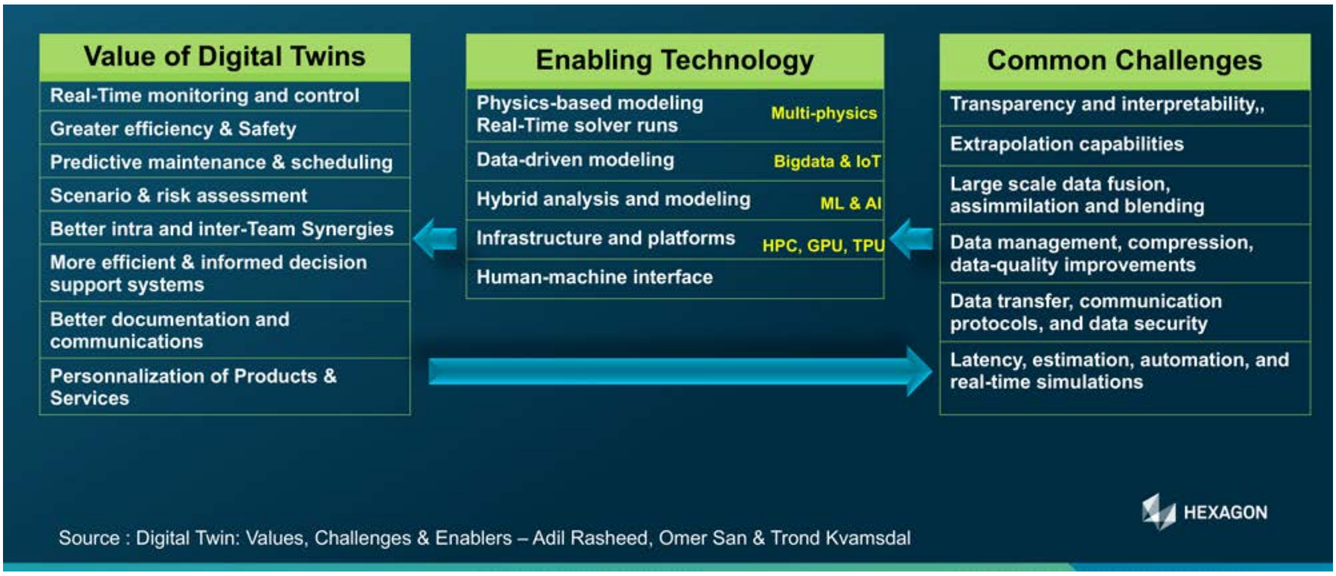


Figure 6: Interaction between Value Generation Challenges and Enabling Technology in Digital Twins

systems with billions of scenarios that cannot realistically be covered by physical tests.

The second benefit of a Digital Twin is the ability to plan and test new features. A digital twin should represent reality, but it can also produce a view into the future. Designers can then virtually create tomorrow's cases for your product and test scenarios. The tests can be tweaked and performed as many times as you like thus finding the most optimal solution that you can then take and make.

Although in their current state Digital Twins can improve Design & Engineering because they are a great and valuable source of data to feed AI models, challenges however remain. Fig. 6 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 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. Indeed, Digital Twins require real-time solver runs. This is where AI and ML techniques applied to CAE bring all the needed and missing value (Fig. 4).

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.

The need for Digital Twin enabling technologies addressing the above challenges becomes a "must" today. In this white paper we will focus on the first two, namely physics-based modeling and data-driven modeling.

1. **Physics-based Modeling:** This approach consists of observing a physical phenomenon of interest, developing a partial understanding of it, formulating the 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 (e.g., MSC Nastran, Marc, Adams, Actran, Romax, Cradle CFD, etc). A great advantage of any physics-based modeling approach is that it is generally less biased than data driven models since they are governed by the laws of nature. However, the choice of which governing equation should be applied for a given physical system might be biased in the sense that different scientists/engineers have different judgments, but this kind of bias is transparent as long as the governing equations are stated. At the same time, however, these models are subject to numerical instability, can be computationally demanding, have huge errors owing to uncertainty in modeling and inputs, and the lack of robust mechanisms to assimilate long term historical data. Another problem associated with numerical modeling is the incompatibility between the way 3D geometries are modeled in CAD-systems and the need for modeling adjustments to solve (eg. de-featuring).

   The very concept of partial differential equations (PDE, as governing equations) has some limitations too. We should recall that a PDE is defined over infinitesimal increments and represents only the "change" and not

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

the physical phenomena itself. Additionally, it is subject to boundary conditions (space or time related) while the real-world is not. At best, a CAE model represents only a small portion of space-time and is subject to the initial constraints imposed on it. This is not the case for real data which are continuous in space and time and vary not only according to a law of infinitesimal changes but also due to real changes in the environmental conditions which can only be investigated if continuous data from the observed environment is available.

Despite their immense success, the use of high-fidelity CAE techniques 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 qualify (many of) them to be denoted "high-fidelity" models that can serve to develop a so called "Reduced Order Models" (ROM) which we will further develop in the coming pages and which may be used efficiently to establish predictive digital twins.

2. **Data-driven Modeling:** While physics-based models are the workhorse of CAE at the design phase, with an increasing supply of data in a digital twin context, open source cutting edge and easy-to-use libraries (eg tensorflow, openAI), cheap computational infrastructure (CPU, GPU and TPU) and high quality, 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 the 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 are 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.

A new approach can be developed we believe to combine physics-based 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.

What is AI/ML and how does it work?

Any ML can be broadly categorized into basically 3 types of techniques (summarized in Fig. 7): supervised learning, unsupervised learning and reinforcement learning

It is worth at this stage to go through each of these topics:

1. **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.

   Supervised learning uses both classification and regression techniques to develop predictive models:

   Classification techniques predict discrete responses—for example, whether an email is genuine or spam, or whether a tumor is cancerous or benign. Classification models classify input data into categories.

Regression techniques predict continuous responses—for example, changes in temperature or fluctuations in power demand.

One of the shortfalls of supervised algorithms is the need of 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.

2. **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

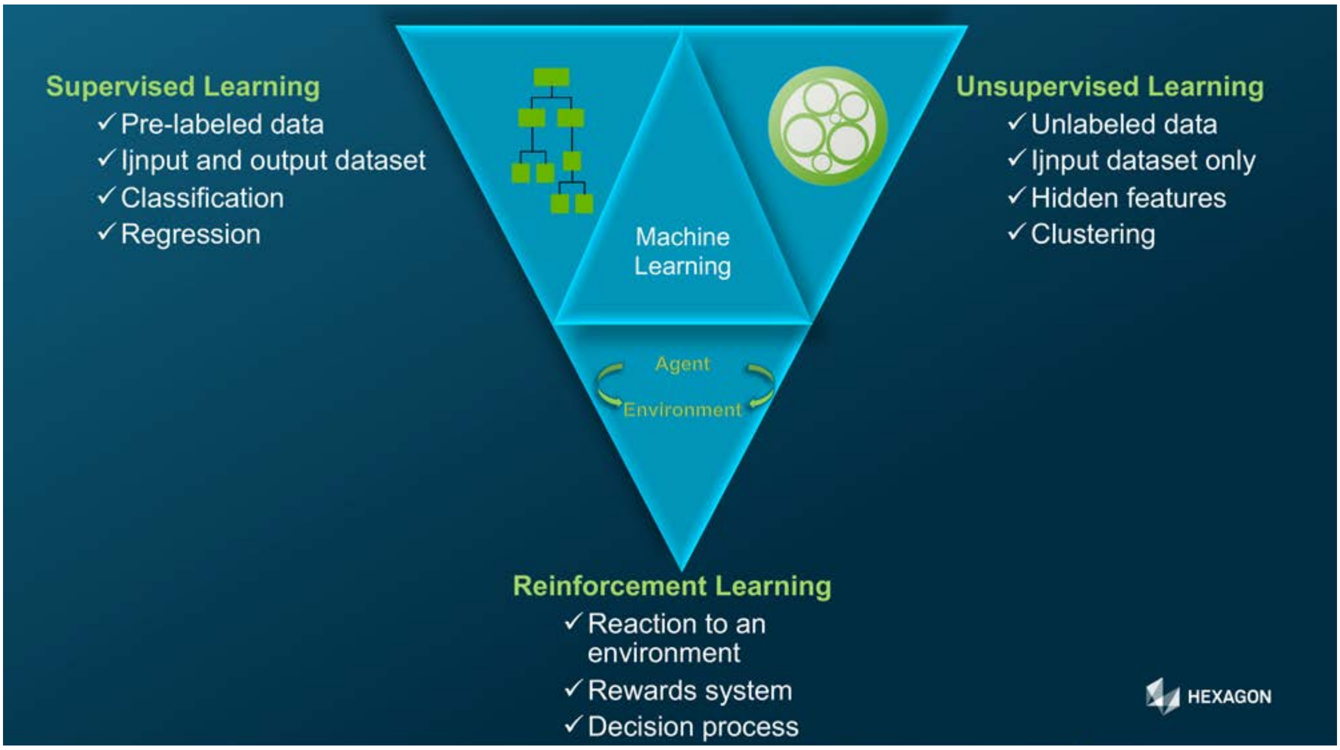


Figure 7: The 3 main Machine Learning technics

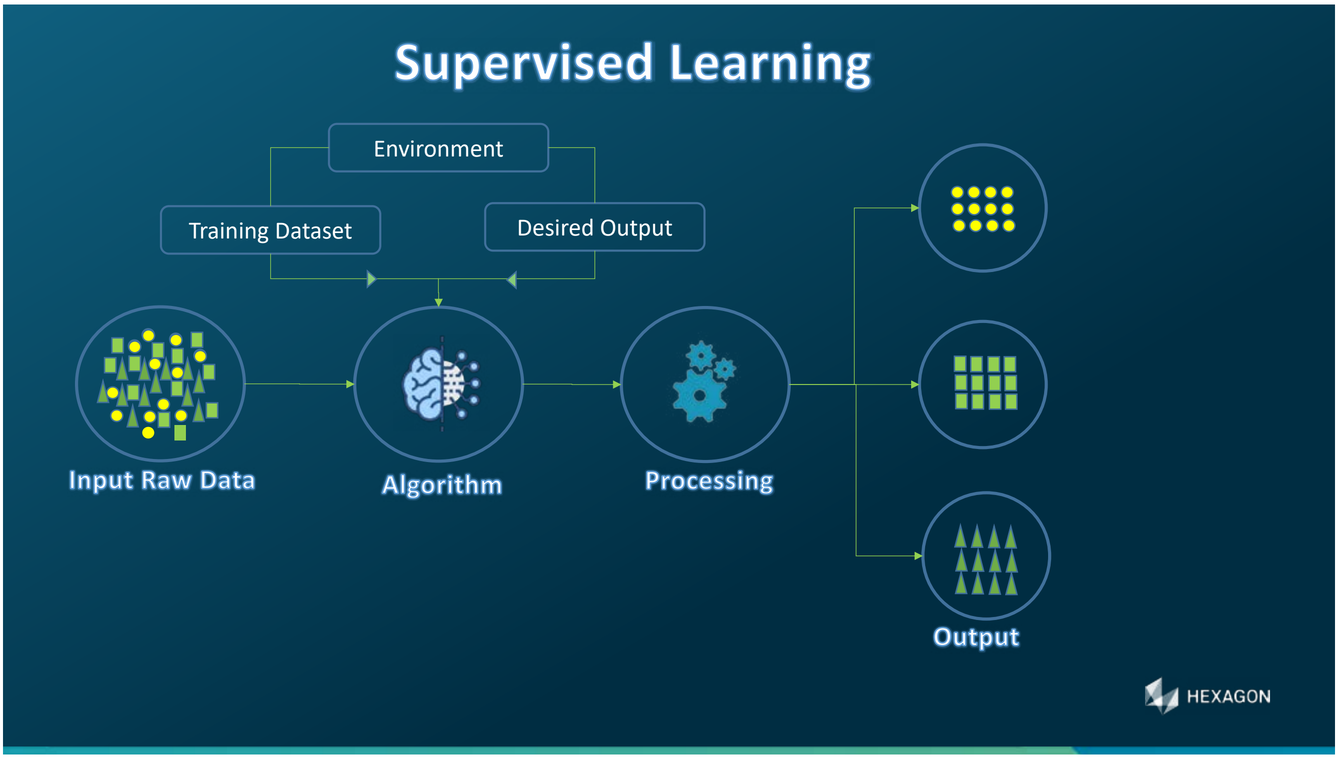


Figure 8: The Supervised learning approach

So, in a nutshell, the 2 approaches can be summarized as follows:

Criteria	Supervised learning	Unepervised learning
Method	Input and output variables given	Only the input data is given
Goal	The output is predicted using the labled input datasheet	The ouput is predicted based on the patterns in the input datasheet

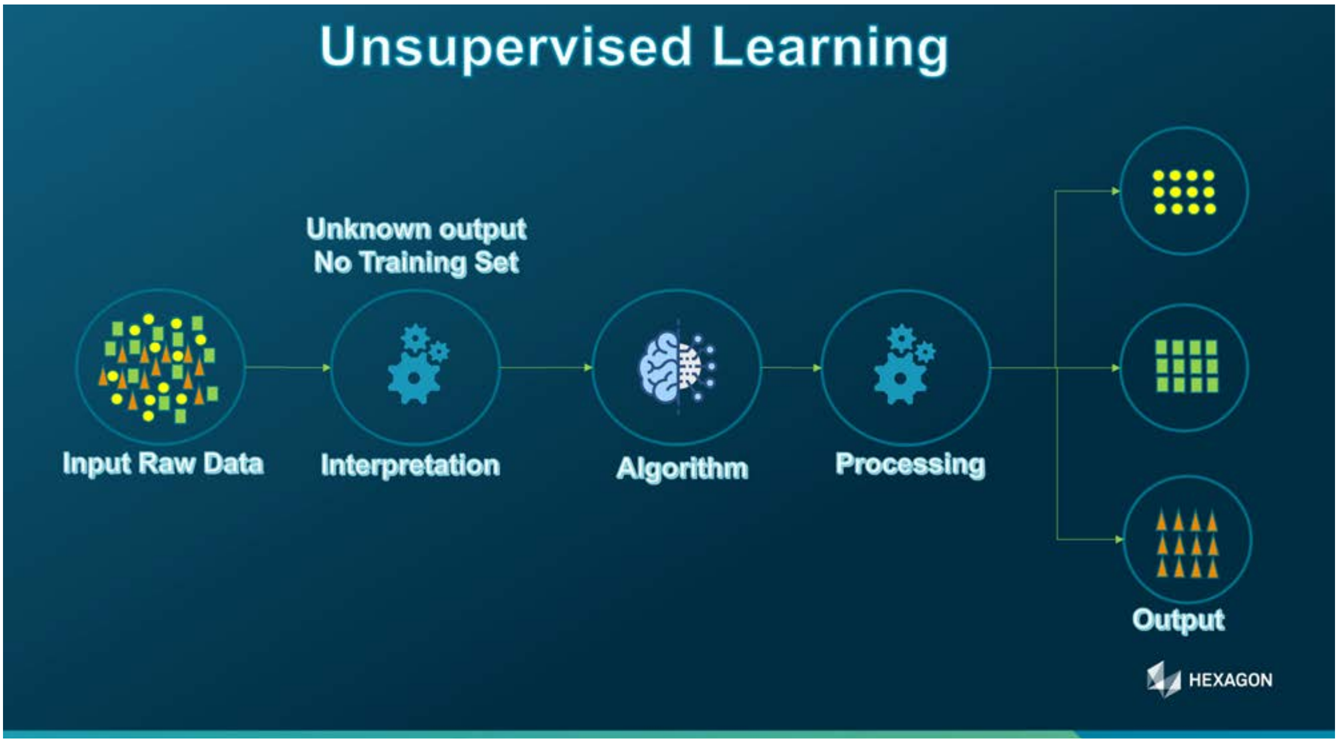


Figure 9: Unsupervised learning approach

Figure 10: Reinforcement learning approach

The following table gives an overview of the difference between supervised, unsupervised and reinforcement learning:

Criteria	Supervised learning	Unupervised learning	
Definition	The machine learns by using labeled data	The machine is trained on unlabeled data without any guidance	An agent interfacts 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 predfined 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

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.

Reinforcement learning: 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. Reinforcement Learning has the potential to help in such a data-deprived situation. Actually, it is based on neither supervised learning nor unsupervised learning... in fact, reinforcement learning, algorithms learn to react to an environment on their own. To be more specific, reinforcement learning 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.

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

Aren't Machine Learning techniques offering a Smarter approach to Design? Doubtless, the answer is 'yes', 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 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 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 have taken 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 approaches. By using multi-scale 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. 11). Then, as real-world 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. Several of these have been small incremental improvements one upon another. These developments can be grouped into a few major types.

1. There has been a substantial increase in the amount of training data underlying the AI models that are interesting to many industrial applications.

2. Computing hardware is significantly more powerful today making it more realistic to train models for longer.

3. Mathematicians have discovered ways to accurately train neural networks with more than one hidden layer (deep learning).

4. 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 is in automation of CAE activities, respectively in the reduction of the duration of these activities. The first development 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 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 (fig. 11). 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 Multiphysics phenomena in practical engineering applications is often particularly challenging, as repeated numerical simulations are required. A remedy is a simplification of the physics-based 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. 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 simulations of multi-body dynamics and test automotive hardware-in-the-loop. It simply was not feasible before with a full 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 at the intersection between now classical pure data-based ML and physics-based modelling (often based on available PDE's) for high fidelity simulations and data drive 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 (Design of Experiment 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 intrusively 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 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

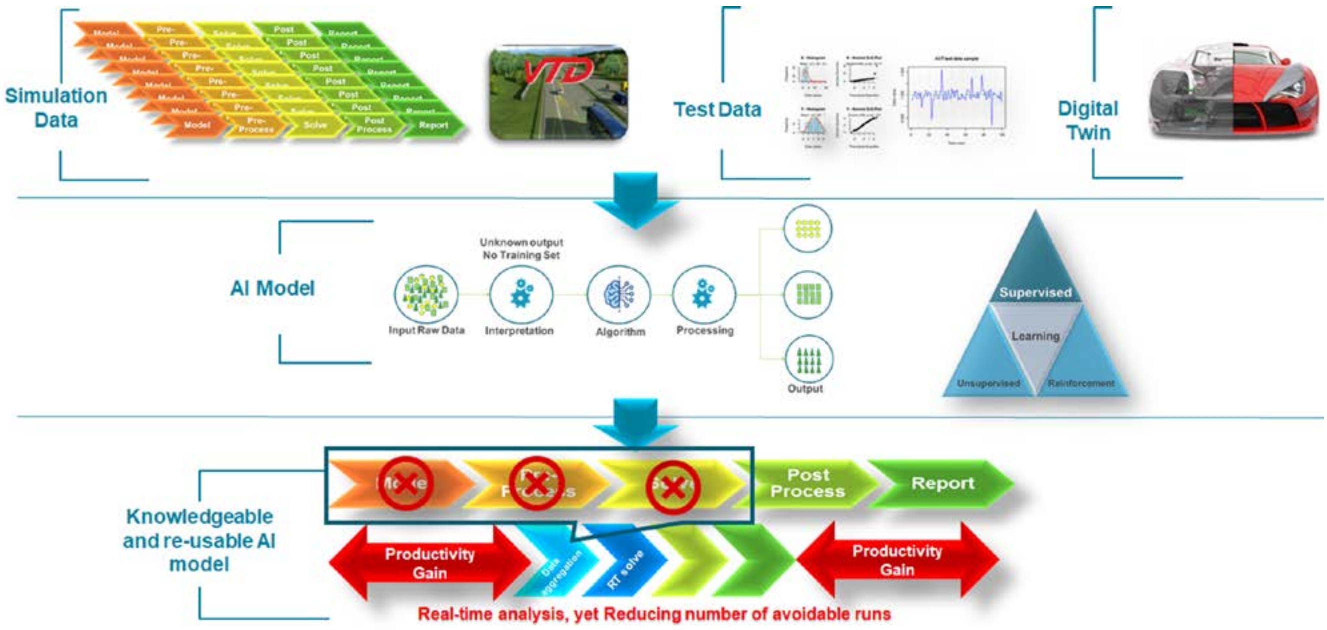


Figure 11: Simulation is the best and most realistic yet cost effective source of data – virtual test driving of car environmental simulation scenarios

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.

One can also claim 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 (ex. Finite element) with huge application potential.

Whatever the affiliation, combining Reduced Order Model (ROM) methods and more traditional Machine Learning 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.

In the following text we will explain how a new and innovative technology based on CADLM's ODYSSEE platform combined with Hexagon | MSC Software solutions can enhance and optimize current traditional approaches, without excessive interfacing and scripting effort. Since the primary goal of all above techniques (fig. 12) 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 on model fidelity.

In the past decade, much effort has been made to develop various methods for model reduction in CAE. CADLM & Hexagon | MSC Software have partnered to develop model reduction approaches for a variety of engineering problems while remaining agnostic to the underlying physics type. Using this approach, the engineering simulation community can 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 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 based on algebraic fitting 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, used to create Reduced Order Models (ROMs) of vehicle behavior (fig. 13) (ref. [9] & [10]).

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, Optimization and uncertainty quantification as well as 3D-0D links.

R.O.M. 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 (D.O.E) are needed to feed the machine learning with data considering: well balanced (space filling) samples of "X" and compute "Y" (fig. 14)

We then achieve:

- Real-time computing – almost zero computing effort for parametric 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 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 are integrated with other system level components for building

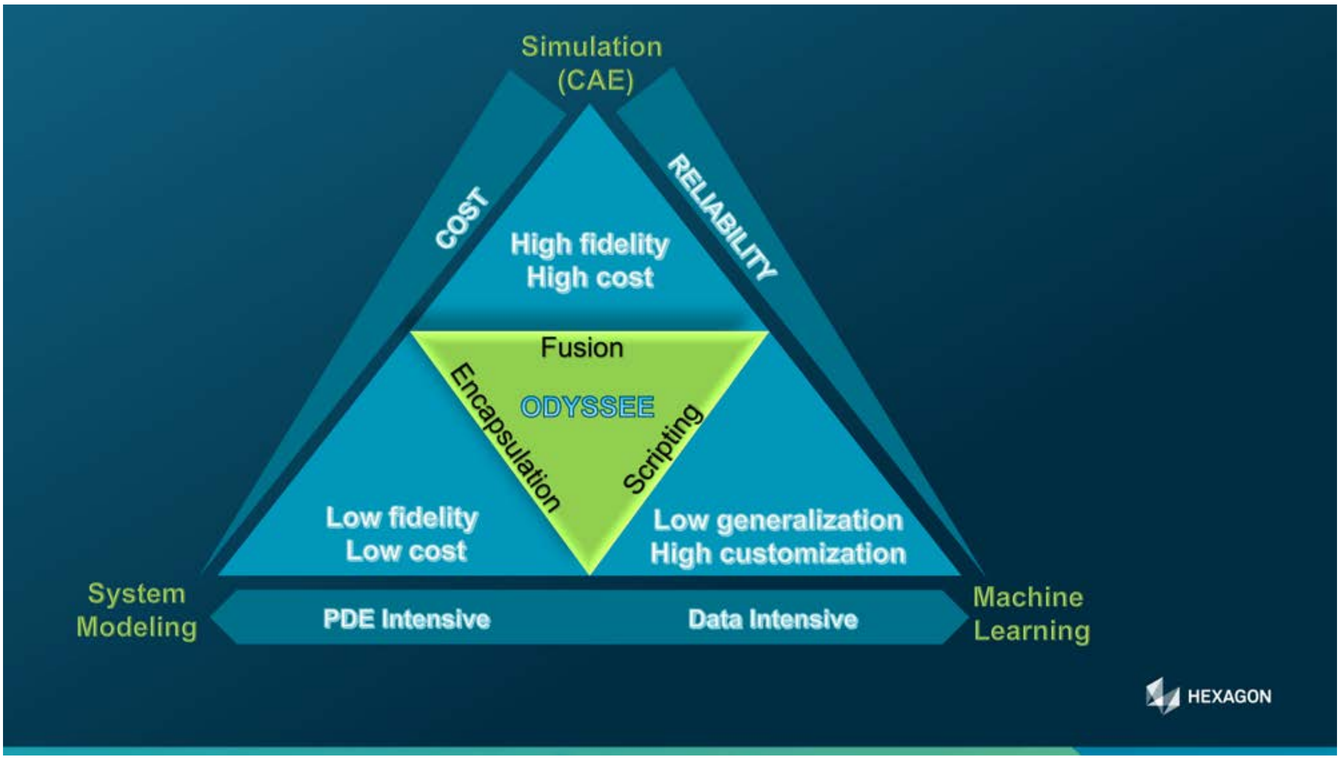


Figure 12: New and innovative AI based approach with CADLM's ODYSSEE platform

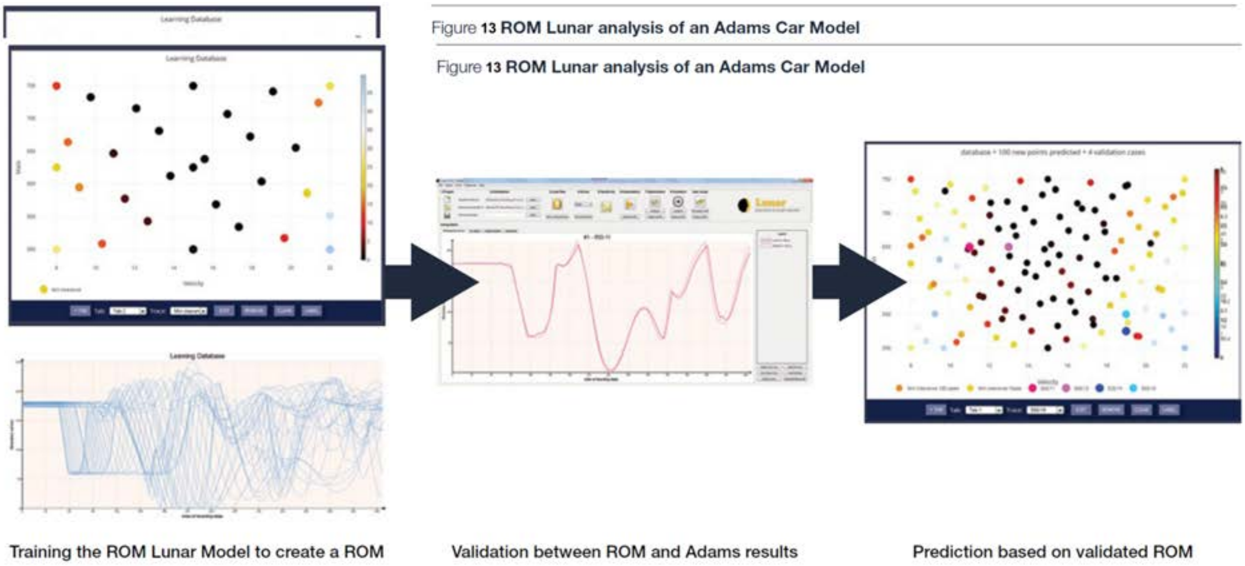


Figure 13: ROM Odyssee Lunar analysis of an Adams Car Suspension Model

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 real and 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 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 get back to the challenges faced by CAE based engineering listed before, fig. 15 is complementing the list of assets offered by ML to address those challenges when it comes to making data available, and besides this fact, it 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.

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. 16):

Fig. 5 illustrated the benefit of virtual test drive vs. real test drive. Developing autonomous driving simulation algorithms is hugely computationally intensive. On the one hand, a real Design of Experiments (DOE) approach is 'brute force' and requires thousands of road test scenarios to be solved in parallel. However, by applying machine learning to the few resulting datasets, data scientists can automate the process of spotting trends and patterns. Such training is intense and requires specialist Graphic Processing Unit (GPU) or Tensor Processing Unit (TPU) hardware that will typically need to be provisioned by a cloud provider, while the output of well-applied AI pinpoints subsequent simulations actually needed to be performed to identify 'edge cases' and avoid testing thousands of unnecessary virtual test drive miles.

At the end of the day, 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. This is a paradigm shift from simulation validated by test to DOE-fed AI models validated by simulation and test.

Strategic Alliance between Hexagon | MSC Software and CADLM: "I-CAE SCALE" to deliver real incremental value

Our **I-CAE SCALE** initiative is born from a Hexagon | MSC Software and CADLM strategic Alliance in order to bring to market the most advance ROM and ML technologies at the service of CAE. This allows both companies to merge their long 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 this expected "Digital Twins" component.

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, FV, 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

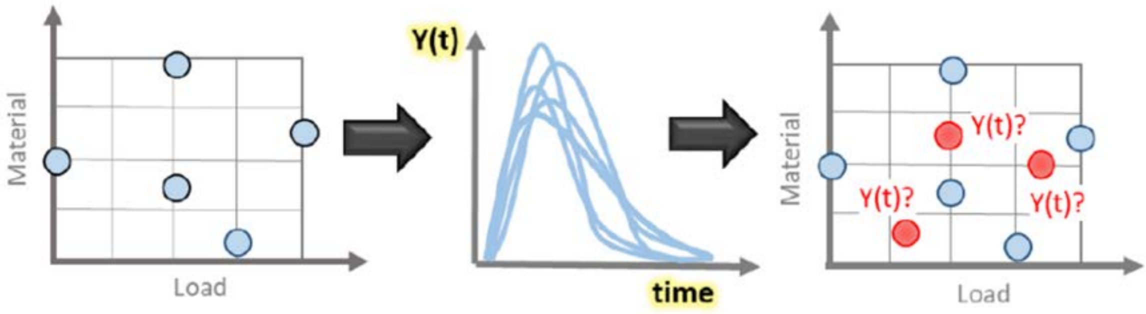


Figure 14: Data sampling through D.O.E. Design

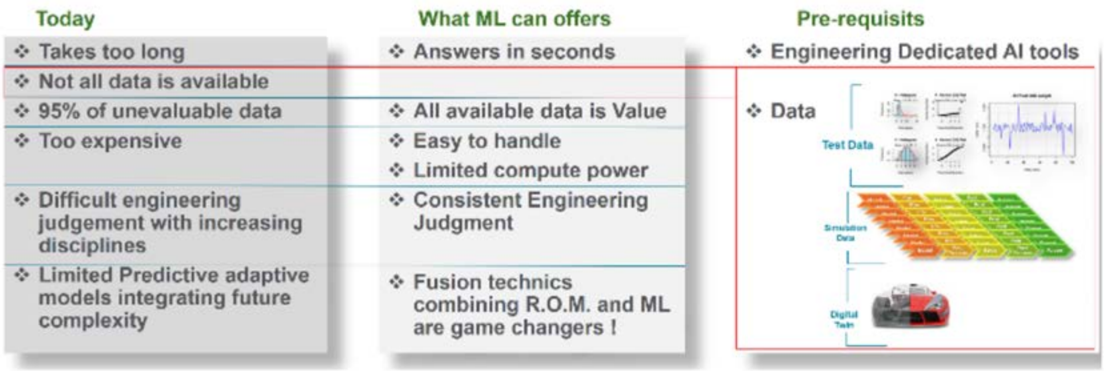


Figure 15: AI/ML addressing challenges Physics-based Engineering is facing today

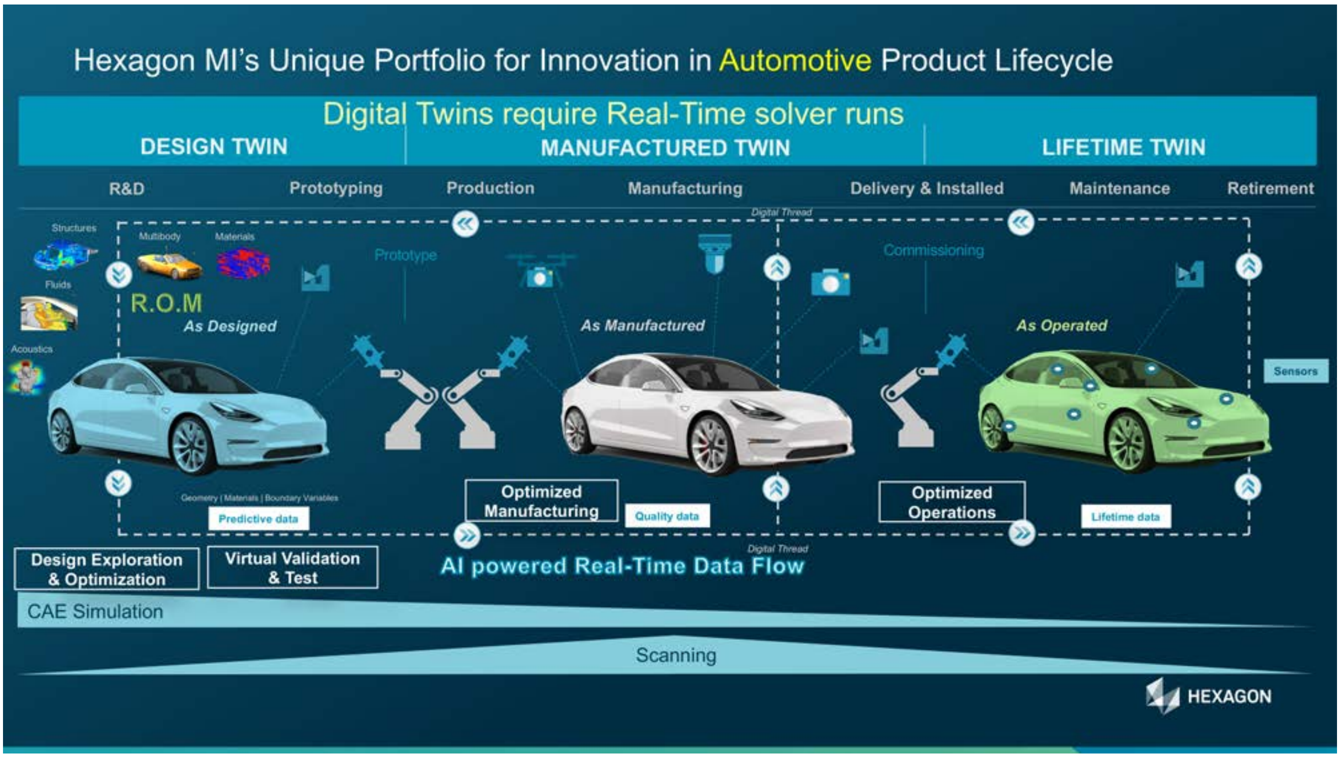


Figure 16: It requires real-time solver runs to move from Digital models towards Digital Twins

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.

Both Hexagon | MSC Software and CADLM collaborate in order to create a comprehensive series of dedicated applications with proven solutions (use cases) for various CAE domains ranging from mechanical to crash, from CFD to vibroacoustic including Multiphysics and multi-disciplinary optimization. Here are a few examples demonstrating the range of the use-cases:

Crash and safety simulation

One of the main bottlenecks in crash simulation is the computational effort due to the “explicit time integration” methods with small time marching steps, thus requiring many iterations. A typical 150 millisecond simulation (approximately the duration of a frontal car crash) may require a couple of hundred thousand of iterations, each iteration running over millions of degrees of freedom. Typical application domains are transport accident reconstruction studies (automotive, rail and aeronautics). Using CADLM’s ODYSSEE Lunar technology the simulation of a car crash has been reduced to last only a few seconds of clock time, making it feasible for parametric design purposes as well as optimization with many hundreds of simulations.

Real (physical) sensor data are rich sources of information when laboratory testing (materials, functionality, fatigue, reliability and homologation) is conducted. In many cases it may be even faster to acquire direct data rather than construct and exploit CAE models. ODYSSEE (and its components Lunar, Quasar and Nova) have proven to be very efficient tools capable of comparing experimental data independent of their nature (measurements, images, sound, etc.). In particular ODYSSEE Quasar technology has been capable of predicting even the experimental results based on the past history of laboratory tests. In the following example this has been applied to child-seat crash testing reducing the test preparation from days to hours.

Automotive Crash

Goal: Real time parametric analysis of crash/safety scenarios

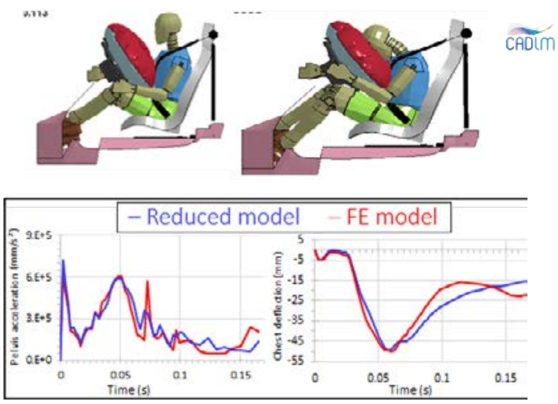
Achievement: Reduce computation time per case from 10hr to 1 second!

Method: ROM + ML (Supervised) using the data of past crash tests

Applications: Explicit/Implicit time dependent and non-linear analysis activities with lengthy time and cost

“ODYSSEE was introduced to the Advanced CAE Division at Toyota Motor Corporation in October 2017.” Toyota Motor Group

Value: Accelerating development speed to market by significantly reducing “what-if” simulations time and associated expenses.



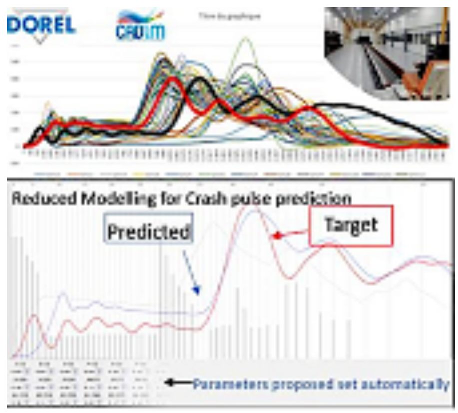
Goal: Reduce preparation effort for sled test + pulse characterization (too complex)

Achievement: Effort reduced from 2 weeks to ½ day. Automated sled parameter (90) settings

Method: CAE + ROM + Inverse Optimization

“The combination of real data and ODYSSEE’s learning capacities has been an important step for the adaptation and permanent improvement of our crash test facilities.”
P. Leman Dorel Europe Laboratory Group

Value: Optimal exploitation of testing facility with reduced preparation time



Fluid-structure interaction

The ALE (Arbitrary Lagrangian-Eulerian) computational effort is even bigger than that of crash. This may be due to high interfacing costs of two different physical problems with different time scales requiring sub-cycling or many other adjustments to the time integration process. A simulation of a ballistic impact (lasting physically a few milliseconds) may take days or week to complete. In the following example this simulation has been reduced to a fraction of second which has made it exploitable for a physics-based simulator used for training in defense or other domains.

Ballistic Missile Real-time Simulation

Goal: Conduct realistic damage modes for defense simulators

Achievement: Real-time simulation with damage modelling

Method: ROM of ALE model, Radial Basis Functions, combined reduced model with domain decomposition techniques

Applications: Gaming, Simulators, ...

Value: Enhance simulator environments with real-time and cost effective "real" physical behavior

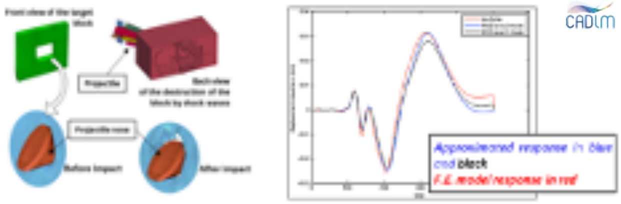


Image-based Fault Detection

One of the major domains of application of ML is understanding the origins of manufacturing faults. In particular, detecting surface imperfections and their origins, as in the following chip manufacturing example, is a challenge to quality production. CADLM's Quasar technology helped with classification of faults in order to understand their origins. Images and processes were recorded, and their interaction was compared, understood and the origin of the problem is identified via reliable ML algorithms developed specifically for this problem. This has helped to remove 50% of the faults within six months of an introduction into the production chain.

Electronics PCB Fault Detection

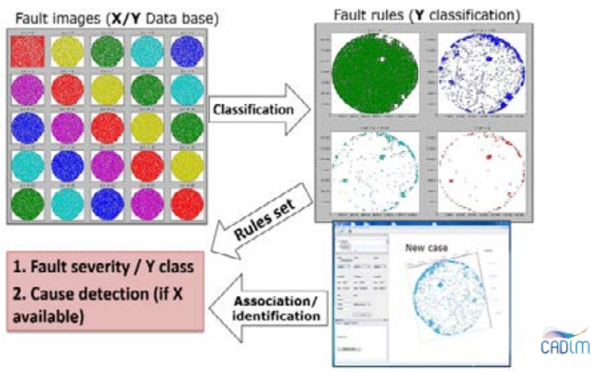
Goal: Determine origin of defects (scratches) and improve quality

Achievement: Classification of fault categories, automated elimination of low-quality products via camera inspection very early in the process

Method: Image treatment, classification (unsupervised)

Applications: health monitoring, quality control, fault detection, image-based prediction

Value: Reduce faulty (low quality) products by photography and ML



Injection simulation process

Injection Process Optimization (IPO) involves finding process parameters for optimal injection performance. IPO of reinforced plastics has a direct financial impact on the involved businesses: Injection Moulded Part quality, reliability, raw material and energy savings etc. Virtual testing is a great candidate for predicting virtually key properties of the injected part like Fiber Orientation Tensor, FOT. FOT has a first order impact on the final mechanical and geometric performance of the injected part. However, Virtual injection process optimization is CPU heavy when based on actual injection simulations. This cost is amplified when dealing with optimization because of the need of running numerous cases.

ML & ROM's using Quasar for emulating actual injection simulations

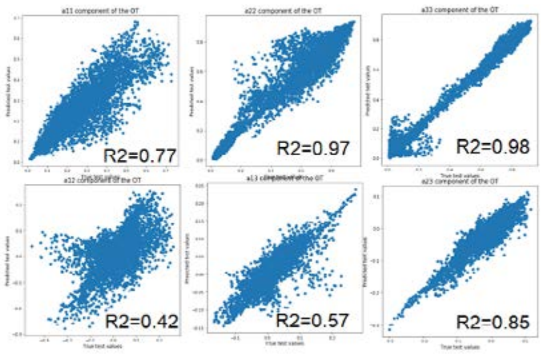
Goal: Predict Fiber Orientation Tensor for exhaustive statistical process scenarios at minimum CPU cost

Method: Quasar (from CADLM) & Reduced Order modeling techniques

Achievement: Stable statistical performance at N=20 simulations when typical DoE studies (Design of Experiment), rely usually on at least N=150 injection simulations.

Applications: Injection simulations process optimization

Value: Real-time optimization of injection process
Cut by more than 5 the requested CPU cost yet keeping a good statistical performance (from 1 week to less than 24 hours)



Autonomous Driving

One of the major obstacles in the field of autonomous driving is that of evaluation of the reliability and uncertainty identification of autonomous sensor systems. There are two ways to do this: either drive the car through millions of kilometers and conduct a real evaluation of all possible scenarios or implement a virtual testing environment based on all data (signals) captured during a few test drives. The data generated needs to be filtered and unified based on fusion technologies in such a way that they become easy to interpret and visualize. CADLM develops state-of-the-art technology integrating various sources of data such as cameras, radar, lidar, etc.

Criteria

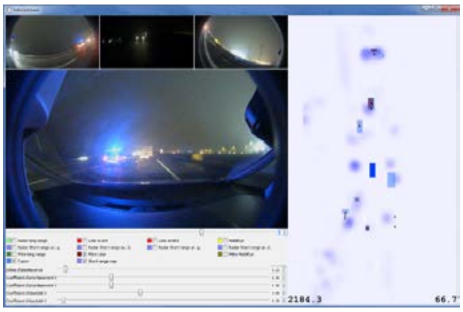
Goal: Fusion of Lidar, Radar, Camera, GPS sensor data

Achievement: reconstruction and visual reconstruction of most realistic scene with visual distance (front/behind of 50 meters)

Method: Non-supervised, signal processing and Kalman Filters for reconstruction of missing information, spatial and temporal filters to interpreted movements from sensors + combination of various proprietary algorithms

Applications: Autonomous driving, dashboard applications

Value: Convert full sensor data source into a usable dashboard view yet maximizing reliability



Biomechanics and Healthcare

CADLM has developed both Finite Element and equivalent ROM models for fast personalized prosthesis design based on real time biomechanics models (also developed by CADLM) as well as efficient image processing solutions exploiting only small data bases for pneumonia detection from MRI's. The ODYSSEE Quasar engine is again applied for an advanced object detection (defect zones) and in conjunction with realistic and personalized CAE models, many solutions for advance medical engineering and diagnostics are provided.

Globally CADLM & Hexagon | MSC Software's AI based technology is available today with many applications in the fields of Structural integrity, Manufacturing, On-board data recording analysis, Health Care, Consumer product, Civil Engineering and Transport. This is a unique and outstandingly mature technology in service of design, manufacturing and life sciences closing the loop for real Digital Twin applications.

Criteria

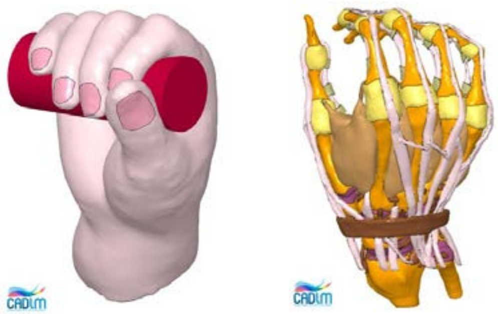
Goal: Provide real-time human body models for prosthesis and ergonomic design applications

Achievement: Encapsulation of complex human body models based on FE simulations

Method: CAE Model based & ROM/ML based compression and prediction techniques

Applications: Medicine, ergonomic, sports, injury assessment

Value: Convert time consuming human body models requiring an implicit/ explicit solver environment



Criteria

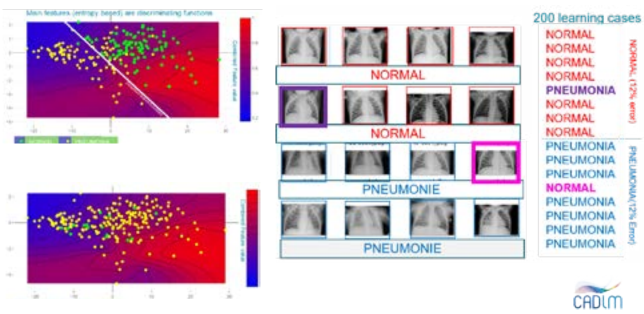
Goal: provide fast and easy to use diagnostics for lung related disease.

Achievement: identification of pneumonia cases exploiting only a small data base (200 individuals) with ~90% accuracy.

Method: Entropy based image compression and object detection

Applications: Medicine

Value: support for early disease identification and associated diagnostic



Summary: 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 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 harnesses the complexities of multi-disciplinary product design. The CAE industry therefore needs to evolve with growing expectations across its many disciplines because systems-oriented approaches make conventional 'engineering judgement' less feasible and less scalable. 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 complexity is Integrated Computational Materials Engineering (ICME). Today, we can use calibrated multi-scale modelling to predict how a new composite will behave as manufactured. New customer projects are using machine learning to decide which material system should be used to make a specific part with a specific process. For example, in Additive Manufacturing (AM) 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 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 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, 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 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.

Conclusions

In the introduction to this white paper we outlined 10 key success factors for the adoption of AI/ML in manufacturing. Let's review them and see how they apply to the CAE industry:

Key success factors from 1 to 4 and 10 involving executive ownership and bottom up as well as top down initiative, go without saying and are valid for any main initiative as long as the value is proven. It will be critical to always start with the business case for CAE and AI with a well-defined business problem where analytics can bring value by showing measurable results.

Key success factors 5 involving education is actually being addressed step-by-step today with the academic curriculum trending towards educating design engineers in the art of data science to enable them to implement best practices in ML. Although it is expected that many steps in the ML model developing process will be automated over the next few years and not only on the solving side, but certainly in increasing ML driven automated pre-processing tasks, the design engineer still needs to have some fundamental knowledge about ML models and how to properly implement them.

Key success factors 6 & 7 involving trusted data imply a significant change in adapting to engineering design best practices in term of Simulation & Process Data Management like that offered by MSC SimManager or MaterialCenter. If we are to rely on virtual data, then we as CAE users must become adept at routinely collecting data and managing it effectively. The volume of data needed to feed machine learning will mandate efficient storage and careful management of data as a strategic asset within an organization. This can only be achieved if data management becomes routine through an automated, and ideally invisible, part of design and engineering workflows. No matter if it is the outcome of a design failure (whatever failure means here; non-optimal, not matching design criteria, etc) or success. The data from CAE simulation runs has to be captured as it is a precious asset to train ML models in best exploring the design space. This change goes along with the Digital-twin constraint, where FE/FV simulation changes from being only a tool in the design cycle to a tool of data generation across it as well, making product design living and functioning as part of an "end-to-end digital twin platform" contributing to fusing the real and the digital world in real time.

Finally, related to success factors 8 & 9, which involves automation and business benefits, are certainly ones that Hexagon as a company is the best equipped to serve the market needs of manufacturing with its unique technology stack. It overlaps the influence of IoT and digital twins on product design and engineering and the intersection of ML towards the goal of fusing both worlds; the digital and the real one. For sure, sensors generating data about a product's performance during operation will need to be integrated in new data management platforms to be used in training ML models for the design of new, improved, highly innovative products, and the performance optimization of existing products in manufacturing operations. This is the brave new world and opportunity that stretches out before us all.

References

1. "Sizing the prize - What's the real value of AI for your business and how can you capitalize?" Price Waterhouse Coopers Report 2019:

https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf

2. "Artificial Intelligence Global Adoption Trends and Strategies", IDC Report 2019:

https://www.idc.com/getdoc.jsp?containerId=US45120919

3. 'Winning with AI', MIT Sloan Management Review 2019:

https://sloanreview.mit.edu/projects/winning-with-ai/

4. McKinsey 2019 AI Council Report:

https://www.mckinsey.com/featured-insights/artificial-intelligence/global-ai-survey-ai-proves-its-worth-but-few-scale-impact

5. "Industry 4.0: Reimagining manufacturing operations after manufacturing operations after COVID-19", McKinsey & Company Report, 2020:

https://www.mckinsey.com/business-functions/operations/our-insights/industry-40-reimagining-manufacturing-operations-after-covid-19

6. "Ten Keys to AI Success in 2020: Insights from Industry Leaders", Datarobot analysis of Gartner Report 2020:

https://www.datarobot.com/resource/aisuccess2020/?cta_id=2020_key_trends&cta_position=post-blog

7. Tech-Clarity 'Empowering Design Engineer infographic' 2019:

https://tech-clarity.com/design-engineers-infographic/7877

8. Gartner: top 10 Strategic Technology Trends for 2019:

https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2019/

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:

https://www.mscsoftware.com/sites/default/files/enabling-accurate-design-decisions-while-compressing-engineering-timelines-with-cadlm-technology.pdf

10. "Model Order Reduction Techniques for Real-time Parametric Crash and Safety Simulations", Kambiz Kayvantash, CARHS Biomechanics conference, Dublin, 2018

Hexagon is a global leader in sensor, software and autonomous solutions. We are putting data to work to boost efficiency, productivity, and quality across industrial, manufacturing, infrastructure, safety, and mobility applications.

Our technologies are shaping urban and production ecosystems to become increasingly connected and autonomous – ensuring a scalable, sustainable future.

MSC Software, part of Hexagon's Manufacturing Intelligence division, is one of the ten original software companies and a global leader in helping product manufacturers to advance their engineering methods with simulation software and services. Learn more at mscsoftware.com. Hexagon's Manufacturing Intelligence division provides solutions that utilise data from design and engineering, production and metrology to make manufacturing smarter.

Learn more about Hexagon (Nasdaq Stockholm: HEXA B) at hexagon.com and follow us @HexagonAB.