DATA DISCIPLINE: MANAGING ENGINEERING DATA FOR AI-POWERED DESIGN

Historically, the field of product design and engineering has found itself behind other industries in terms of artificial intelligence (AI) and machine learning (ML) adoption due to the volume and nature of the data required. Banking, finance, retail, and even futuristic tools like AI-augmented facial recognition utilize mostly two-dimensional datasets. Lightweight data, and often regulations requiring the logging and maintenance of the data for audit, mean these industries have a wealth of historical and real-time data for which machine learning tools can readily be applied.

Conversely, be it for architecture, aerospace, or consumer goods, engineers work regularly with massive amounts of data. The design of three-dimensional products often requires highly complex 3D CAD models, FEA meshes consisting of thousands or millions of elements, high-fidelity simulation of multiple physics, and optimization runs exploring multiple variants of a design. This all adds up to a glut of data and too often, little in the way of an enterprise-level roadmap for how to utilize it, share it with other groups, or even if and where to save it at all.

The advancements in the fields of AI and ML, combined with the increased availability of robust simulation, testing, and field data sets has made engineering data science a critical component of the modern product development lifecycle, but in order to extract maximum value from these exciting tools, companies need a plan to store, manage, and utilize their data efficiently. They need data discipline.

Machine Learning Revolves Around Data

Due to fast advancements in sensor and bandwidth technology, increased storage capacity, and computation power, the use of ML for engineering has become more feasible. Continuous method development and access to open-source codes have opened the doors for companies to experiment with ML and its capabilities. There are, however, still challenges pertaining to data:

Data on demand. When engineers do not have enough data to run an ML model, often the approach is to produce the data by using physics-based simulations. This method can be time consuming and could present issues with a lack of data. For this reason, extensible or adaptive sampling is critical to create the optimal set of data that maximizes the learning and minimizes simulation costs.

Data in hand. This refers to use of the historical data. Often data has been collected over a long period of time, perhaps 30 years or more, and stored in different places and formats created by different software versions. As a result, reliability can be an issue. Additionally, this data needs to be converted to metadata for it to be used by an ML model.

Data in flight. Internet of Things (IoT) sensors are typically responsible for collecting and producing large amounts of fast data from operations such as live telemetry data from F1 cars during a race. Volume and quality of data can be an issue in implementing an effective ML model.

## Machine Learning Challenges within Engineering

Specific to CAE and CAD processes, ML can provide smarter, faster ways to process data and build optimized designs. However, these applications have challenges unique to engineering:

Volume. Incorporating ML methods into various CAE/CAD tasks can require access to large sets of data. Simulations must be run for useful data to be extracted, making this process computationally expensive and time consuming.

Format. Due to the non-tabular nature of CAD, the format of this data poses a problem when being used to teach an ML model. 3D shapes do not translate well to columns and rows, so establishing how to best achieve this is imperative.

Predictions. ML applications within engineering need to predict more than a simple yes or no answer when dealing with simulation. For example, when running a transient simulation, it is possible that millions of elements multiplied by thousands of time-steps need to be considered and calculated.

Variety. One aspect of ML within engineering that is advantageous, however, is the possibility to create intentional design failures through simulation. Unique to this discipline, leveraging this data leads to a more accurate ML model that can be used to further the development of software features and processes.
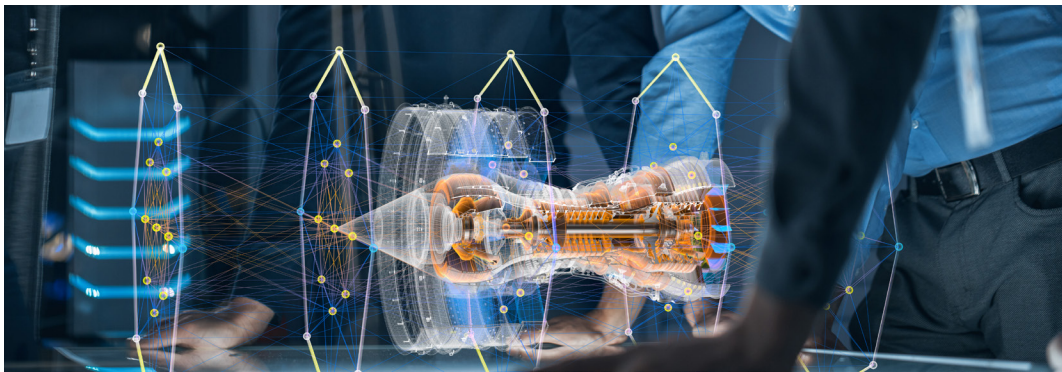
Leveraging Engineering Data for AI-powered Design Innovation

Computer-aided engineering augmented by AI is now offering manufacturers the ability to discover machine learning-guided insights, explore new solutions to complex design problems through physics and AI-driven workflows, and achieve greater product innovation through collaboration and design convergence. Here are a few examples of AI-powered design tools offered by Altair:

Design Generation: AI technology helps product designers explore a broader population of customer pleasing, high performing, and manufacturable new product design alternatives. By applying the same physics-based tools used for verification from concept to design, and through to sign-off and guided by ML using organizational specific constraints, Altair® DesignAI™ enables faster design convergence by confidently rejecting low-potential designs earlier in development cycles.

Design Exploration: For high-fidelity modeling of complex geometries, analysts can use Altair® HyperWorks® Design Explorer, an end-to-end workflow for real time performance prediction and evaluation. Automating repetitive tasks using ML, Design Explorer intuitively performs direct modeling for geometry creation and editing, mid-surface extraction, surface and mid-meshing, mesh quality correction, combined with efficient assembly management and process guidance.

Design Optimization: From design fine-tuning through to design synthesis, including complex multiphysics projects or the study of sets of data, Altair® HyperStudy® helps multidisciplinary teams gain insight from complex models, explore and create new concepts with a variety of inputs, determine best compromises, and support decision-making. Simulation technology combined with design exploration and ML enables engineers to meet time-to-market challenges effectively, and helps teams deliver higher performing products that consider more design dimensions throughout the development process.

Exploiting Private and Hybrid Cloud Solutions

A disciplined data strategy requires HPC tools to orchestrate, visualize, optimize, and analyze demanding workloads, especially the computationally intensive jobs required of some ML-augmented simulation and optimization runs. Altair's comprehensive enterprise computing toolset helps product developers easily access the HPC power they need, migrate to the cloud, and eliminate I/O bottlenecks.

Many organizations struggle to manage and mine data that comes from modern technology platforms. The data arriving into the organization may be in the form of a small amount of very large files, or in the form of millions of very small files arriving every day, or even every minute. Platforms such as Apache Spark™ are seen by data scientists as preferred solutions to manage and process these vast amounts of data to quickly generate insight from data found in distributed file systems. Its ability to work in-memory with extremely large datasets is in part why Spark is included in big data architectures. Altair's workload management tools such as Altair® PBS Professional® enable organizations to work efficiently with big data in high-performance computing, modern processing and storage platforms, and cloud environments.

Altair® Unlimited™ is a turnkey, state-of-the-art private appliance, available in both on-premises and cloud-based formats. Altair Unlimited delivers unlimited use of a wide range of solver technology for simulating mechanics, fluids, electromagnetics, and more — including modeling, visualization, and optimization. To keep it all working at maximum efficiency, HPC resource management and user-friendly web portal software comes included with every system. Altair Unlimited boxes up software, system administration, and infrastructure as a service into a single, intuitive platform.

Cloud Bursting and Migration: Peak-time workload conflicts can be resolved with intuitive cloud bursting, and vendor-agnostic tools like Altair® NavOps® help enterprises automate the migration of compute-intensive HPC workloads to the cloud. NavOps is application, resource, and budget-aware, providing real-time insights into workloads and spending with complete visibility into HPC cloud resources.

By combining sophisticated automation with cloud spend management, organizations can boost efficiency, reduce cloud costs, and improve time-to-results, ultimately improving revenue and profitability.

Big Data Storage and Avoiding I/O Bottlenecks: IT admins can also employ software to find rogue applications and eliminate bottlenecks, take control of file system dependencies, and get live per-job application telemetry for HPC and hybrid cloud. With Altair Breeze™, every engineer is an I/O expert, quickly solving software deployment problems and resolving file and network dependencies. With detailed data for storage exports and summary reports for sharing, Breeze identifies good and bad I/O for easy wins. Because monitoring is increasingly important as distributed systems and compute clusters become more complex, Altair Mistral™ monitors I/O, CPU, and memory, quickly locating rogue jobs and storage bottlenecks and keeping track of what's running on clusters day-to-day.

Managing Seamless Data Access

Equally important to storing the data necessary for machine learning solutions, the data also needs to be accessible to multiple stakeholders along the product development process to ensure it can be linked to ML-augmented CAE to improve engineering processes and extract useful insights.

Altair One™ is a turnkey platform that brings together powerful engineering and analytics tools and scalable computing resources, all in one place. With Altair One you get dynamic, collaborative access with a central data repository. Companies can provision scalable resources in the cloud or on-premises through a single pane of glass, run cloud versions of Altair software like DesignAI, and run simulation jobs in the cloud. Altair One accelerates discovery by broadening access to all the data, software, and HPC resources needed throughout the product development lifecycle.

Working with Altair: Empowering Data Discipline and ML-augmented CAE

Simulation-driven design has changed product development forever, enabling engineers to reduce design iterations and prototype testing. Increasing scientific computing power has expanded the opportunity to apply analysis, making large design studies possible within the timing constraints of a program. Now engineering data science is transforming product development yet again.

Augmented simulation features inside CAE tools are accelerating the design decision process with machine learning. The power of ML-based, AI-powered design combined with physics-based, simulation-driven design leveraging the latest in high-performance computing is just being realized. At the same time, predictive data analytics techniques long associated with business-centric data are being aggressively deployed on asset-centric data to enhance manufacturing, warranty, and testing performance.

To learn more, check out the presentation "The Future of AI in Product Design." A panel of AI experts from Altair discuss the current state of engineering data science and the adoption of augmented simulation, AI-powered design, and predictive data analytics.

AI Panel Discussion
Watch Now