

# EE392b

# Seminar on

# Industrial AI

April 1, 2025

Prof. Daniel O'Neill  
dconeill@stanford.edu  
Stanford University

Prof.  
Daniel  
ONeill



- AI for operations, AIOps
- AI preference modeling
- Genomics
- Digital Twins, DT's
- Agentic AI & Mixture of Experts

PhD Stanford EE  
MBA UC Berkeley



ORACLE



# AI Snapshot 1Q 2025



# January 2025: Stargate

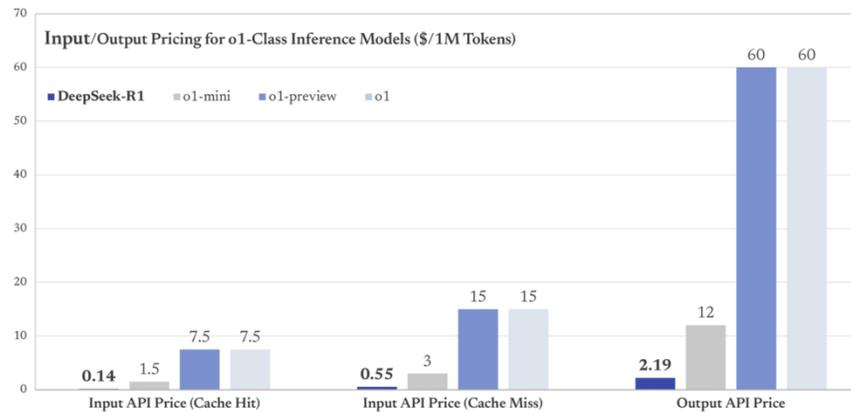
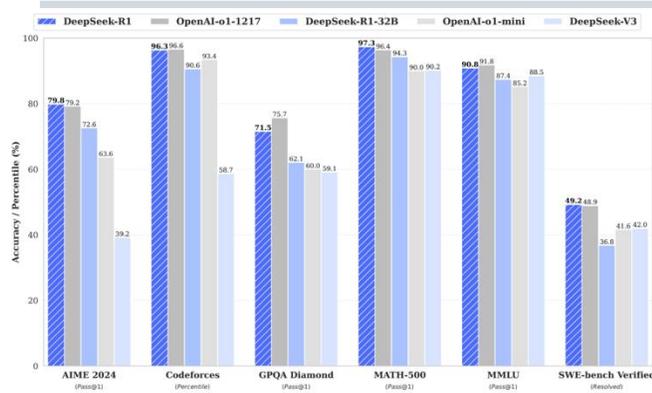
“...initial technology partners will include [OpenAI](#), [Microsoft](#), [Nvidia](#) and [Oracle](#), as well as semiconductor company [Arm](#) and venture capital firms. They said they would invest \$100 billion to start and up to \$500 billion over the next four years.”



- Cost
  - Scale
  - Proprietary
  - Government
  - Industrial Structure
- } Bigger is Better!

# Also in January 2025:

## DeepSeek-R1



- Lower cost
- Comparable performance
- Open source
- Government
- Industry???

# BANG! Bigger is Better – Not so Sure

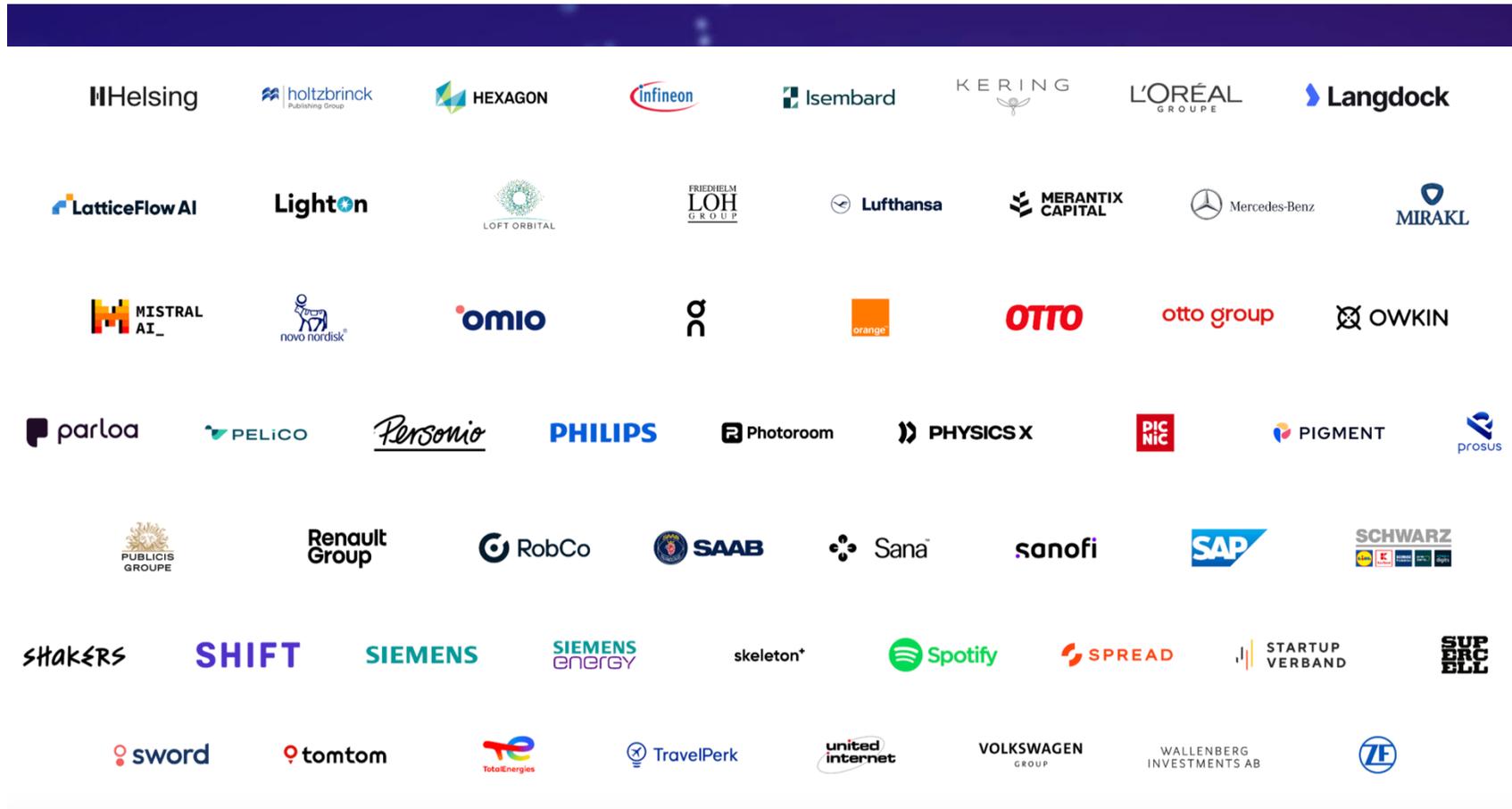


Mixture of Experts  
Open Source

Training costs

- DeepSeek: ~ \$6M
- OpenAI: ~ \$100M

# In Feb 2025



The current supporters of the EU AI Champions Initiative represent over **70+** organizations, **\$3+ trillion** in public market cap, and over **3.7M jobs** across Europe—and that list is growing. All participating companies are united in showing strong leadership in AI.

# And Corporate Spending...

Google will spend more than \$100 billion on AI, DeepMind CEO says

Apr 16, 2024 — Just last month, Google DeepMind CEO Demis Hassabis said the billions of dollars being poured into AI is...

QZ Quartz



Microsoft Slated to Invest \$80 Billion in AI-Enabled Data Centers ...

Jan 3, 2025 — Microsoft Slated to Invest \$80 Billion in AI-Enabled Data Centers This Fiscal Year. ... Sarina Trangle starte...

 Investopedia



Meta plans to invest up to \$65 billion in AI infrastructure in 2025, a significant increase from the \$38-40 billion spent in 2024, with the goal of expanding its AI capabilities and data centers. 

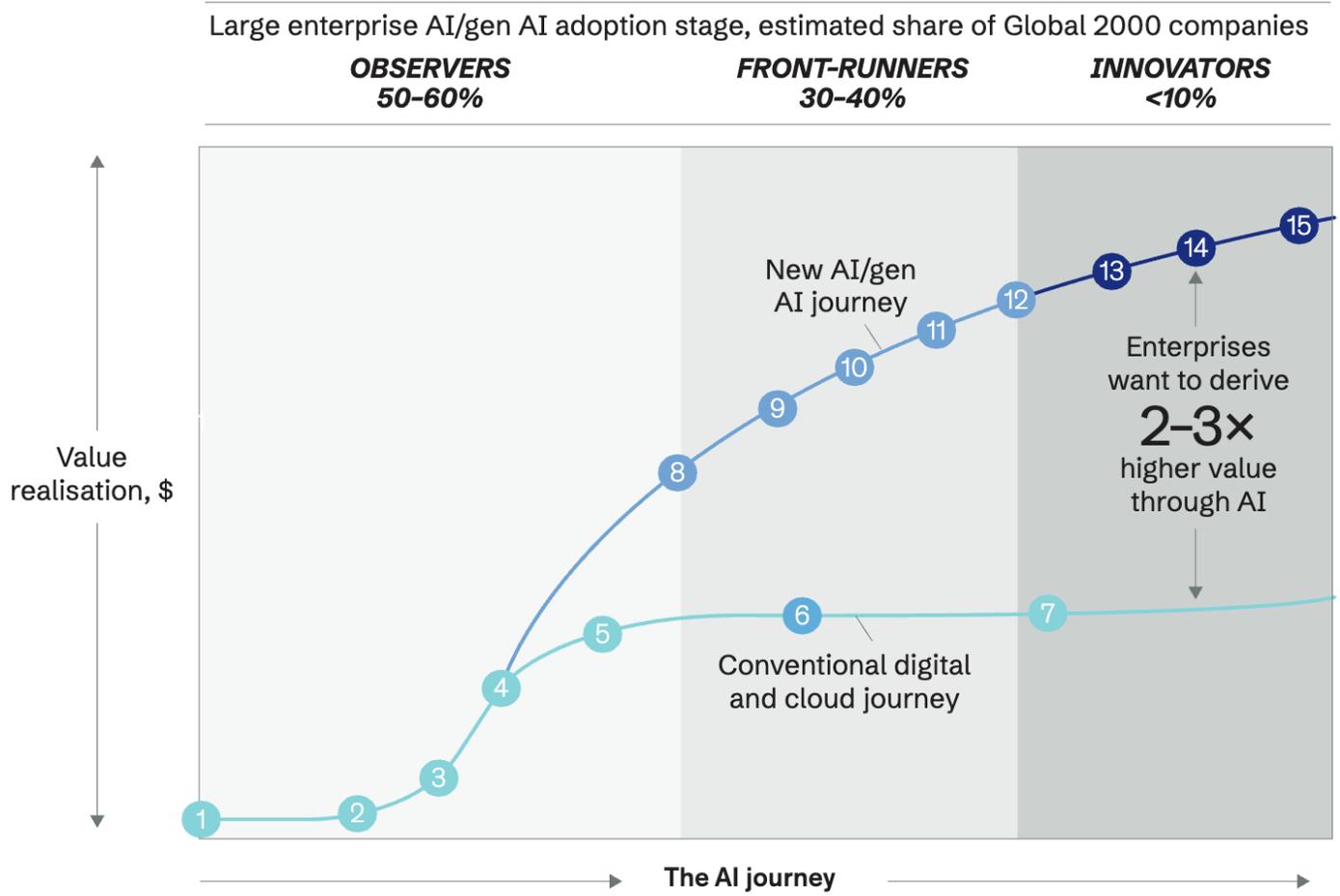
# AI Startups

As of January 2025,  
there were about  
**10,095** AI startups  
(Gemini)



- [Abstract AI](#) — 100% AI-handled customer support
- [ggml](#) — run AI models anywhere
- [Lindy](#) — AI personal assistant
- [Pika Labs](#) — cutting-edge generative video
- [Mathpix](#) — AI-powered document automation
- [Julius](#) — AI data scientist
- [Hume AI](#) — AI toolkit to understand emotional ex
- [Cofactory](#) — factory for generating companies witi
- [Induced](#) — AI-first browser RPA platform
- [Zeta Labs](#) — automate routine online tasks
- [Speakshyft](#) — real-time accent translation
- [Common Sense Machines](#) — game-engine ready 3
- [Guru AI](#) — easy-to-use video analysis models
- [Curio](#) — AI-powered toys
- [Echo Labs](#) — human-level transcription
- [Reality Defender](#) — deepfake detection
- [Andiron AI](#) — e-commerce optimization
- [Lightpaper](#) — AI assembly lines for knowledge wc
- [Portola](#) — AI-powered creative tools for kids
- [Espresso AI](#) — optimize Snowflake queries using I
- [Jenni](#) — AI workspace for researchers
- [AutogenAI](#) — generate bids proposals using LLMs
- [Merlin](#) — AI-led user interviews
- [Tutor Intelligence](#) — AI cobots
- [RunPod](#) — serverless GPU platform
- [Akool](#) — personalized visual marketing content
- [Coframe](#) — automated A/B testing
- [OpusClip](#) — AI video clipping tool
- [Freed](#) — AI medical scribe
- [Perplexity](#) — the fastest way to get an answer
- [Cursor](#) — AI-first code editor
- [Replicate](#) — cloud infrastructure for ML models
- [Animato](#) — video chat with AI characters
- [Lexica.art](#) — make AI art
- [Minion.ai](#) — automated browser assistant
- [Recraft](#) — generate vector art and 3D images
- [Flair](#) — AI design tool for branded content
- [ValueBase](#) — AI property valuation models for municipal governments
- [WOMBO](#) — magical consumer AI experiences
- [Chroma](#) — programmable memory for AI
- [Poly Corp](#) — AI-generated textures
- [Sieve](#) — AI video API
- [Sameday](#) — appointment scheduling AI
- [Play.ht](#) — AI voiceover for podcasts
- [Ghostwrite](#) — automatic email composer
- [BuildShip](#) — low-code visual backend builder
- [Birch](#) — automating complex call center operations in regulated industries
- [Vizcom](#) — AI-powered engineering drawings
- [Circle Labs](#) — generative AI discord friends
- [Samaya AI](#) — knowledge discovery platform for financial services
- [Secret Weapons](#) — AI video tools used by Hollywood
- [Pixelcut](#) — AI-powered product photos
- [AniML](#) — NeRF-generated product videos
- [Dust](#) — browser copilot for teams
- [Forefront](#) — enterprise chatbot

# Why?



# What is I-AI?

- No precise definition
- Use evolved very quickly to “Table Stakes”
  - Manufacturing
  - Genomics
  - Systems
  - Maintenance
- Technology
  - DNN’s everywhere
  - Mixed with traditional methods



# Enter your prompt:

## List major applications of Generative AI in industry

- **Llama Response:**

- Deep Neural Networks (DNNs) have transformed numerous fields by providing powerful tools for automating complex tasks. Here are some major applications of DNN-based AI:

- 1. **Image and Video Processing**:

- - **Image Classification**: Identifying objects in images, used in systems like Google Photos and Instagram.
- - **Object Detection**: Locating objects within an image or video, utilized in autonomous vehicles and augmented reality.
- - **Facial Recognition**: Used in security systems and for unlocking devices.
- - **Image Segmentation**: Dividing an image into segments for tasks in medical imaging and autonomous vehicles.
- - **Style Transfer and Generation**: Creating artistic images, as seen in apps like Prisma.

- 2. **Natural Language Processing (NLP)**:

- - **Text Classification**: Spam detection, sentiment analysis, and topic categorization.
- - **Machine Translation**: Automatic translation between languages, used in tools like Google Translate.
- - **Speech Recognition**: Converting spoken language into text, used in virtual assistants like Siri and Alexa.
- - **Chatbots**: Automated systems for customer service and engagements.
- - **Natural Language Generation**: Producing human-like text, seen in applications like OpenAI's GPT models.

- 3. **Audio Processing**:

- - **Speech Synthesis**: Text-to-speech applications, providing natural-sounding audio outputs.
- - **Music Generation**: Creating music using AI models.
- - **Audio Classification**: Recognizing audio samples, such as identifying songs and detecting sounds in environments.

# TLDR

# I-AI Application Segments



The  
Economist

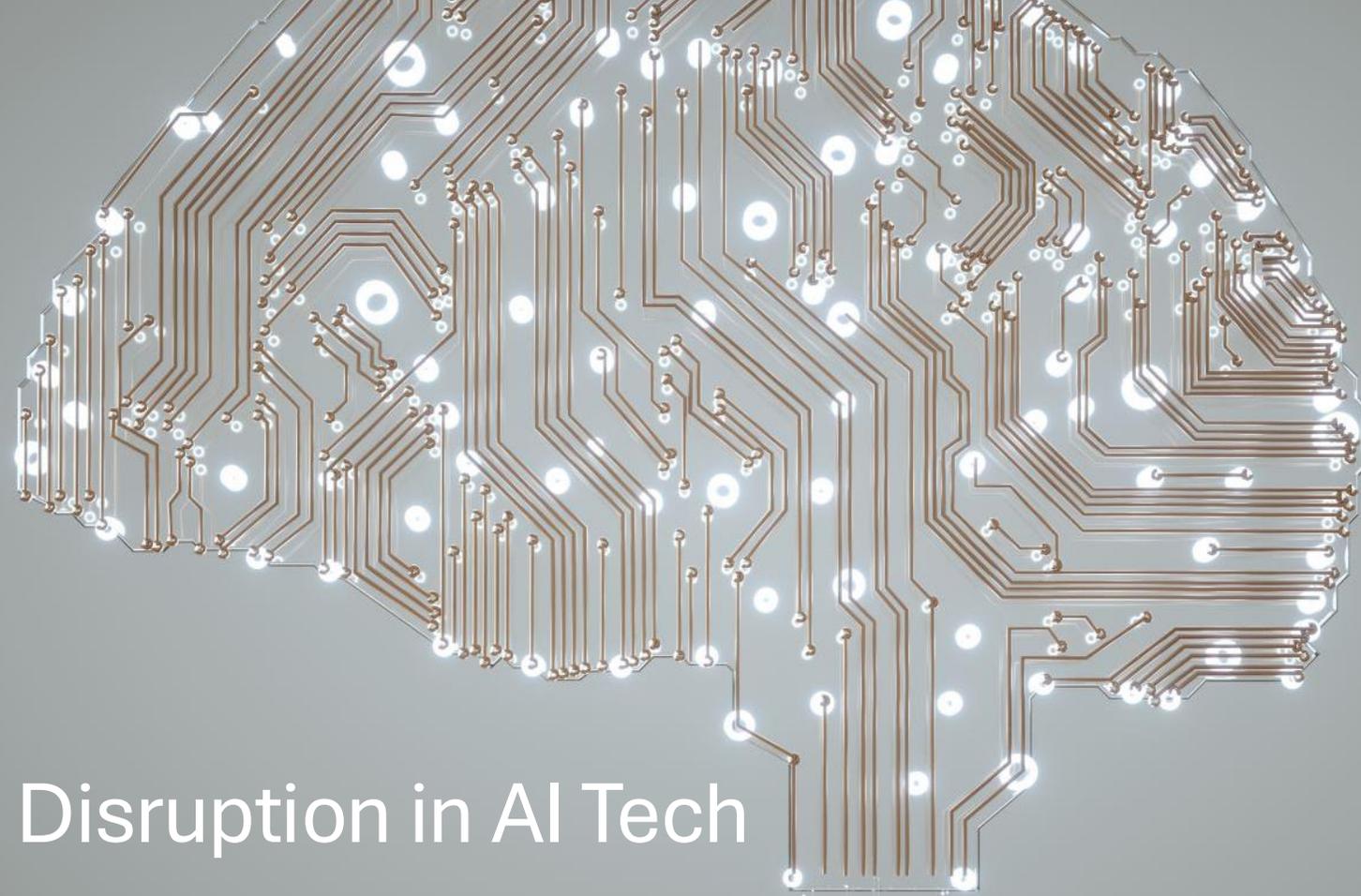
≡ Menu

Business | Meet your new copilot

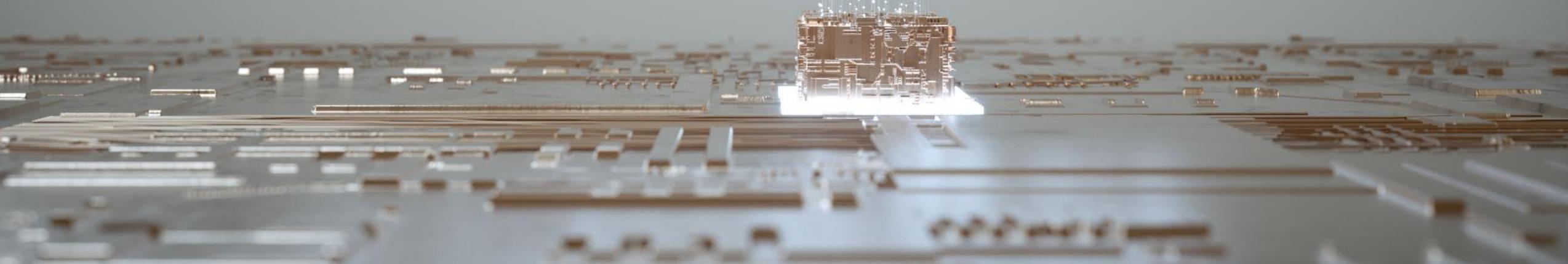
Speakers

# AI in operations is not (just) chat...

- AI is not a fancy control system
    - It doesn't understand RISK
    - Unpredictable
    - Doesn't know if it is wrong
  - Brilliant at finding unmodeled patterns in complex data (Microsoft)
- 
- Best used when
    - Downside can be managed.
    - Integrated with other tools (KLA)
    - Glue to integrate different tools
  - Great for multi-modal applications
  - Great for human interface



# Disruption in AI Tech



# Skip AI Tech Review: Types of Models

## • Generative models

- Models of the distribution of data:  $P(X | \alpha)$ 
  - Unlabeled data  $X$
  - Find the best  $\alpha$  to fit data
- Examples
  - LLM e.g. ChatGPT, Dall E
  - VAE

## • Discriminative models

- Predict a value from a set of inputs:  $y = F(X | \beta)$ 
  - Labeled data
    - Input  $X$
    - Output  $y$
  - Find the best  $\beta$  to fit data

# AI Tech Review

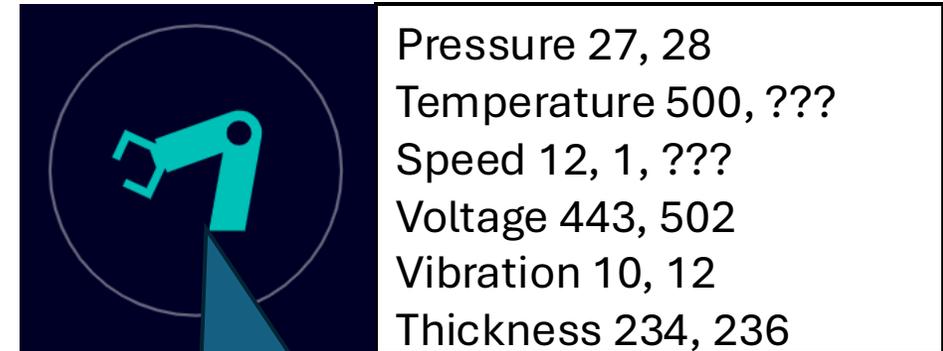
Guess ???

Enter your prompt: find the next word in the following: The cat in the

GPT-4 Response:

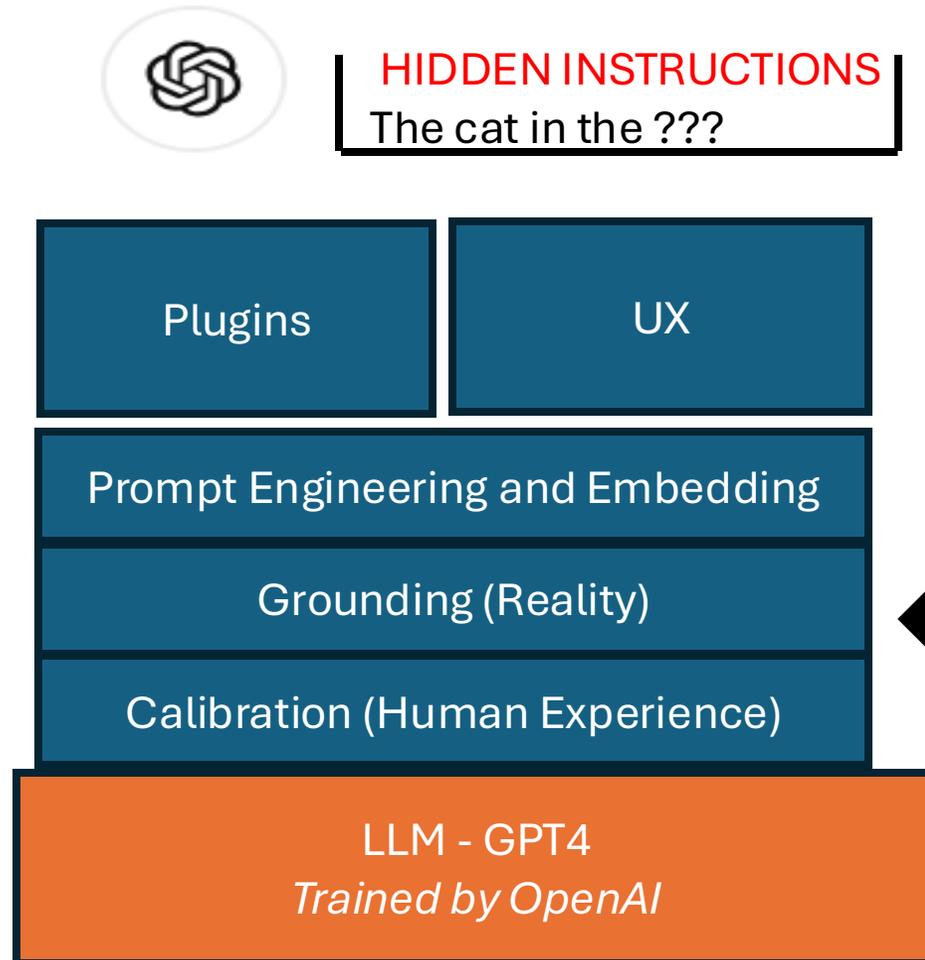
The phrase "The cat in the" is most commonly completed with the word "hat," as popularized by Dr. Seuss's children's book "The Cat in the Hat." This is a well-known book that many people recognize, which makes "hat" a likely candidate for the next word.

- **Industrial applications more complex**
  - **Generative model => Encoder & Decoder**
    - Telemetry is not text
    - Probabilistic trajectories  $p(z_1, \dots, z_N)$



Predict  
Anomaly?

# CoPilots and chatBots



- CoPilots/Agents
- Specialized
- Many platforms (open source)

## Tools:

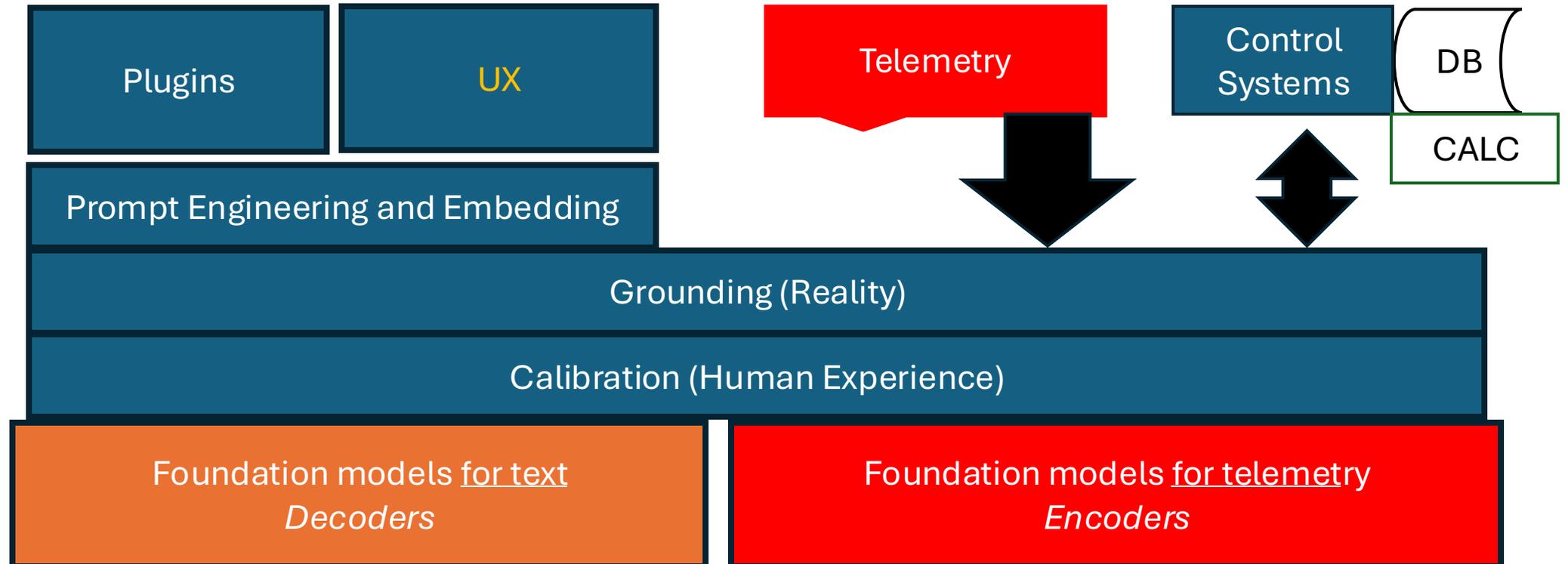


Retrieval-Augmented Generation (RAG)  
Commonly used to embed corporate or service information

chatBot

# Industrial Agents

AI Agent



# Flood of Papers. Hugging Face

## Daily Papers

by  AK and the research community

🔍 Search any paper with AI

Daily Weekly **Monthly**

◀ March ▶

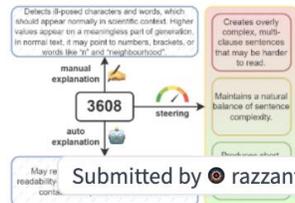
### Feature-Level Insights into Artificial Text Detection with Sparse Autoencoders

Kristian Kuznetsov<sup>1,2</sup>, Laida Kushnareva<sup>2</sup>, Polina Druzhinina<sup>1,5</sup>, Anton Razzhigaev<sup>1,5</sup>, Anastasia Voznyuk<sup>3</sup>, Irina Piontkovskaya<sup>2</sup>, Evgeny Burnaev<sup>1,5</sup>, Serguei Barannikov<sup>1,4</sup>,

<sup>1</sup>Skolkovo Institute of Science and Technology, <sup>2</sup>AI Foundation and Algorithm Lab  
<sup>3</sup>Moscow Institute of Physics and Technology, <sup>4</sup>CNRS, Université Paris Cité, France  
<sup>5</sup>Artificial Intelligence Research Institute (AIRI)

#### Abstract

Artificial Text Detection (ATD) is becoming increasingly important with the rise of advanced Large Language Models (LLMs). Despite numerous efforts, no single algorithm performs consistently well across different types of unseen text or guarantees effective generalization to new LLMs. Interpretability plays a crucial role in achieving this goal. In this study, we enhance ATD interpretability by using Sparse Autoencoders (SAE) to extract features from Gemma-2-2b's residual stream. We identify



▲  
212

### Feature-Level Insights into Artificial Text Detection with Sparse Autoencoders

 · 8 authors

🗨️ 2

### Transformers without Normalization

Jiachen Zhu<sup>1,2</sup>, Xinlei Chen<sup>1</sup>, Kaiming He<sup>3</sup>, Yann LeCun<sup>1,2</sup>, Zhuang Liu<sup>1,4,†</sup>

<sup>1</sup>FAIR, Meta, <sup>2</sup>New York University, <sup>3</sup>MIT, <sup>4</sup>Princeton University  
<sup>†</sup>Project lead

Normalization layers are ubiquitous in modern neural networks and have long been considered essential. This work demonstrates that Transformers without normalization can achieve the same or better performance using a remarkably simple technique. We introduce Dynamic Tanh (DyT), an element-wise operation  $DyT(x) = \tanh(\alpha x)$ , as a drop-in replacement for normalization layers in Transformers. DyT is inspired by the observation that layer normalization in Transformers often produces tanh-like, S-shaped input-output mappings. By incorporating DyT, Transformers without normalization can match or exceed the performance of their normalized counterparts, mostly without hyperparameter tuning. We validate the effectiveness of Transformers with DyT across diverse settings, ranging from recognition to generation, supervised to self-supervised learning, and computer vision to language models. These findings challenge the conventional understanding that normalization layers are indispensable in modern neural networks, and offer new insights into their role in deep networks.

Date: March 14, 2025  
Project page and code: [jiachenzhu.github.io/DyT](https://jiachenzhu.github.io/DyT)  
Correspondence: [jiachen.zhu@nyu.edu](mailto:jiachen.zhu@nyu.edu), [zhuangl@princeton.edu](mailto:zhuangl@princeton.edu)

Meta

Submitted by  akhaliq

▲  
128

### Transformers without Normalization

 · 5 authors

🗨️ 3

### RuCCoD: Towards Automated ICD Coding in Russian

Aleksandr Nesterov<sup>1</sup>, Andrey Sakhovskiy<sup>2</sup>, Ivan Sviridov<sup>3</sup>, Airat Valiev<sup>4</sup>, Vladimir Makharev<sup>1</sup>, Petr Anokhin<sup>1</sup>, Galina Zubkova<sup>1</sup>, Elena Tutubalina<sup>1,2,5</sup>

<sup>1</sup> AIRI, Moscow, Russia  
<sup>2</sup> Sber AI, Moscow, Russia  
<sup>3</sup> Sber AI Lab, Moscow, Russia  
<sup>4</sup> HSE University, Moscow, Russia  
<sup>5</sup> ISP RAS Research Center for Trusted Artificial Intelligence, Moscow, Russia

#### Abstract

This study investigates the feasibility of automating clinical coding in Russian, a language with limited biomedical resources. We present a new dataset for ICD coding, which includes diagnosis fields from electronic health records (EHRs) annotated with over 10,000 entities and more than 1,500 unique ICD codes. This dataset serves as a benchmark for several state-of-the-art models, including BERT, LLaMA with LoRA, and RAG, with additional experiments examining transfer learning across do-



▲  
122

### RuCCoD: Towards Automated ICD Coding in Russian

 · 8 authors

🗨️ 2

# Issues and Tech Solutions

## Issues

- Hallucinations – what and why!
- Honesty – learn to say “I got nothin”
- Grounding – not connected to reality
- Matching human expectations
- Instruction following. Not logical
- Performance
- Cost in a later section

## Response

- Co-Pilots
- Reinforcement Learning with Human Feedback
- Chain of Thought
- Mixture of Experts

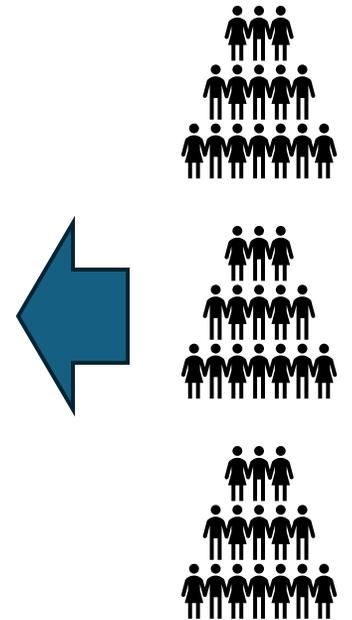
# Skip: Meeting Human Expectation: RLHF

**HIDDEN INSTRUCTION**  
The cat in the ???

LLM - GPT4  
*Re-Trained by OpenAI*

car: No!  
↔  
hat: Yes

Reinforcement  
Learning with  
Human Feedback



- A kind of average over the population
- Lots of data required or a Rewards Model
  - Simplest is Bradley Terry
- Skews the LLM probability model

# Instruction Following: Chain of Thought: CoT

## HIDDEN INSTRUCTIONS

Q: Do cats like to hide?

A: Yes, especially in warm places

Q: Can that be clothing?

A: Yes, coats, hats, and other things

Q: Do children's books show this?

A: Yes, Dr. Seuss is an example

Q: Complete "the cat in the ???"

A: hat

Can be done  
automatically  
and internally to  
AI engine!

LLM - GPT4  
*Trained by OpenAI*

# Mixture of Experts: MoE

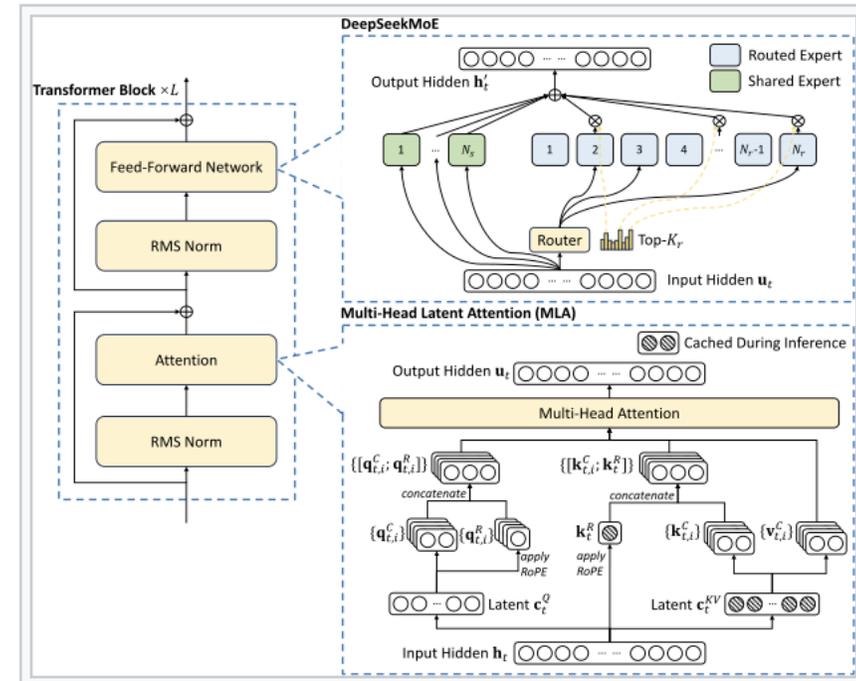
- Analogous to an internal research department
- Learn a set (100's) of special experts and query them as needed
- Let the GenAI model automatically select the experts



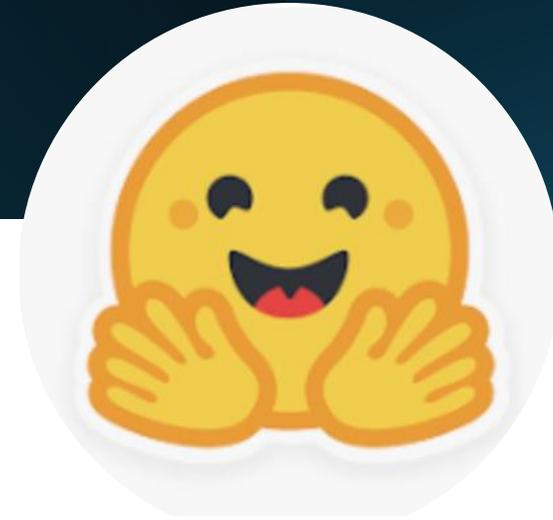
*Dado Ruvic | Reuters*

# Mixture of Experts: MoE

- Analogous to an internal research department
- Learn a set (100's) of special experts and query them as needed
- Let the GenAI model automatically create and select the experts



# Open Source Building Blocks



- Open source engines
- HuggingFace
- GitHub
- As of January 2025, there were about **10,095 AI startups** (Gemini)

# Open Source GenAI Engines

- [Open Source Leader Board](#)



## Open LLM Leaderboard

Comparing Large Language Models in an open and reproducible way

🔍 Search by model name • try "meta @architecture:llama @license:mit" 4275 / 4275 ⌵ Advanced Filters ⓘ

Supports strict search and regex • Use semicolons for multiple terms

# Open Source Sentiment

Llama has now been downloaded over **1 Billion times!**

A note to:

The **researchers** at Meta training these models — and those building on the research in other labs.

The **developers and enthusiasts** on r/LocalLlama, [@huggingface](#) and more; experimenting with new models and creating derivatives.

The **small startups and big enterprises** alike who are creating a new wave of AI-powered products, built with Llama.

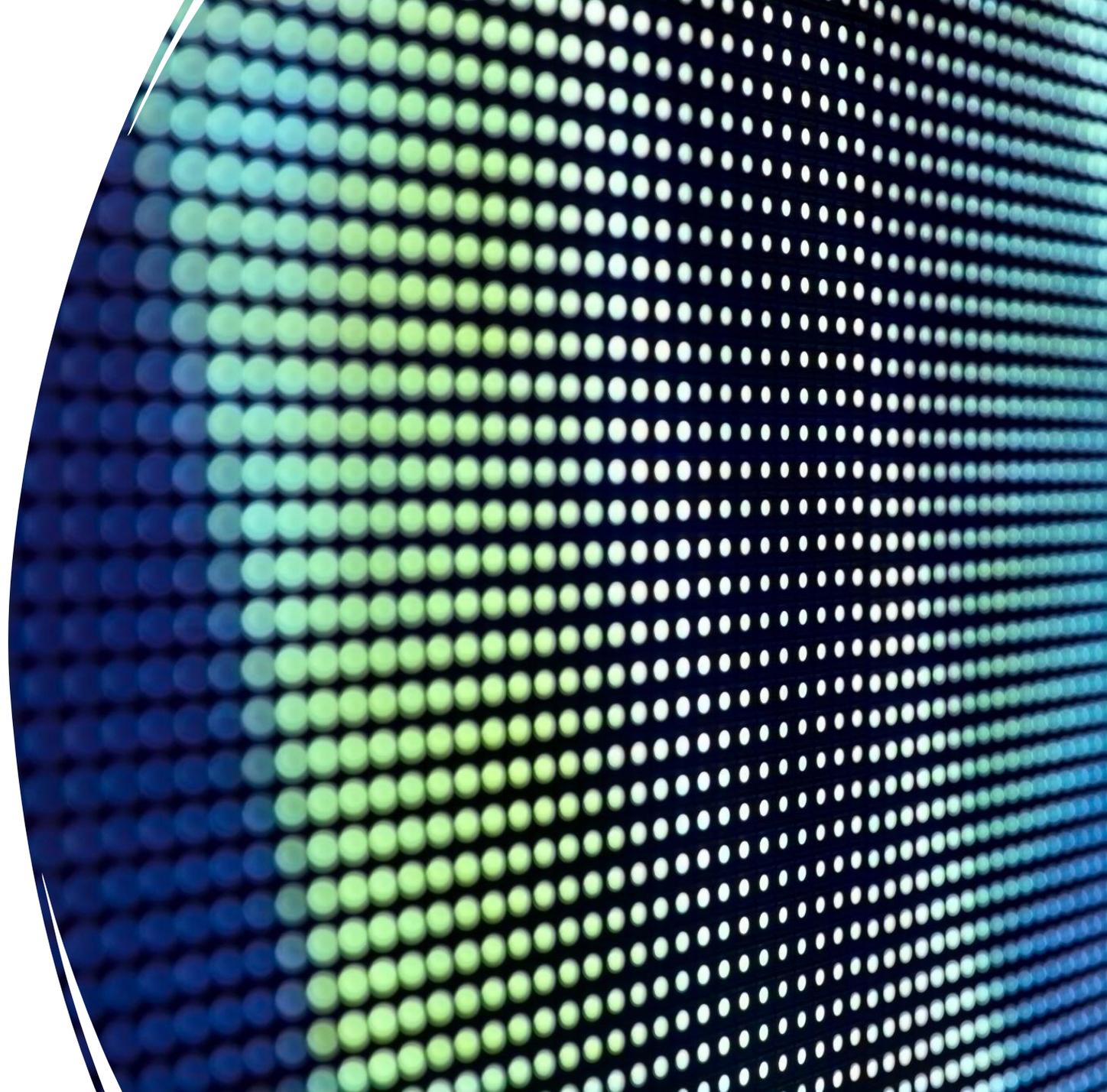
**The global AI community.**

Your actions speak louder than words, thank you for making it abundantly clear — a billion times over — that open source AI is how we'll create the next wave of world changing technologies, together. 🐪 ❤️

# Speaker Applications

---

- AIOps - Microsoft
- DT and semiconductor manufacturing
  - KLA
  - Aitomatic
- Supply chain – Walmart
- Autonomy - Northrup Grumman
- AI cost - Payi



# Overall

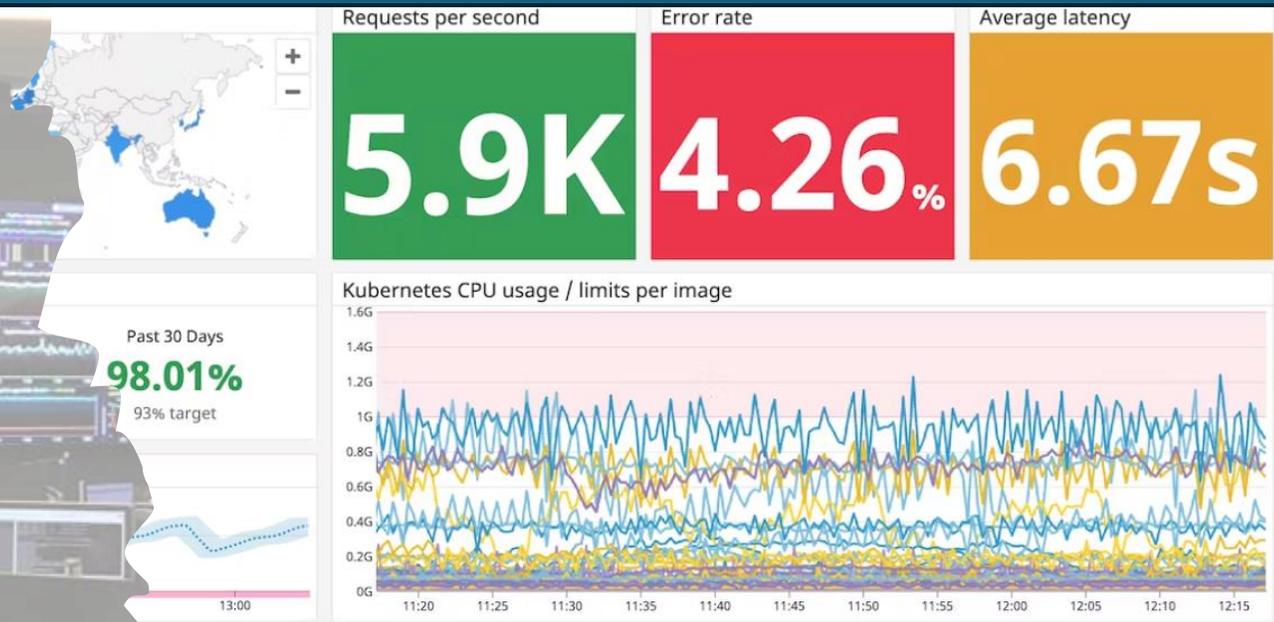
- Ubiquity
- Lego
- Integration
- Cost management
- Autonomy



# Data Center is the Factory of the 21<sup>st</sup> century

# AIOps

# Managing Large Scale Computer Systems



Clone Dashboard

|           |            |             |            |              |
|-----------|------------|-------------|------------|--------------|
| Clusters  | Namespaces | Services    | Pods       | Kubelets up  |
| 16        | 43         | 740         | 3635       | 172          |
| Nodes     | DaemonSets | Deployments | Containers | Kubelet Ping |
| 184 nodes | 48         | 452         | 5070       | 164          |

Events

Running pods per namespace

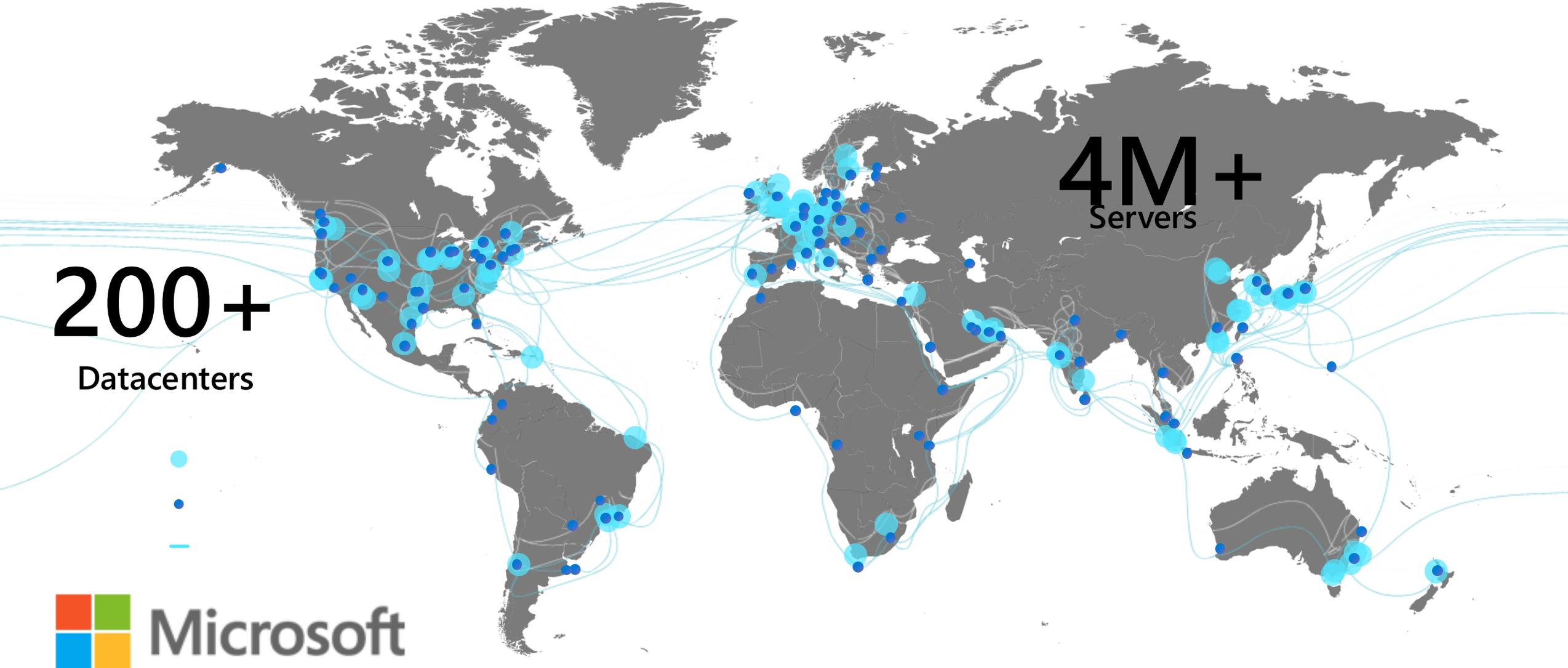
Pods running by namespace

Pods in ready state by node

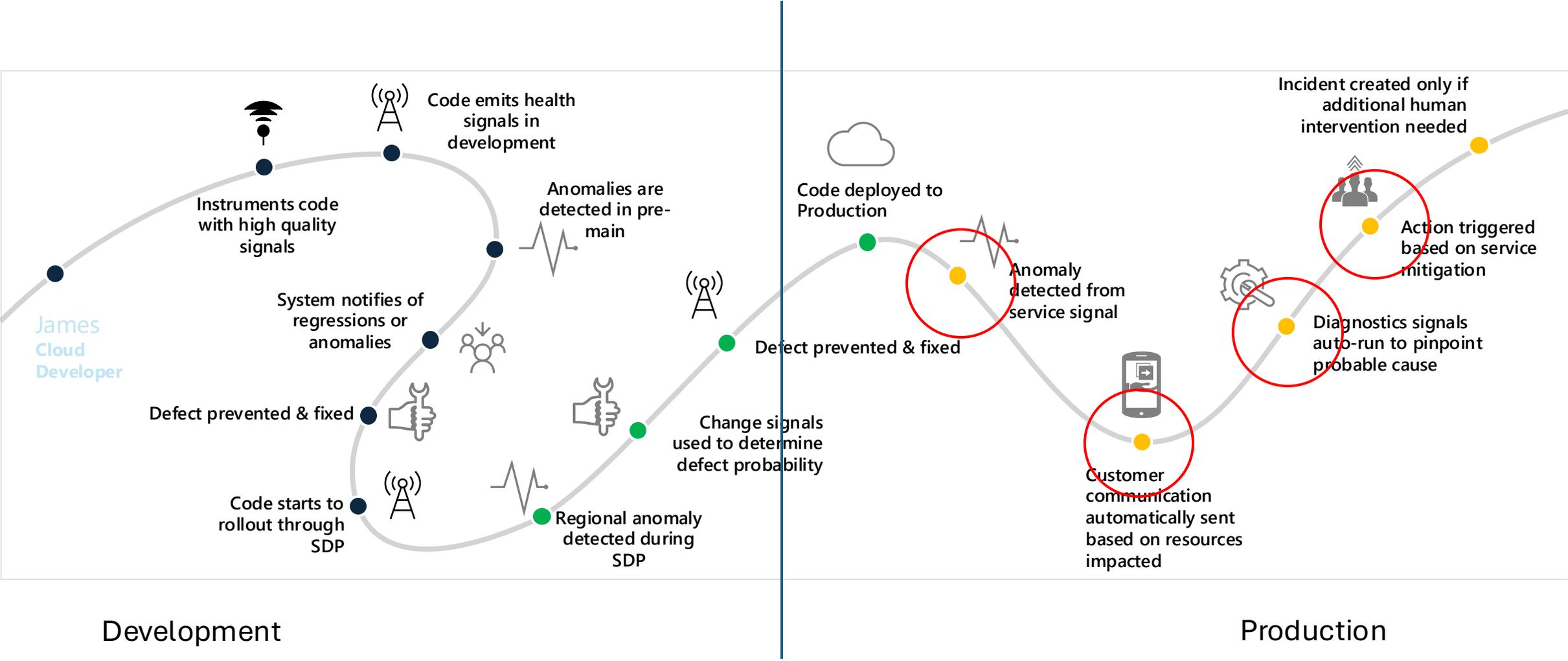
Service: web-store x ERROR x

| DATE                | HOST  |
|---------------------|---|
| Dec 17 15:16:27.646 | gke-demo-11287-us-prod-e-default-pool-4180c160-b... |
| Dec 17 15:16:26.953 | 1-e0e7142ed027a9996                                 |
| Dec 17 15:16:26.488 | 1-0915828a295f4e873                                 |
| Dec 17 15:16:26.451 | gke-demo-11287-us-prod-west-pool-2-2b19f4ee-xygy... |
| Dec 17 15:16:26.448 | gke-demo-11287-us-prod-west-pool-2-2b19f4ee-xygy... |
| Dec 17 15:16:26.376 | 1-0e57866e977b5f467                                 |
| Dec 17 15:16:26.168 | vm-48732de3-deb-4ab4-7b71-f478f876976.c.datadog...  |
| Dec 17 15:16:25.828 | gke-demo-11287-us-prod-e-default-pool-59bb8b03-p... |
| Dec 17 15:16:26.027 | gke-demo-11287-us-prod-e-default-pool-59bb8b03-p... |
| Dec 17 15:16:25.848 | gke-demo-11287-us-prod-e-default-pool-59bb8b03-p... |
| Dec 17 15:16:25.797 | gke-demo-11287-us-prod-e-default-pool-59bb8b03-p... |
| Dec 17 15:16:25.649 | gke-demo-11287-us-prod-west-pool-2-ac11bc48-uyk6... |
| Dec 17 15:16:25.553 | 1-0f4f1a0ccc998021                                  |
| Dec 17 15:16:25.553 | gke-demo-11287-us-prod-e-default-pool-4180c160-1... |
| Dec 17 15:16:25.552 | 1-072c0868ce156c179                                 |
| Dec 17 15:16:25.344 | gke-demo-dpn-us-west-default-pool-438ab055-1nsf...  |
| Dec 17 15:16:25.243 | 1-0e57866e977b5f467                                 |
| Dec 17 15:16:25.242 | 1-0e57866e977b5f467                                 |
| Dec 17 15:16:25.182 | gke-demo-11287-us-prod-e-default-pool-4180c160-b... |
| Dec 17 15:16:25.182 | gke-demo-11287-us-prod-e-default-pool-4180c160-b... |
| Dec 17 15:16:25.128 | gke-demo-11287-us-prod-e-default-pool-59bb8b03-p... |
| Dec 17 15:16:25.128 | gke-demo-11287-us-prod-e-default-pool-59bb8b03-p... |
| Dec 17 15:16:24.964 | 1-0a320338826c12                                    |
| Dec 17 15:16:24.929 | gke-demo-dpn-us-west-default-pool-1bd696c-sodd...   |
| Dec 17 15:16:24.766 | gke-demo-11287-us-prod-west-pool-2-2b19f4ee-xy74... |

# Data Center is the Factory of the 21<sup>st</sup> century



# AI Ops: Monitoring and Prediction



# Digital Twins in Semiconductor Industry

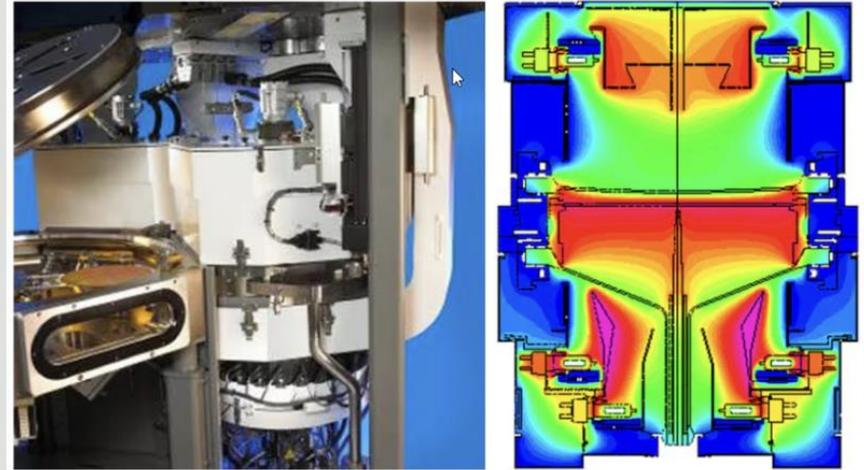
- Digital Twins
- Co-optimization
- ~~Predictive maintenance~~
- Predictive management



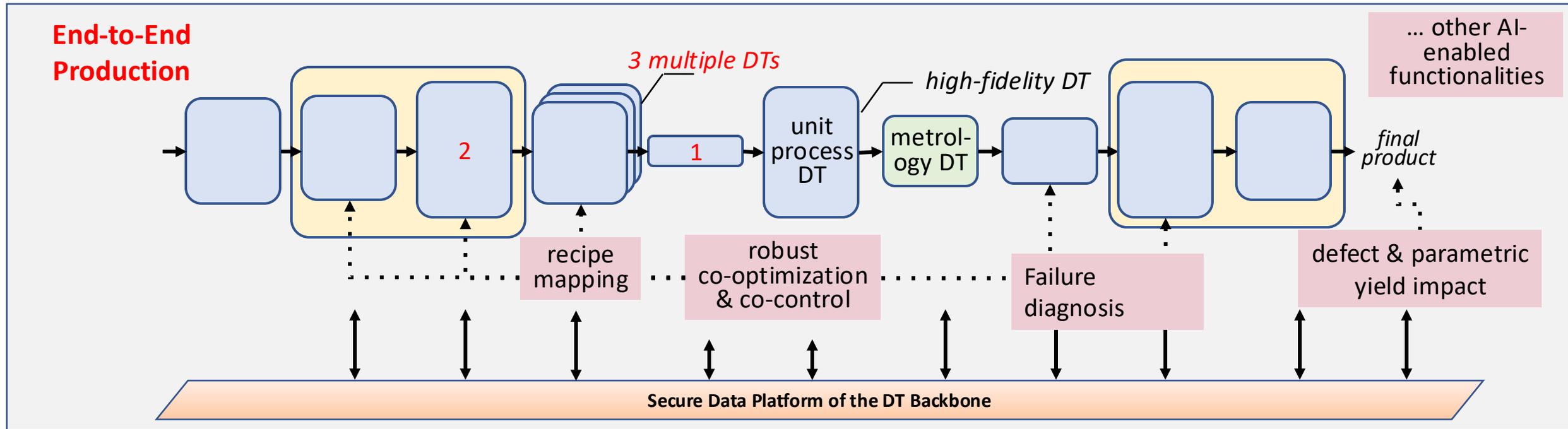
# Digital Twin

- A software twin of an object
  - Physical
  - Logical
- Mix
  - AI Encoder/Decoder
  - Simulation software
- At a specified level of fidelity
- Integrated with physical device

Why mix different models?

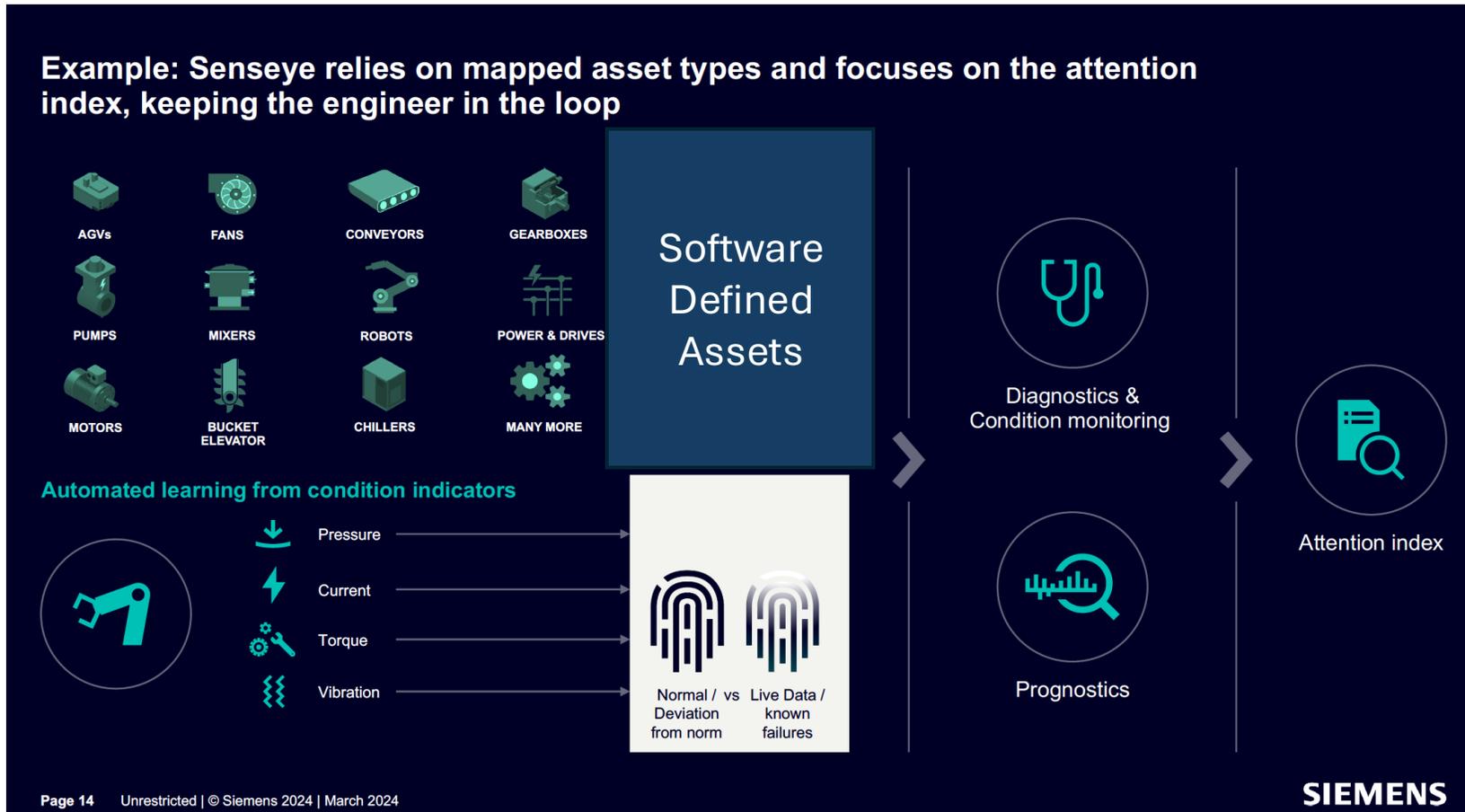


# Digital Twins in Semiconductor Industry



Adapt to what came before  
Co-optimize by looking across the system

# DT's of many types



# Traditional Supply Chain Problem



# AI as Glue: Integration (Translation)

- Translating and connecting different applications
- Anomalies
- Patterns



# Agentic (Multi-agent)

Local Agent 1  
Warehouse

Agent 2  
Trucking

Human  
Supervisor



Local Agent 4  
Retail

# Agents for Everything

## AI Agents

Take action with autonomous AI agents that work for you.

## Customer Service Management

Empower self-service, boost agent productivity, and speed up resolution.

## HR Service Delivery

Give employees instant answers, guidance, and fast issue resolution.

## IT Asset Management

Improve technology use and spend over the IT asset lifecycle.

## Security Operations

Defend against security threats and attacks.

## IT Service Management

Transform service management for productivity and ROI.

## IT Operations Management

Deliver proactive digital operations with AIOps.

## Strategic Portfolio Management

Gain insights to move from strategy to business outcomes.

## Integration Hub

Connect anything to the ServiceNow platform.

## Field Service Management

Reduce field service costs and improve efficiency.

# Cost

- This a hardware issue being addressed in several ways
  - Federated computing
  - Faster chips
  - Student – Teacher AI models



# Costs

- Cost of proprietary engines

## Artificial Analysis LLM Performance Leaderboard

Independent performance benchmarks & pricing across API providers of LLMs. Definitions are below the table.  
For further analysis and [methodology](#), see [artificialanalysis.ai](#).

Filter, e.g. GPT, Meta

EXPAND COLUMNS I→

| API PROVIDER ↑↓   | MODEL ↑↓   | FEATURES I→       | MODEL INTELLIGENCE I→                     | PRICE I→                 | OUTPUT TOKENS/S I→ | LATENCY I→                |
|---|--|-------------------|---|--------------------------|--------------------|---------------------------|
|   |  | CONTEXT WINDOW ↑↓ | ARTIFICIAL ANALYSIS INTELLIGENCE INDEX ↑↓ | BLENDED USD/1M Tokens ↑↓ | MEDIAN Tokens/s ↑↓ | MEDIAN First Chunk (s) ↑↓ |
|  Microsoft Azure   |  o3-mini (high)     | 200k              | 66  | \$1.93                   | 108.7              | 69.33                     |
|  OpenAI            |  o3-mini            | 200k              | 63  | \$1.93                   | 183.2              | 12.57                     |
|  Microsoft Azure   |  o3-mini            | 200k              | 63  | \$1.93                   | 122.6              | 22.14                     |
|  Microsoft Azure |  o1               | 200k              | 62  | \$26.25                  | 95.8               | 32.04                     |
|  deepseek        |  DeepSeek R1      | 64k               | 60  | \$0.96                   | 25.1               | 67.28                     |
|  Hyperbolic      |  DeepSeek R1      | 128k              | 60  | \$2.00                   | 64.7               | 26.19                     |
| <b>NEBIUS</b>   |  DeepSeek R1 Base | 128k              | 60  | \$1.20                   | 11.7               | 93.38                     |

## DeepSeek-R1 API

Input API Price :

cache hit

**\$0.14** / 1M tokens

cache miss

**\$0.55** / 1M tokens

Output API Price :

**\$2.19** / 1M tokens

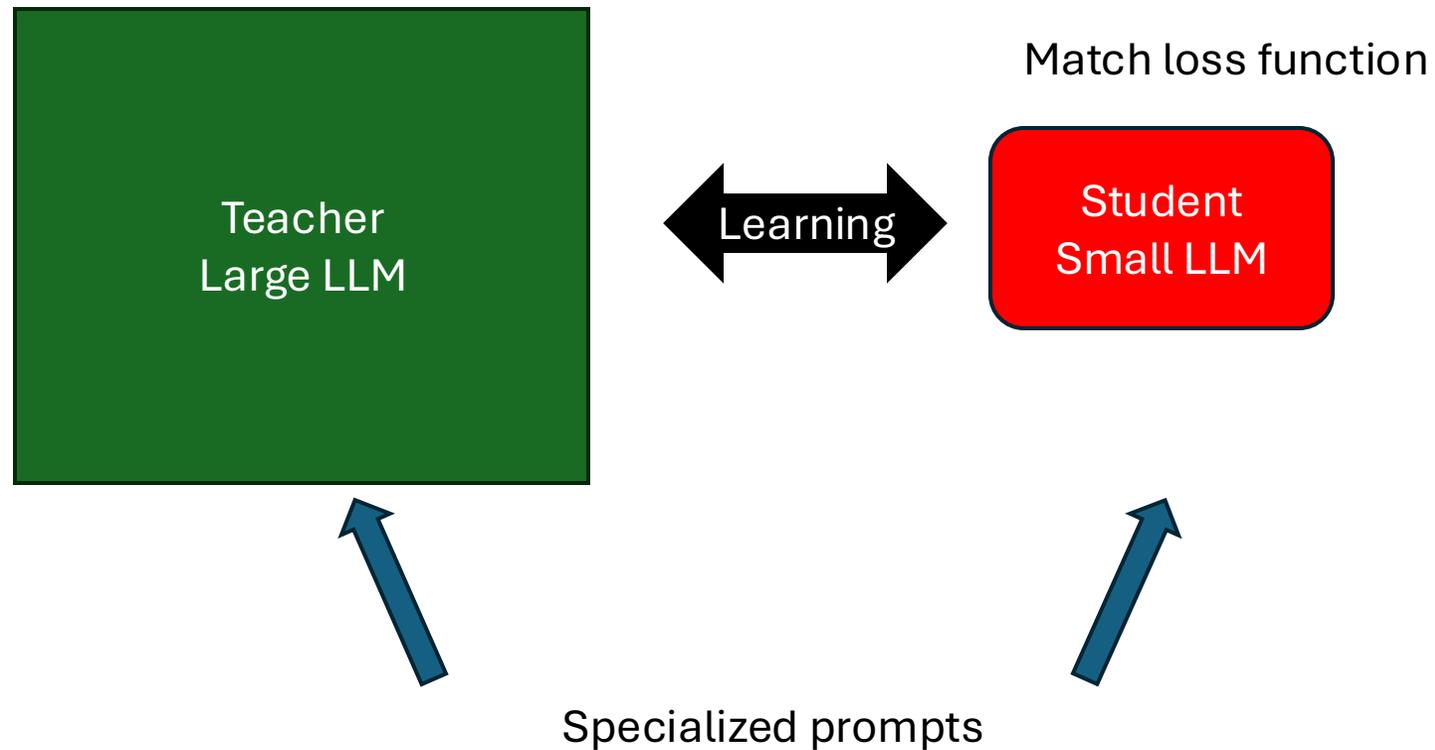


platform.deepseek.com

# Types of Costs and Trends



# Distillation: Reducing Inference Cost



# Swarms...



**Autonomous  
Systems**



# I-AI for Genomics (Not presenting this year)

- Blood tests for
  - Cancer - Anomaly detection
  - Genetic disease - Patterns
  - Epigenetic problems – “predictive maintenance”
- Drug discovery
  - A separate talk!

Primary structure



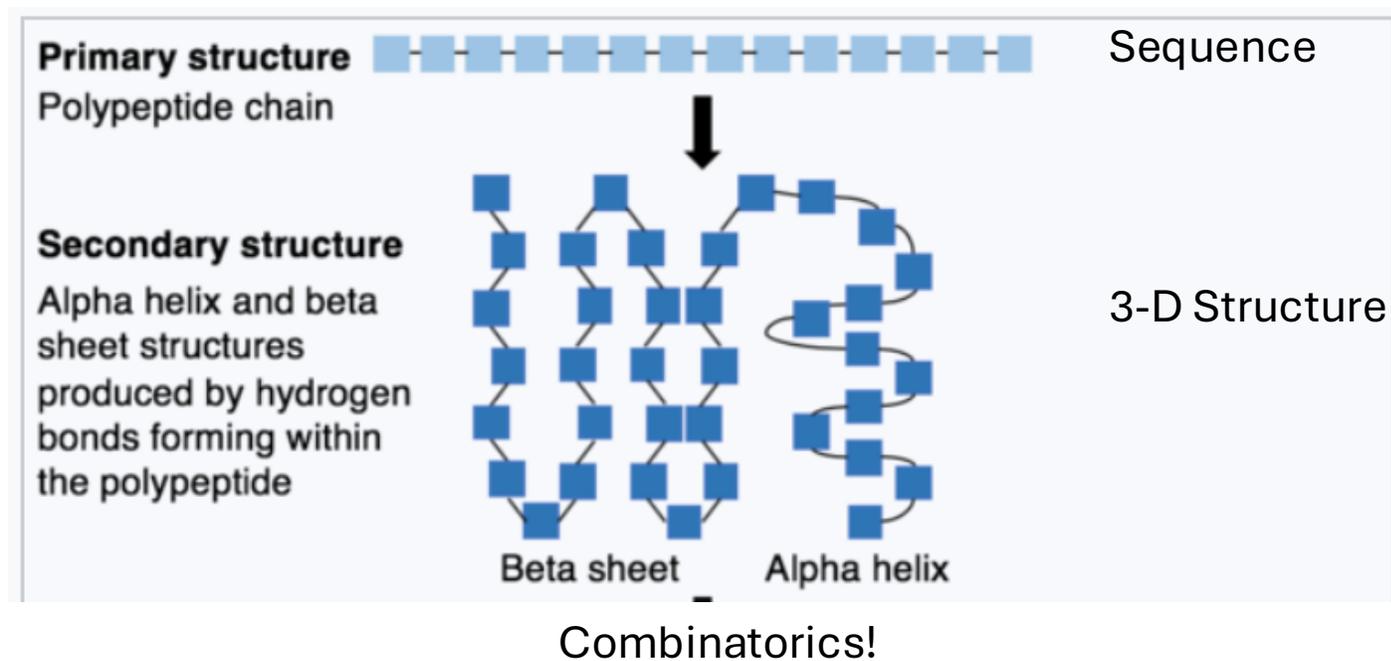
Sequence of nucleotides / amino acids

If we think of biology as fundamentally an information processing system – one that transmits information and maintains structure – we can start to see how it might share a basic underlying structure, or an ‘isomorphic mapping’, to information science.

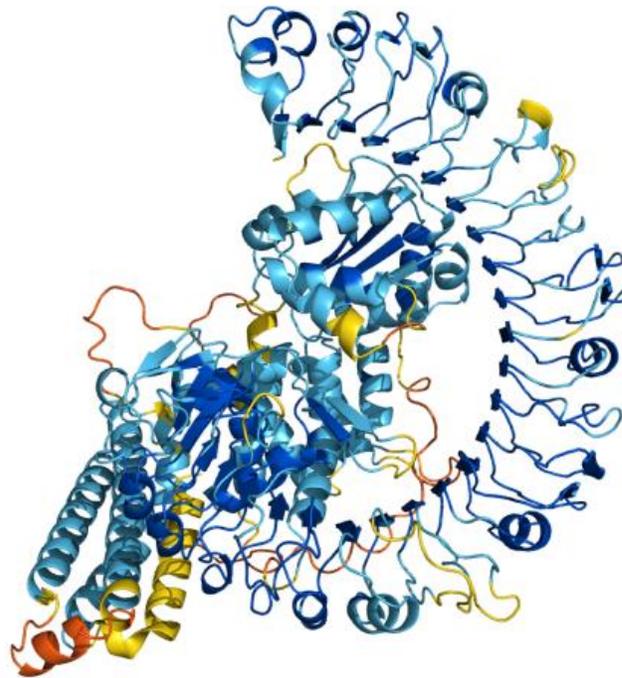
Isomorphic Labs (Google)

# AlphaFold 2

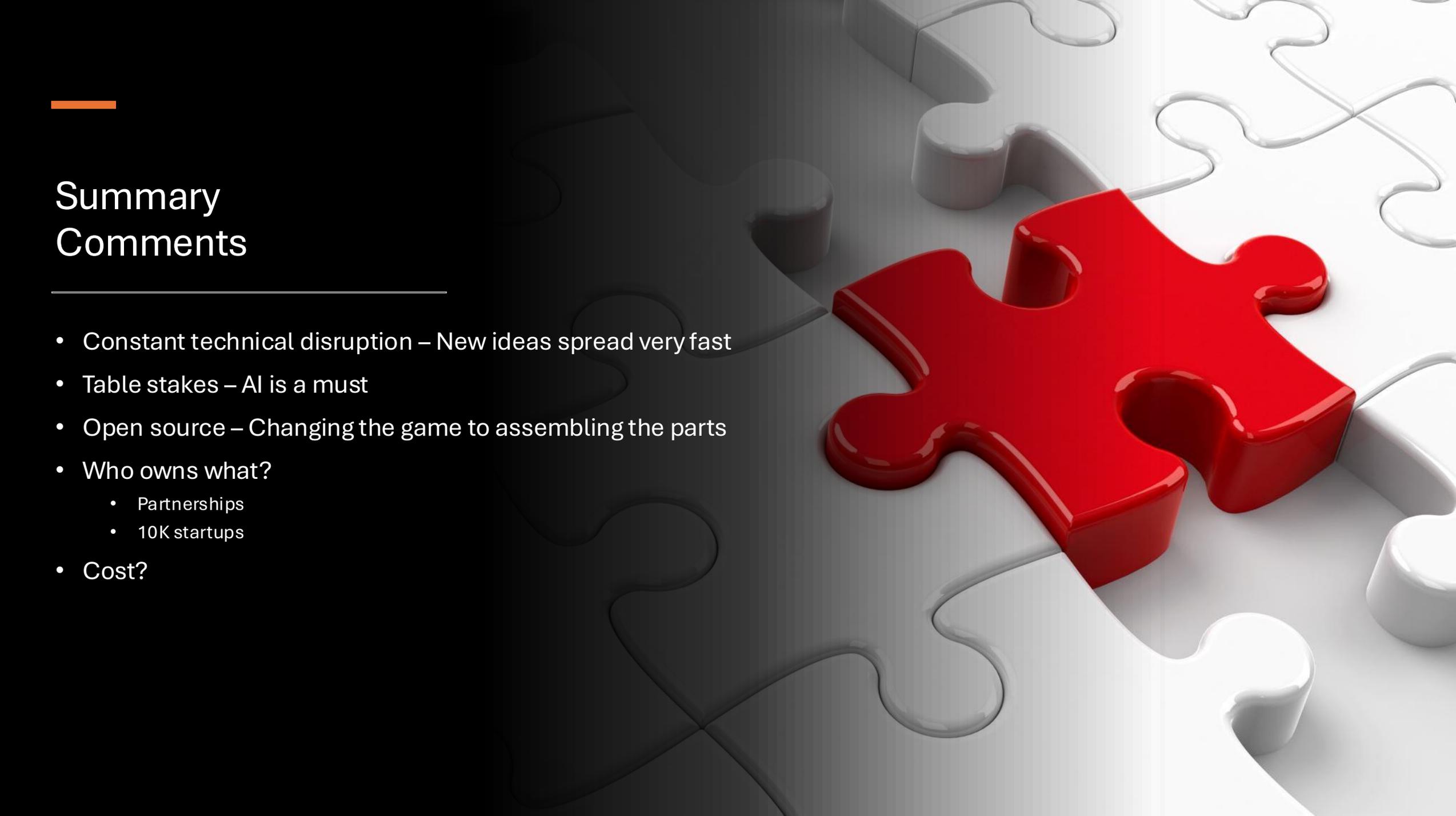
AlphaFold is an AI system developed by DeepMind (Google) that predicts a protein's 3D structure from its amino acid sequence.



CoPilot  
Agent



Note: For energy reasons  
there can exist several  
different configurations



# Summary

## Comments

---

- Constant technical disruption – New ideas spread very fast
- Table stakes – AI is a must
- Open source – Changing the game to assembling the parts
- Who owns what?
  - Partnerships
  - 10K startups
- Cost?

# Nick Landolfi, PhD

---

- Nick is a Vice President at **Unusual Ventures**. He received a B.S. in Electrical Engineering and Computer Science with honors from the University of California, Berkeley where he was a Regents' and Chancellor's Scholar. He received both an M.S. and Ph.D. in Computer Science from Stanford University, where he was a Stanford Graduate Fellow. A Silicon Valley native, Nick started a company in high school and has worked for and advised various early-stage companies over the past decade.

