AI for Automotive Manufacturing

A Practical Guide:
Uses, Risks, and Opportunities today

American Automotive Summit

Jon Sobel

SIGHT MACHINE

What is AI?
It’s a production process.

Data

Math

Decision

How is AI different from analytics we've been using for decades?

One definition: AI "reasons like people." (More precise: AI feels like it reasons like people)

AI **imitates** how we think.

But we don't even know how people think.

We are definitely not yet copying how we think.

A second definition is about the mathematical techniques we use for AI

AI generates answers in non-linear, non-deterministic ways.

We often can't go back and traverse how the work goes from data to result.

The more mind-blowing AI techniques are not designed. They are often developed through trial and error

Increasingly, the best approaches involve “ensembles” of techniques, many of which have been used for decades

Technique	What It Does	Uses/Examples
Machine Learning	Prediction in complex environments without explicit programming	• E-commerce recommendations • Credit analysis • Fraud detection • Predicting maintenance failures
Natural Language Processing (a subset of ML)	Generates text using machine learning, statistics, and linguistic rules	• Search • Auto-complete on your phone
Computer Vision	Identifies objects in digital photos. Also combines with other techniques to generate facial recognition (as in airport security lines)	• “Eyes” for robots and self-driving cars • Quality control in manufacturing
Generative AI	Generates text, videos, images, writes software code	• Homework ☺ • Producing software • Orchestrating, reporiting results from other agents
Simulation, Reinforcement Learning	Massive repetition of scenarios with intelligent algorithms	• Weather forecasting • Operations forecasting

In AI, language models have been the star of the show. But many of the interesting issues arise around data*

1980s–Present: Mass digitization of the world's creative material, transactions, and commerce. Digital data has exploded

This data is now aggregated and networked. Depending on its state, data can support analysis and AI in many physical world domains: healthcare, education, industry

Are we entitled to use it?
Is it representative or biased?
Is it enough?
What do we do when it changes?
Is this data suitable for AI?
Can it even be used?

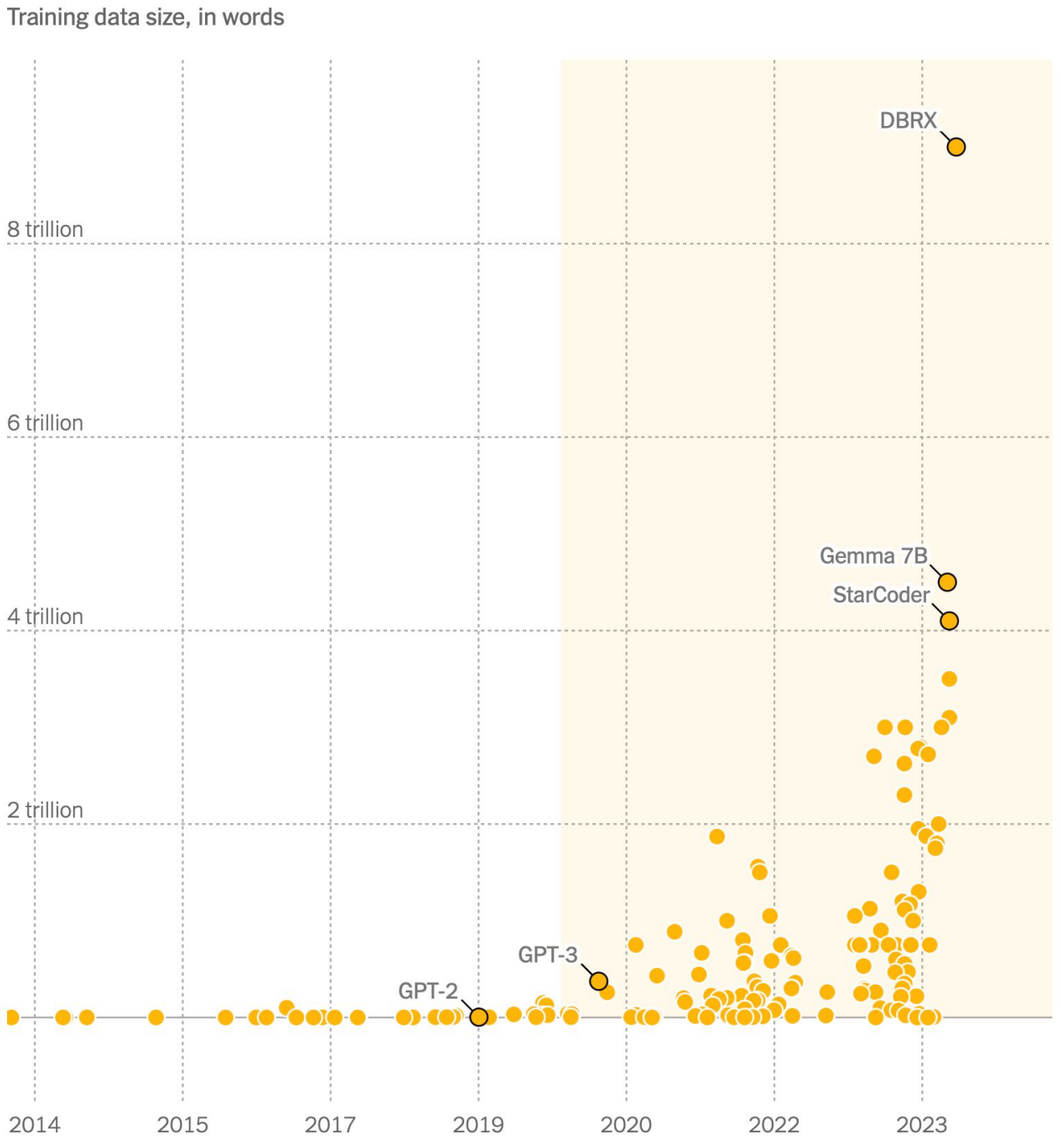
***Must be digital**

Issue #1: Is it enough data? Can we use it?

How Tech Giants Cut Corners to Harvest Data for A.I.

OpenAI, Google, and Meta ignored corporate policies, altered their own rules, and discussed skirting copyright law as they sought online information to train their newest artificial intelligence systems.

The New York Times



Issue #2: In the physical world, it's hard to get enough data
The environment that data represents behaves unexpectedly

Issue #3: physical world data is not like virtual world data. Data is not ready for AI

Whatever Happened to IBM's Watson?

IBM's artificial intelligence was supposed to transform industries and generate riches for the company. Neither has panned out. Now, IBM has settled on a humber vision for Watson.

The New York Times

Remember: LLM's are mostly about words and pictures, trained on the Internet. They are giant "auto-predict" models

Are the results **accurate**?

Fair?

Stable?

Does the use of AI **confuse**, **mislead**, **endanger**, or **inflame**?

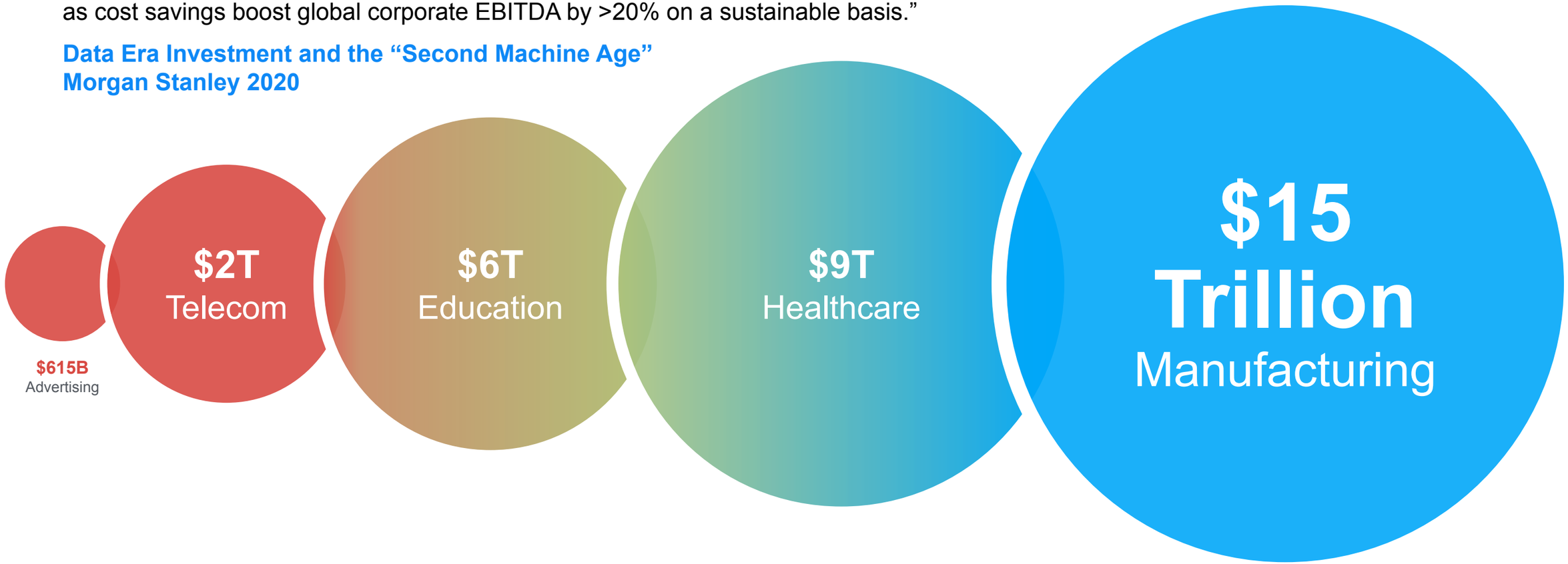
Do we **want** this decision made by AI?

AI in manufacturing: Perhaps the most opportunity of any sector

The most data, the most value — and the least risk

> **The dawning data era investment cycle will be focused on manufacturing ….**
> We think productivity gains worth ~5% of global GDP are possible across all of industries, as cost savings boost global corporate EBITDA by >20% on a sustainable basis."

Data Era Investment and the "Second Machine Age"
Morgan Stanley 2020



Sources: Morgan Stanley, Worldbank.org, GE Minds & Machines

MIT recently concluded “95% of GenAI projects fail” With manufacturing, there are some big differences

- **Abundant “good data” that is mostly numbers. Most AI in plants is not LLMs.** LLMs are useful reporting layers, but don’t do the underlying analysis. The best techniques are not new (e.g., machine learning) and don’t hallucinate. Results are tested through experiments. Plant experiments take minutes and can be run by engineers onsite.

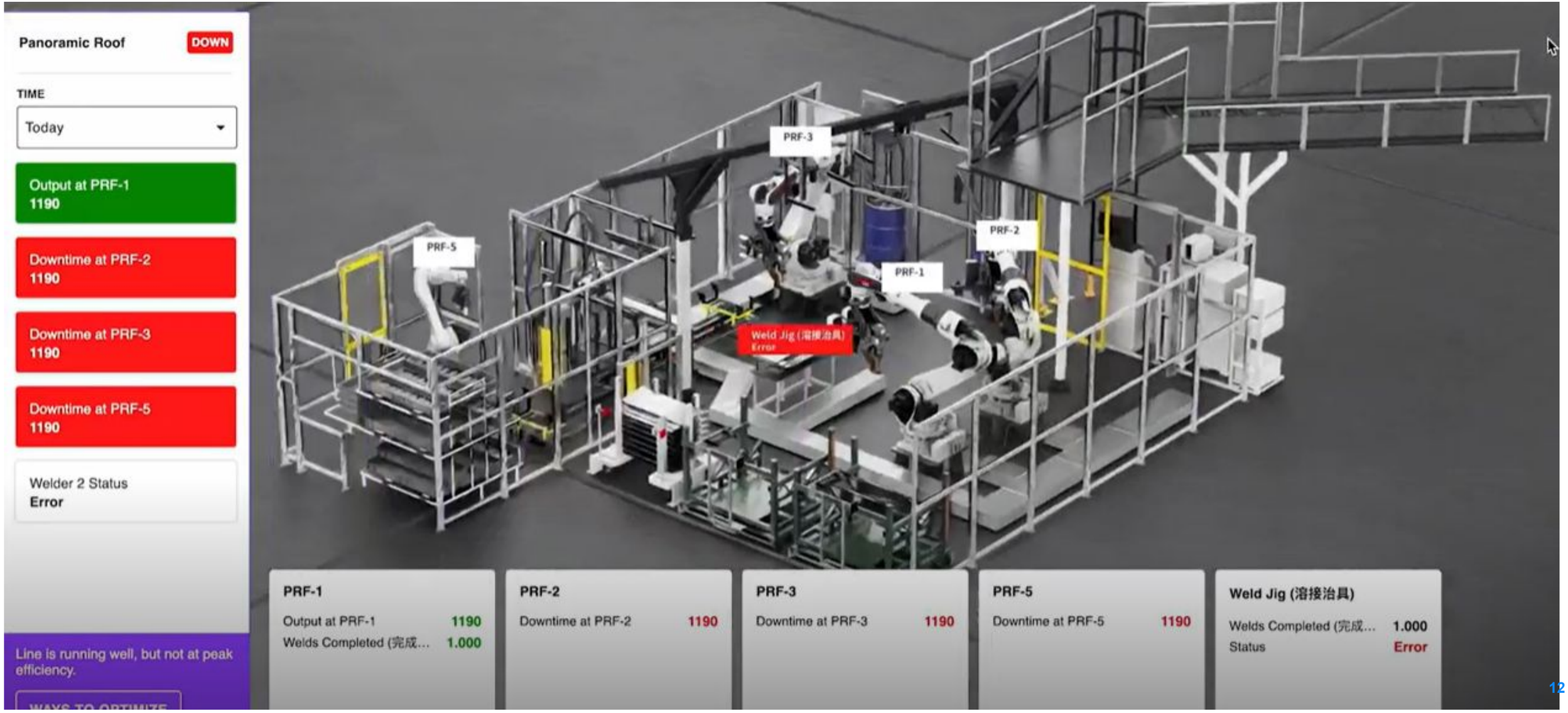
- **Plants start with huge amounts of numerical data. OT data doesn’t start as “good data” – OT data is impossible to work with at scale in its raw state - but with the right solution it can be made into good data. This is the breakout move for AI. Good data unlocks real-time answers to RCA for quality, optimizing plant performance, optimizing energy, process improvements, predictive, visualization and “what-if” simulation through Twins.**

- **Good use cases: In manufacturing AI is applied to complex, well-developed systems. It’s not a party trick**. Manufacturing is one of the best environments for AI because AI is being applied to large, complicated systems that are (kind of) stable and consistent, and we’re looking for explanations of variation

- **Clear goals: the business value is enormous and immediately quantifiable**. Small gains in productivity add up to significant value. 1% improvement in productivity = 3-5% EBITDA gain. Plants putting in full-stack AI (connectivity, data, AI) typically see 5-20% improvement in productivity within a year.

- **Value and fit with workflow. Good data and good AI solve a lot of industry challenges**:
  - Workforce empowerment, hiring, retention
  - Using and optimizing more automation
  - Building in resilience. Being able to rapidly simulate, make “what-if” decisions

Putting it all together:
Real-time data, a 3D twin, prediction, recommendation, simulation



AI for industry: connect plant data, structure, recommend. Results in weeks

Data transformation is automated, early, and universal. Highly variable plant floor data is ingested, labeled, continuously structured into four common building blocks of information (Data Foundation), associated into models (twins), and analyzed in real time

Connect

Securely ingest all plant data

- Energy, Weather, Cost
- ERP
- MES
- Historians
- Quality
- Materials Data
- Controls/Automation

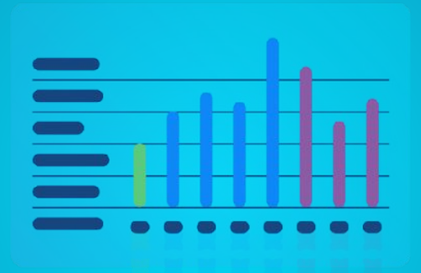
Structure

Automate data transformation

- Stateful, domain-oriented transforms
- Low code/no-code
- Automated feature generation for AI

Recommend

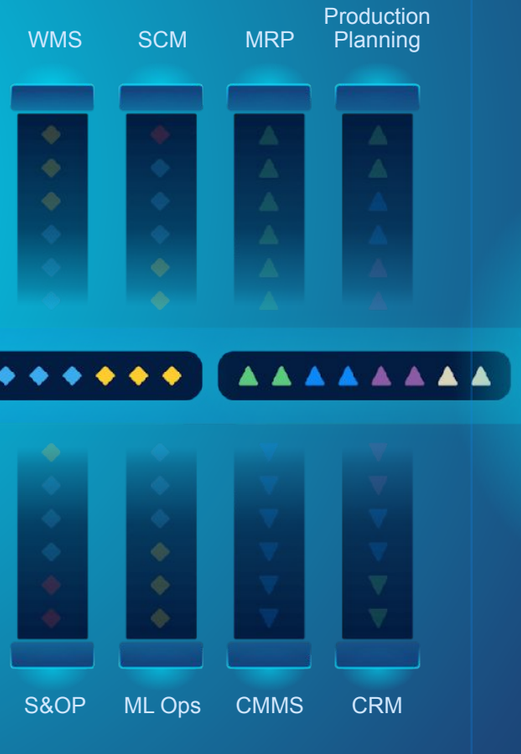
Analyze transformed data with pre-built twins, apps, AI, out-of-box GenAI



Data Foundation
Four domain-specific schemas

Build

Share transformed data with other applications

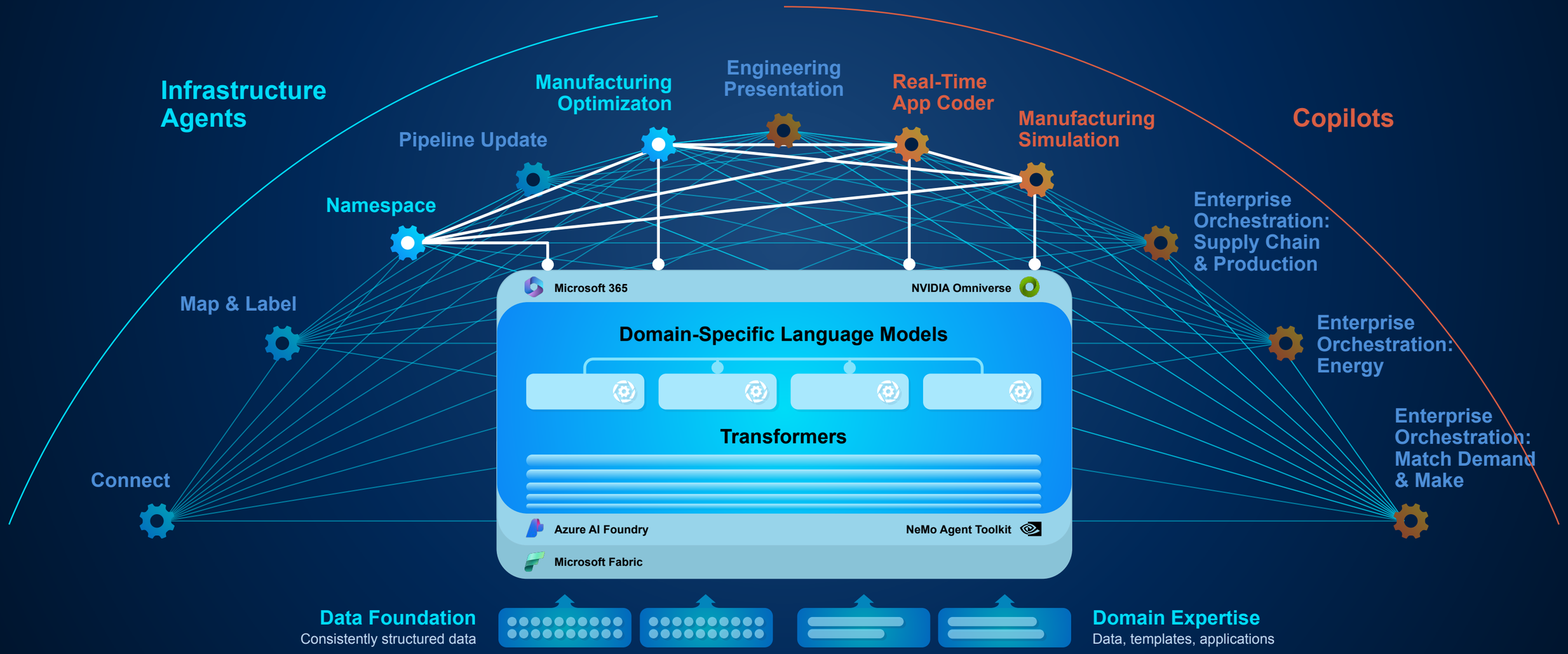


Results

Make data useful for all stakeholders

- Operations Leaders
- Process Engineers
- Operators
- IT Managers
- Data Scientists
- Supply Chain
- Sustainability
- Connected Machines/ After-sales Service
- Value Chain: Cross-firm

Real-time data foundation generates agents, 3D visualization, and simulation
Agents implement and update data foundation and are used to optimize production



Data Foundation supports all production use cases. Structure plant data once, analyze infinitely

	Visibility What's happening?	**Variation** What's changing?	**Cause** What's the cause?	**Rule** What's the rule/recipe?	**Prediction** What will happen?	**Prescription** What steps to take now?	**Optimization** How to optimize the system?
Throughput • Availability • Rate	Global Ops View Asset Throughput Comparison	OEE and KPI Tracker	Alarm Analysis	Centerline Analysis	Predictive Downtime	Prescriptive Uptime	Cookbooks & Recipe
Quality	Top Defect Pareto Daily Meeting Yield Dashboards	Yield Tracker	Root Cause Yield	Centerline Analysis: Recipe Setting	Predictive Yield and Energy Efficiency	Prescriptive Quality	Quality Optimizer, Cookbooks & Recipe
	Descriptive Statistics		**Statistical Inference**			**AI**	
Cost • Energy • Materials • Laborl	Materials Usage	Cost Tracker	Material Reduction, Defect Analysis	Yield Optimization	Predictive Cost	Prescriptive Material Consumption	Energy, Quality Optimization
Flexibility • Changeover • Process Changes • Mix	Changeovers, Network-wide Capacity	Changeover Variation, Network-wide Capacity Tracking	Root Cause Changeover Delays	Run-specific Recipes	Predictive Changeover	Prescriptive Changeover Schedule	Clean in Place (Changeover + Cost)

Lessons from 15 years of bringing AI to plants

- **Beware the glittering AI algorithm.**
  - Set up your pilots so that you are testing for scale during the pilot, not after.
  - Design pilots that demonstrate repeatability across different processes and different data environments.

- **Don't sweat data availability or use case selection. Most plants have abundant data and abundant opportunities to improve. Do think hard about the first plant leader.**

- **Alignment across firm functions will be the most difficult challenge, and it will not get solved fast.**
  - In our company's experience, we typically engage with ~ 15 functions in the manufacturing enterprise
  - Each views data and thinks about how to use data differently.
  - People doing the work in plants – operators and engineers - will be your North Star
  - This is all going to be part of a firmwide data/AI strategy. IT and Ops can build a long-term plan for both
  - The correct paradigm is data. Not software tools

- **Pick a problem to test your approach, pick a date 3 months out or less, set a goal, work backwards.**
  - First projects require 10-15 hours/week in the plant and you can often get funding from Cloud providers
  - Put a clock on it and go.

- **Manufacturing is all numbers. With good solutions, you will know if you have results fast**
  - Manufacturing is the best industry for AI. Results are measurable, concrete, significant.

Thank You

SIGHT MACHINE