Making AI Work for All

Leslie Teo, Senior Director AI Products, AI Singapore

Singapore Training Institute of the IMF

3 December 2025

The Core Question

Will we manage AI's challenges better than we managed globalization?

The Central Thesis: The fundamental challenge we face is not the technology itself, but rather the adoption architecture required to convert disruption into shared prosperity. Will we repeat the mistakes of the past, or will we write a new chapter in how societies manage technological transformation?

Agenda

The Landscape: Why This Time Is Different

Understanding AI's speed, cost, and agency.

Seeing is Believing: AI in Action

Practical demonstrations of current capabilities.

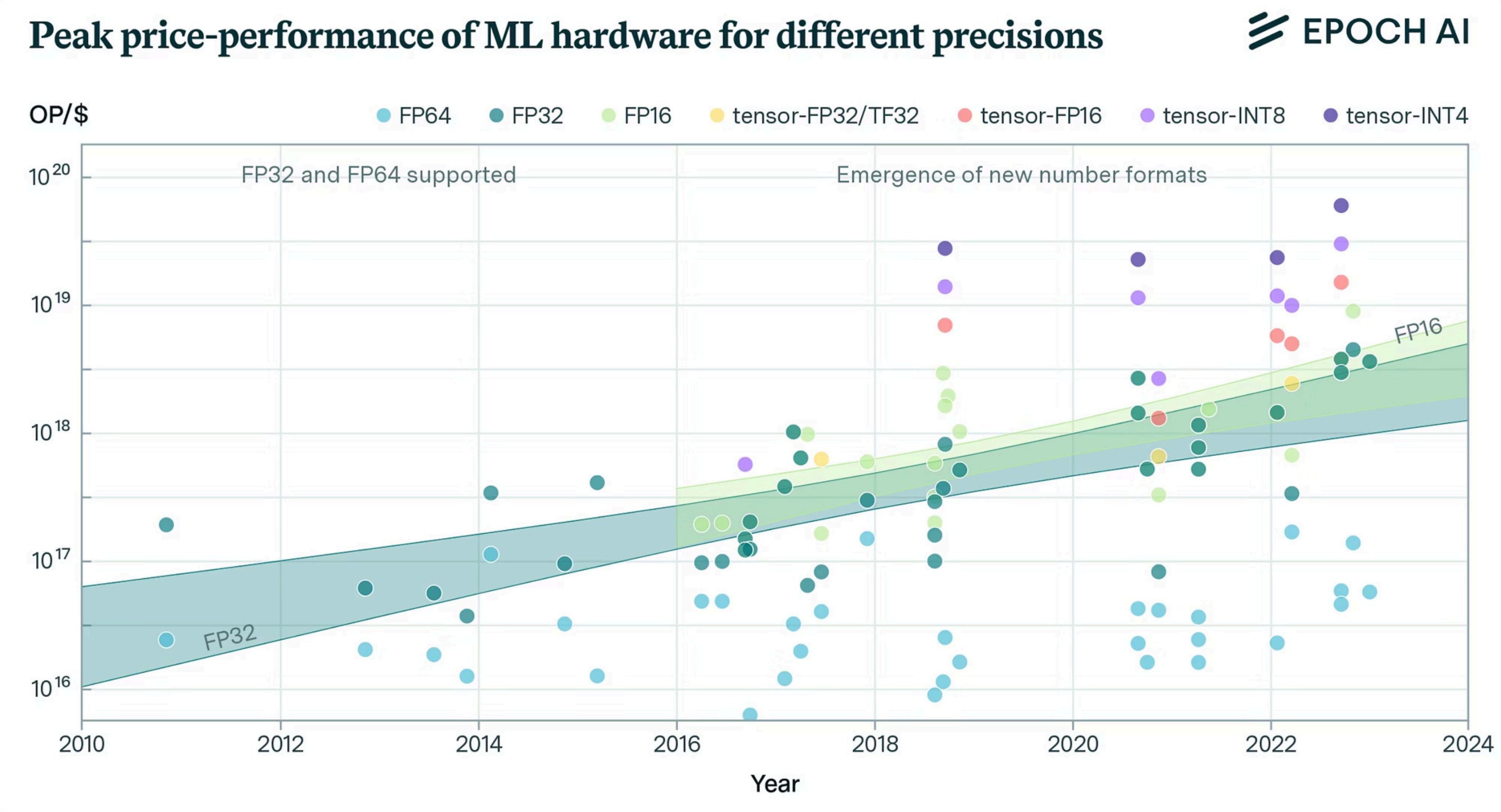
Historical Context: Learning from Past Revolutions

Patterns from previous technological transformations.

Challenges and Possible Policy Responses

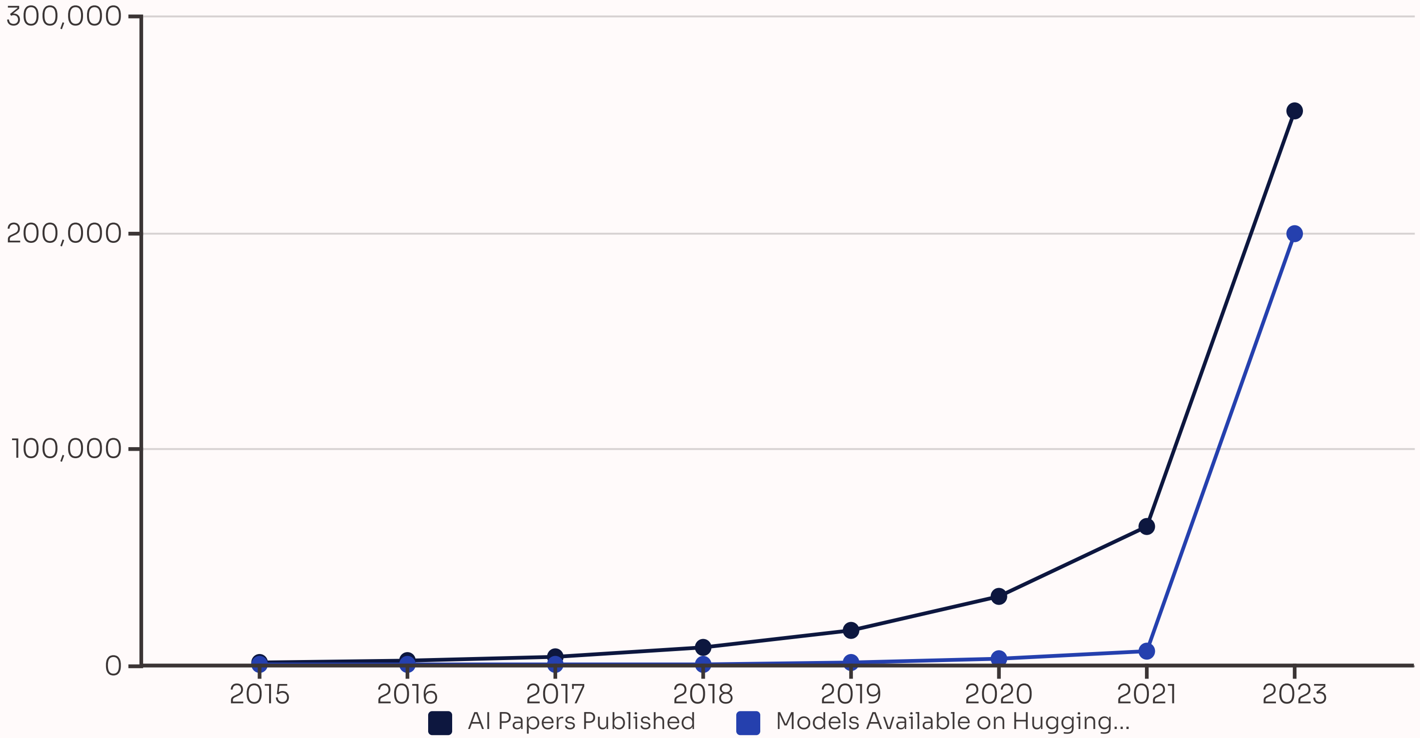
Key obstacles and possible policy solutions.

Speed → Exponential



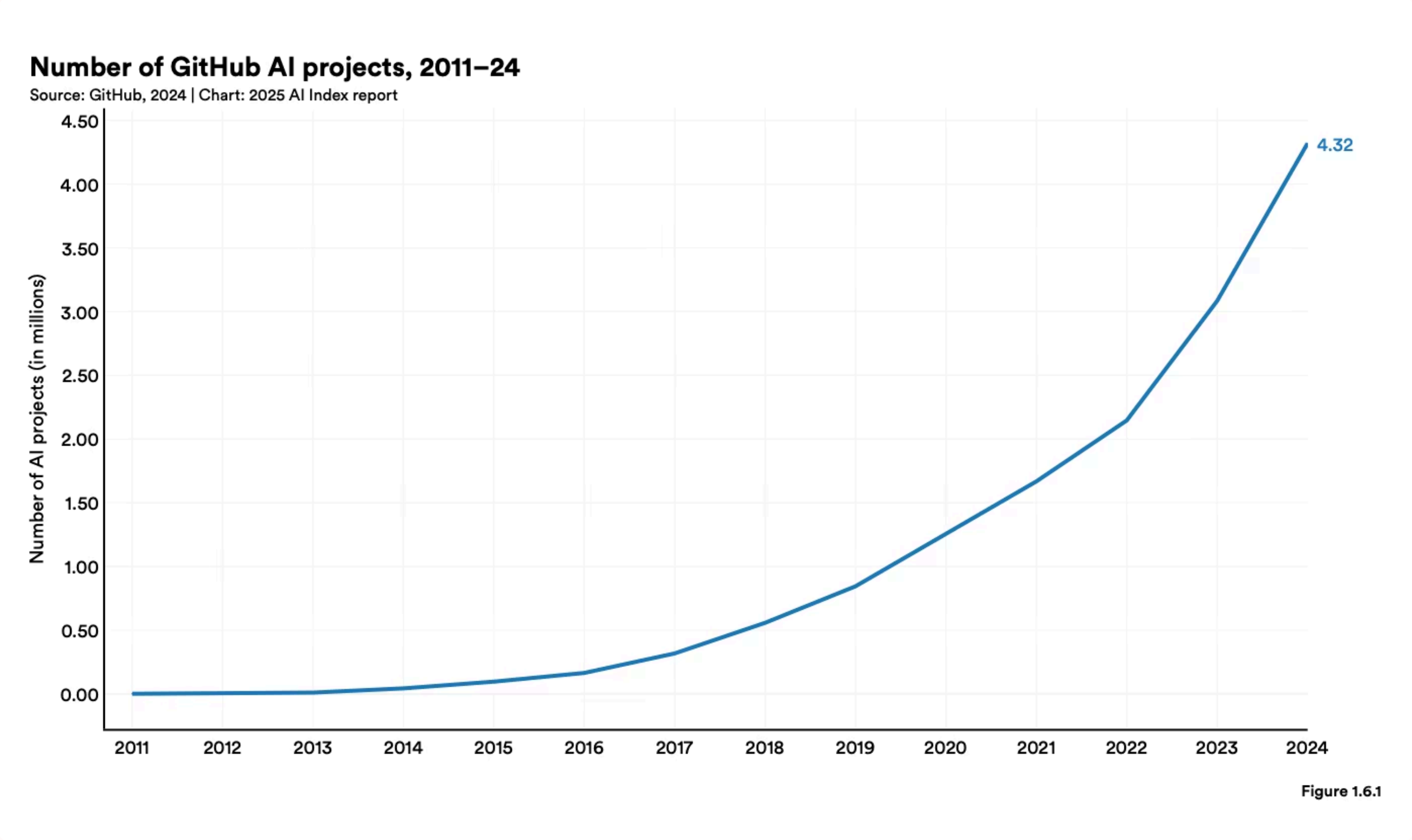
Peak price-performance of ML hardware for different precisions
EPOCH AI
OP/$
FP64
FP32
FP16
tensor-FP32/TF32
tensor-FP16
tensor-INT8
tensor-INT4
FP32 and FP64 supported
Emergence of new number formats
FP32
FP16
10^{20}
10^{19}
10^{18}
10^{17}
10^{16}
2010
2012
2014
2016
2018
2020
2022
2024
Year

Rapid Innovation



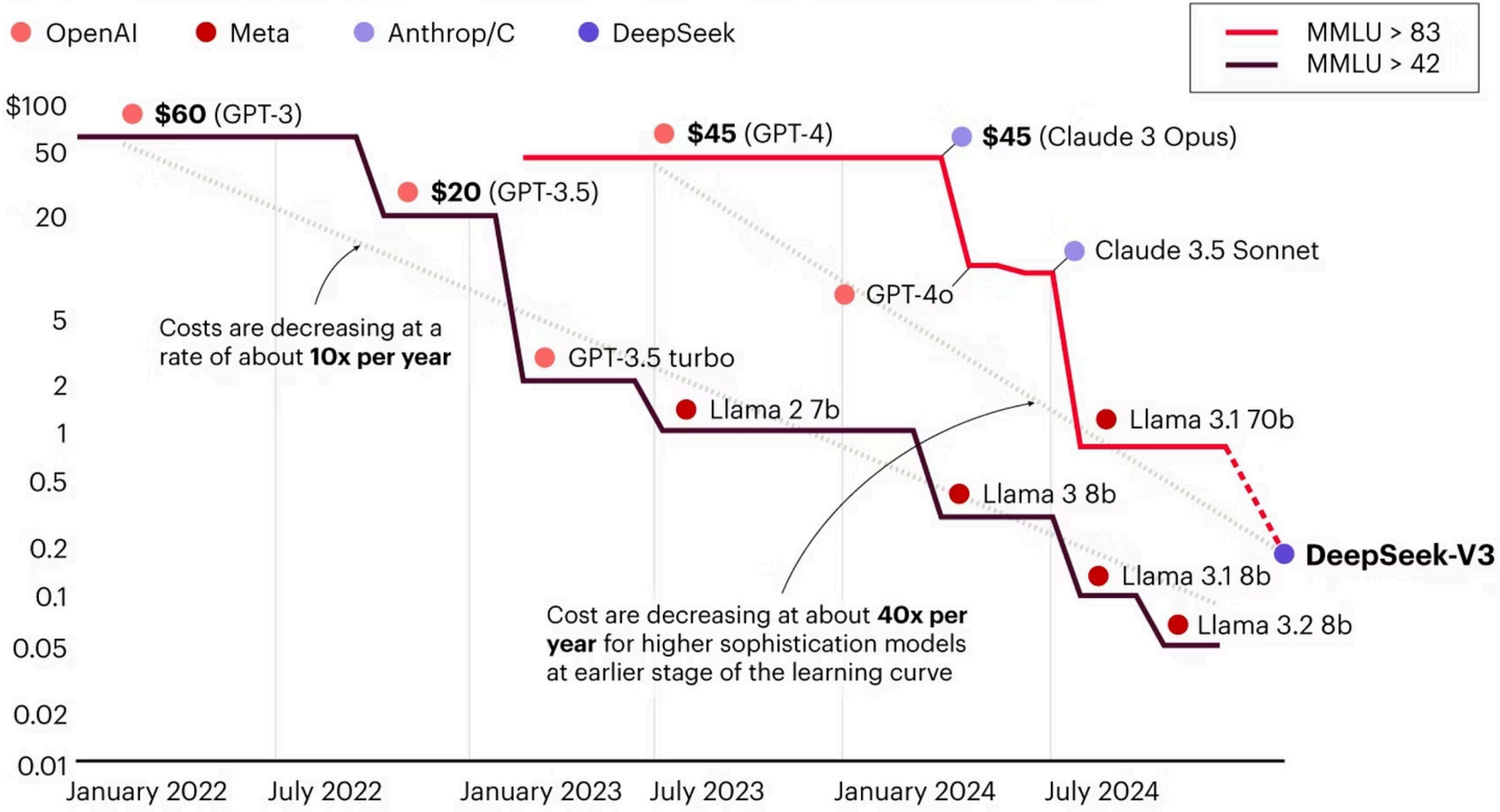
Originally describing the growth of transistors on a microchip, Moore's Law now applies to AI, with capabilities doubling every 3–6 months. This is evident in the rapid increase of AI papers published and models available on Hugging Face.

Rapid Innovation



Cost → Exponential

Cost of the cheapest LLM per million tokens by MMLU (in US dollars, log scale)



Source: Bain

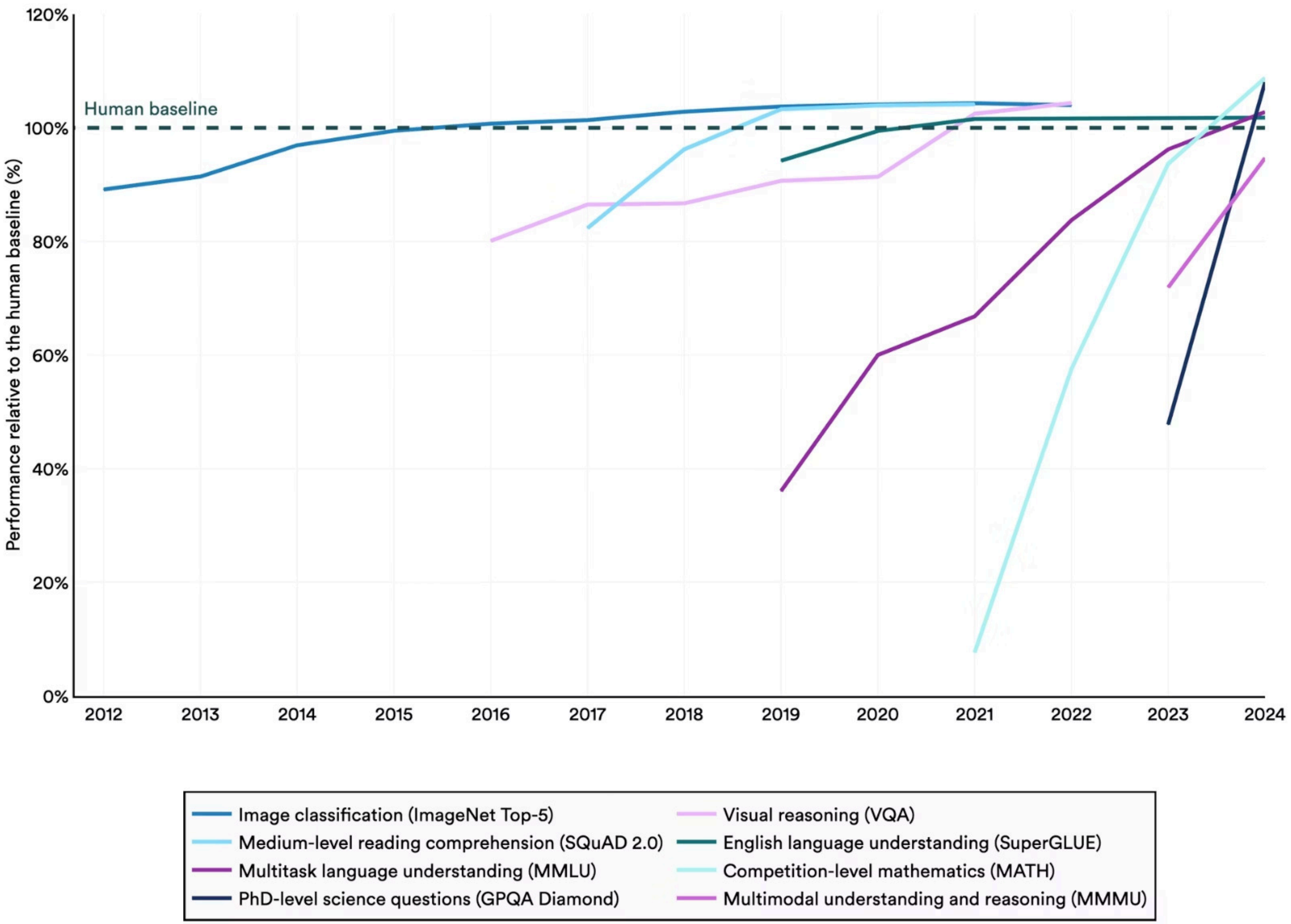
Jan 2025

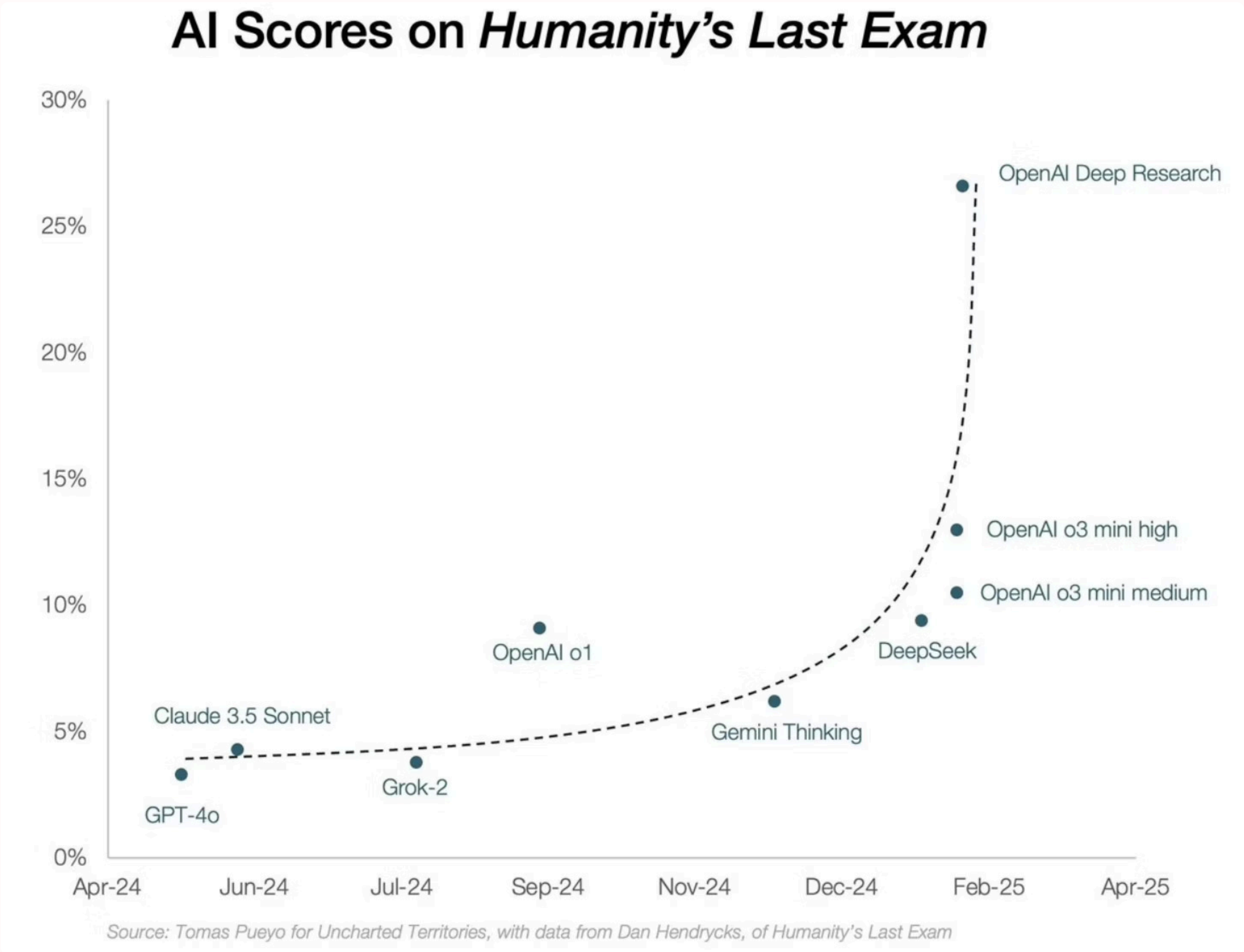
Data source: Epoch AI, Artificial Analysis

Capabilities → Rapid Improvements

Select AI Index technical performance benchmarks vs. human performance

Source: AI Index, 2025 | Chart: 2025 AI Index report




AI Scores on *Humanity's Last Exam*
30%
25%
20%
15%
10%
5%
0%
Apr-24
Jun-24
Jul-24
Sep-24
Nov-24
Dec-24
Feb-25
Apr-25
OpenAI Deep Research
OpenAI o3 mini high
OpenAI o3 mini medium
DeepSeek
OpenAI o1
Gemini Thinking
Claude 3.5 Sonnet
Grok-2
GPT-4o
Source: Tomas Pueyo for Uncharted Territories, with data from Dan Hendrycks, of Humanity's Last Exam

Mathematics

Question:

The set of natural transformations between two functors $F, G : C \to D$ can be expressed as the end

$$Nat(F,G) \cong \int_A Hom_D(F(A), G(A)).$$

Define set of natural cotransformations from F to G to be the coend

$$CoNat(F,G) \cong \int^A Hom_D(F(A), G(A)).$$

Let:
- $F = B_\bullet(\Sigma_4)_{*/}$ be the under ∞-category of the nerve of the delooping of the symmetric group Σ_4 on 4 letters under the unique 0-simplex $*$ of $B_\bullet \Sigma_4$.
- $G = B_\bullet(\Sigma_7)_{*/}$ be the under ∞-category nerve of the delooping of the symmetric group Σ_7 on 7 letters under the unique 0-simplex $*$ of $B_\bullet \Sigma_7$.

How many natural cotransformations are there between F and G?

Emily S
University of São Paulo

Computer Science

Question:

Let G be a graph. An edge-indicator of G is a function $a : 0, 1$
$V(G)$ such that $a(0), a(1) \in E(G)$.

Consider the following Markov Chain $M = M(G)$:
The statespace of M is the set of all edge-indicators of G, an
transitions are defined as follows:

Assume $M_t = a$.
1. pick $b \in 0, 1$ u.a.r.
2. pick $v \in N(a(1-b))$ u.a.r. (here $N(v)$ denotes the open neighbourhood of v)
3. set $a'(b) = v$ and $a'(1-b) = a(1-b)$
4. Set $M_{t+1} = a'$

We call a class of graphs $\mathcal{G}$ well-behaved if, for each $G \in \mathcal{G}$ t
Markov chain $M(G)$ converges to a unique stationary distribu
and the unique stationary distribution is the uniform distributio

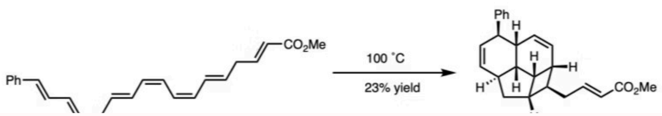
Which of the following graph classes is well-behaved?

Answer Choices:
A. The class of all non-bipartite regular graphs
B. The class of all connected cubic graphs
C. The class of all connected graphs
D. The class of all connected non-bipartite graphs
E. The class of all connected bipartite graphs.

Marc R
Queen Mary University of London

Chemistry

Question:



Linguistics

Question:

I am providing the standardized Biblical Hebrew source text fro
Biblia Hebraica Stuttgartensia (Psalms 104:7). Your task is to
distinguish between closed and open syllables. Please identify
list all closed syllables (ending in a consonant sound) based or

Agents → The Rise of the Infinite Intern

What AI could do in 2016

1 Second

of what it takes a human to do

What AI could do in 2025

Minutes

of what it takes a PhD intern to do

Understand goals

AI agents comprehend complex financial objectives and regulatory requirements without constant human guidance

Plan

Autonomous development of strategies to achieve financial and operational targets

Use tools

AI agents can operate and integrate with various financial tools and systems

Use software and computers

Direct operation of trading platforms, compliance systems, and data analysis software

Learn and remember

Continuous improvement through analysis of trading patterns, market responses, and past interactions

n interns to do m tasks at the same time

Replit

Replit – Build apps and sites with AI

Replit is an AI–powered platform for building professional web apps and websites.

Leave it to Manus

Manus is a general AI agent that bridges minds and actions: it doesn't just think, it delivers results. Manus excels at various tasks in work and life, getting everything done while you rest.

The Collapse of Cognitive Costs

Zero Marginal Cost

The cost of summarizing a 100-page report has fallen from hundreds of dollars for a human analyst to less than one cent in computational resources. This represents a cost reduction of over 99.99%.

Cognitive Automation

For the first time in history, automation directly targets knowledge work at scale—analysis, writing, coding, decision support. These were the "safe" jobs that absorbed workers displaced by previous waves of automation.

From Tools to Agents

We're witnessing a fundamental shift from passive tools like Excel or Word that wait for commands, to active agents that plan, execute, and adapt—systems that can pursue goals with minimal human supervision.

Where Are We Going?

Horizon 1: Copilots (Now - 2025)

AI augments human work with constant supervision. Expect 20-40% productivity gains through tools like writing assistants and code completion.

Horizon 2: Agents (2026 - 2027)

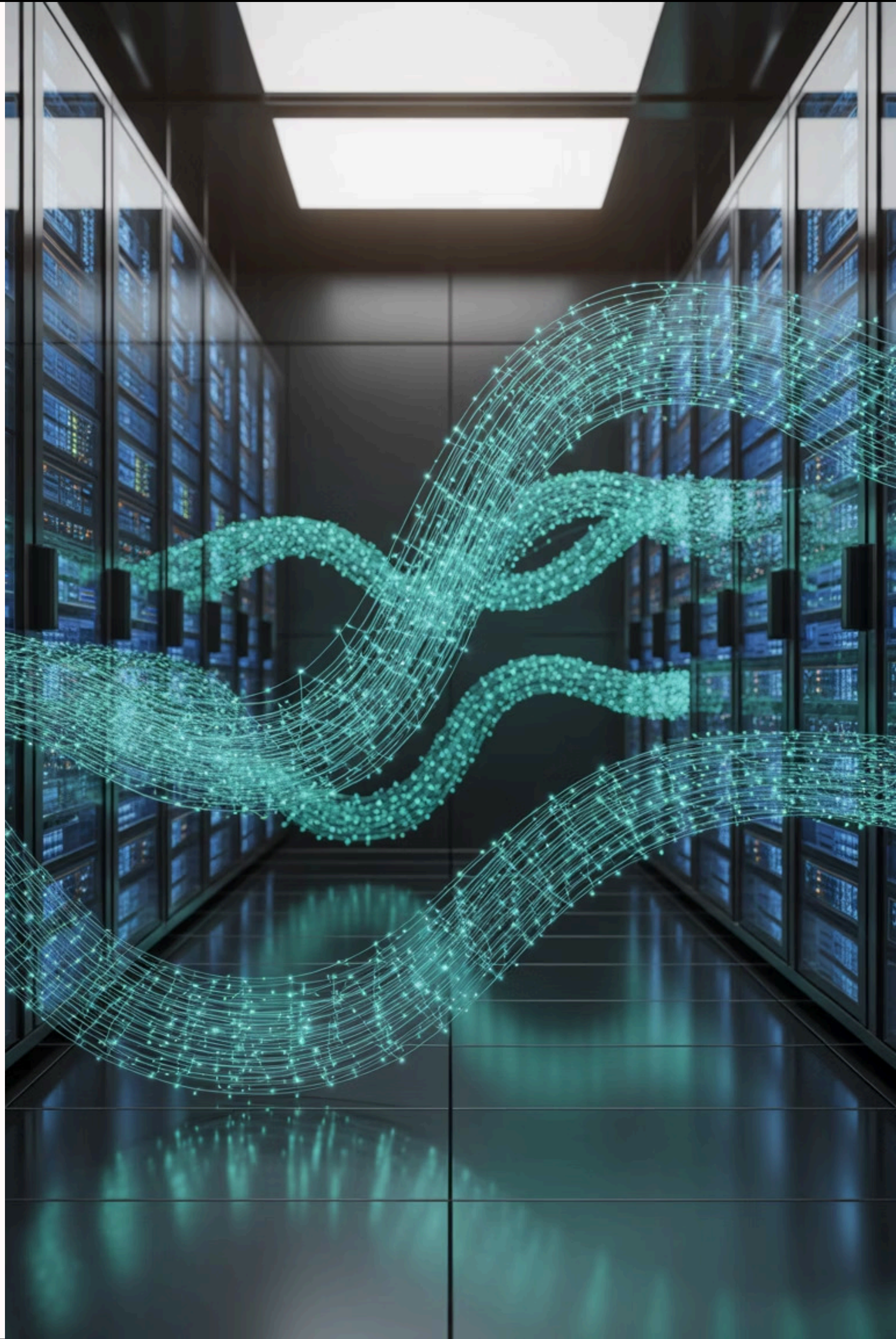
AI systems manage full workflows with periodic oversight, fundamentally transforming job roles. This includes autonomous research and process automation.

Horizon 3: System Reshuffle (2028+)

Multiple AI agents collaborate dynamically, reimagining organizational structures. This leads to automated project teams and adaptive service delivery systems.

Seeing is Believing

Live Demonstrations



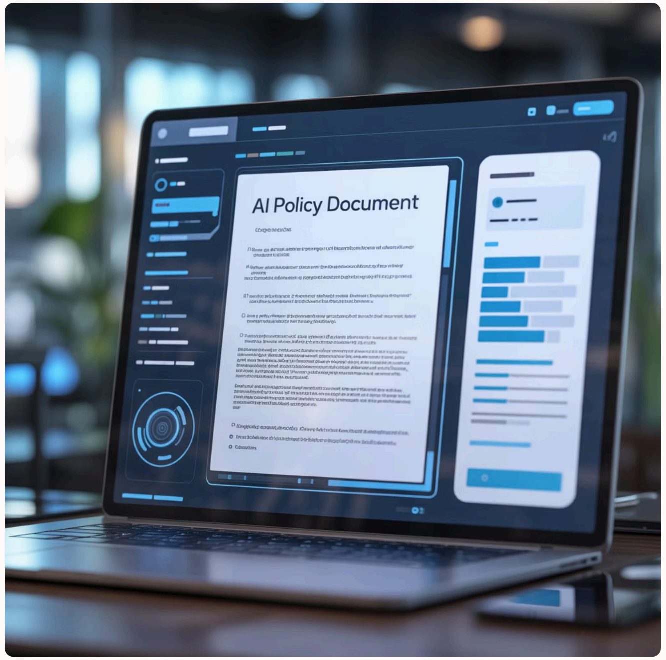
Policy Drafting (The Old Way)

The Task

Draft a briefing note on 'Healthcare Policy for Aging Population'

The Prompt

Act as a senior policy analyst. Draft a briefing note on integrating AI into eldercare. Cover: Privacy risks, workforce augmentation, infrastructure needs. Include executive summary and three policy recommendations.



Implication: The 'Blank Page Problem' is effectively solved. Policymakers shift from being writers to becoming editors and strategic thinkers. Time is redirected from document creation to stakeholder consultation, political judgment, and implementation planning.

The New Way

Our AI framework uses a hierarchical multi-agent architecture to deliver comprehensive economic assessments rapidly, mirroring traditional analytical rigor.

01

Data Collection

Automated real-time economic indicator gathering.

02

Parallel Analysis

Agents concurrently process macroeconomic trends and policy frameworks.

03

Risk Assessment

Identifies vulnerabilities via stress modeling.

04

Policy Recommendations

Evidence-based advice, IMF Article IV compliant.

05

Report Synthesis

Automated report compilation (3-5 minutes).

Specialized AI Agents

1

Data Collection Agent

Connects to authoritative databases, validates economic indicators, and ensures data consistency.

2

Macroeconomic Analysis Agent

Assesses GDP growth, inflation, labor markets, and economic outlook using statistical models.

3

Monetary & Fiscal Policy Agent

Evaluates policy stance, debt sustainability, and monetary-fiscal coordination.

4

External Sector Agent

Examines current account, foreign reserves, exchange rates, and external vulnerabilities.

5

Risk Assessment Agent

Identifies downside scenarios, conducts stress testing, and quantifies risks.

6

Policy Recommendation Agent

Synthesizes findings to formulate actionable, evidence-based policy recommendations.

7

Report Synthesis Agent

Compiles comprehensive analysis into coherent, publication-ready format following IMF Article IV standards, complete with executive summary, detailed findings, and supporting documentation.

Key Capabilities & Value Proposition

Technical Capabilities

Real-Time Data Integration

Maintains live API connections to multiple authoritative sources with automated validation protocols ensuring data accuracy and timeliness for current analysis.

Parallel Processing Architecture

Enables concurrent execution of specialized agents, reducing total processing time to 3-5 minutes while maintaining analytical depth through scalable cloud infrastructure.

Comprehensive Analytical Coverage

Integrates macroeconomic, fiscal, monetary, and external sector analysis with systematic risk assessment and evidence-based policy recommendation generation.

Professional Report Generation

Produces publication-ready output adhering to IMF Article IV format standards, including executive summary and full analytical report with supporting data visualizations.

Value for Central Banks

Speed & Efficiency

Reduces analysis timeframes from weeks to minutes, enabling rapid scenario analysis and real-time economic monitoring for time-sensitive policy decisions.

Quality & Consistency

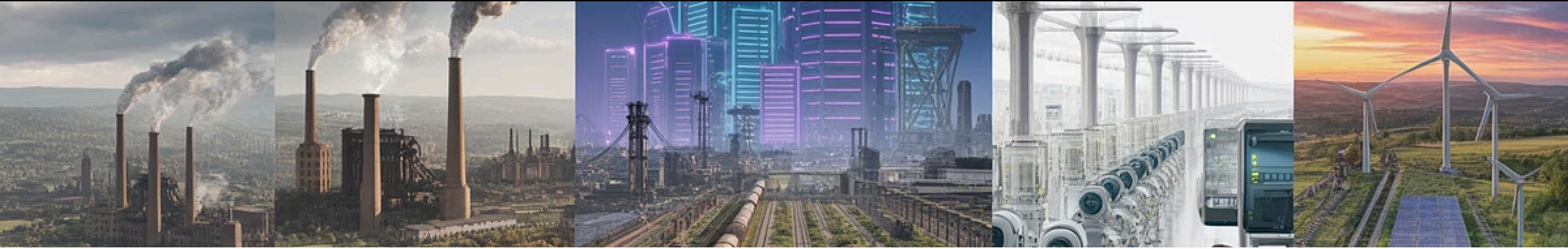
Ensures standardized methodology across analyses, producing reproducible results with comprehensive coverage that maintains institutional knowledge and analytical standards.

Operational Scalability

Supports simultaneous analysis of multiple countries, facilitates regular monitoring cycles, and adapts to custom institutional applications without proportional resource increases.

Practical Applications

Streamlines Article IV consultation preparations, accelerates policy brief development, enhances risk monitoring capabilities, and supports high-quality board paper production.



Understanding Technological Revolutions

Engels' Pause

The Industrial Revolution's Distribution Crisis

The Data (1780-1840)

- Productivity rose 46%
- Wages rose only 12%
- Living standards for workers stagnated or declined
- 60 years elapsed before workers substantially benefited

Why the Gap?

- Institutions lagged behind technology
- Mechanisms for sharing prosperity took decades to develop

Our goal must be to shorten this pause. We cannot accept another 60 year gap between productivity gains and widely shared prosperity. The question is whether we have the institutional imagination and political will to compress this timeline.

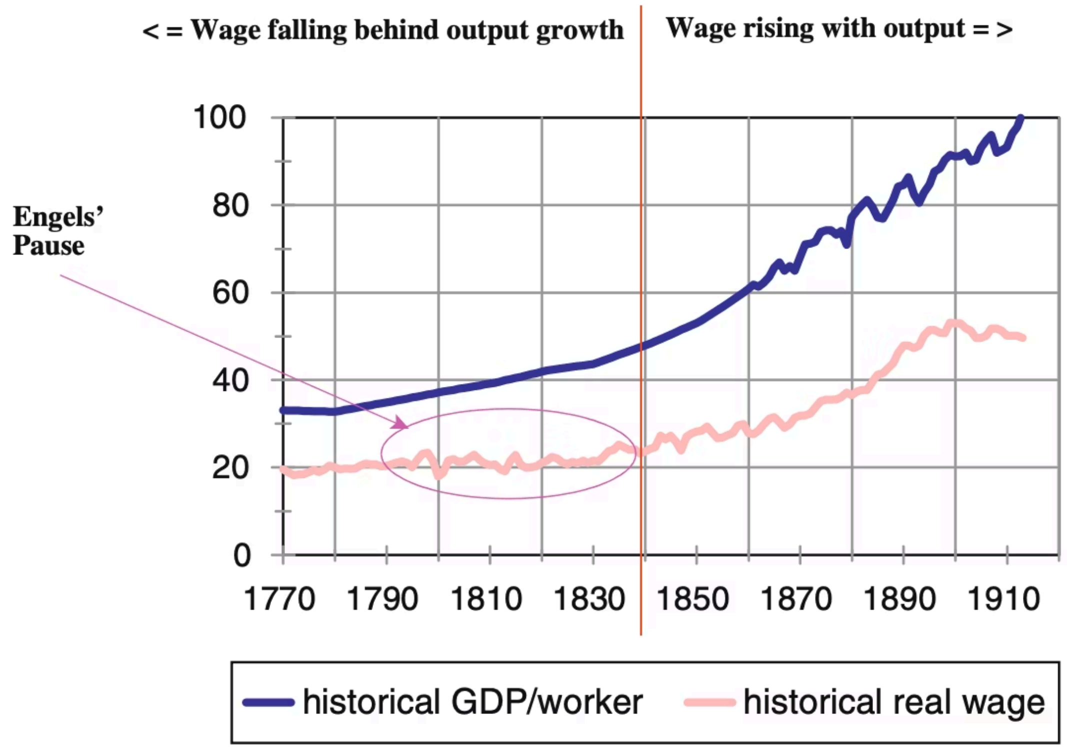


Fig. 1. The two phases of the British industrial revolution.

Allen, R.C. Engels' pause: Technical change, capital accumulation, and inequality in the british industrial revolution. Explor. Econ. Hist. (2009), doi:10.1016/j.eeh.2009.04.004

The Solow Paradox

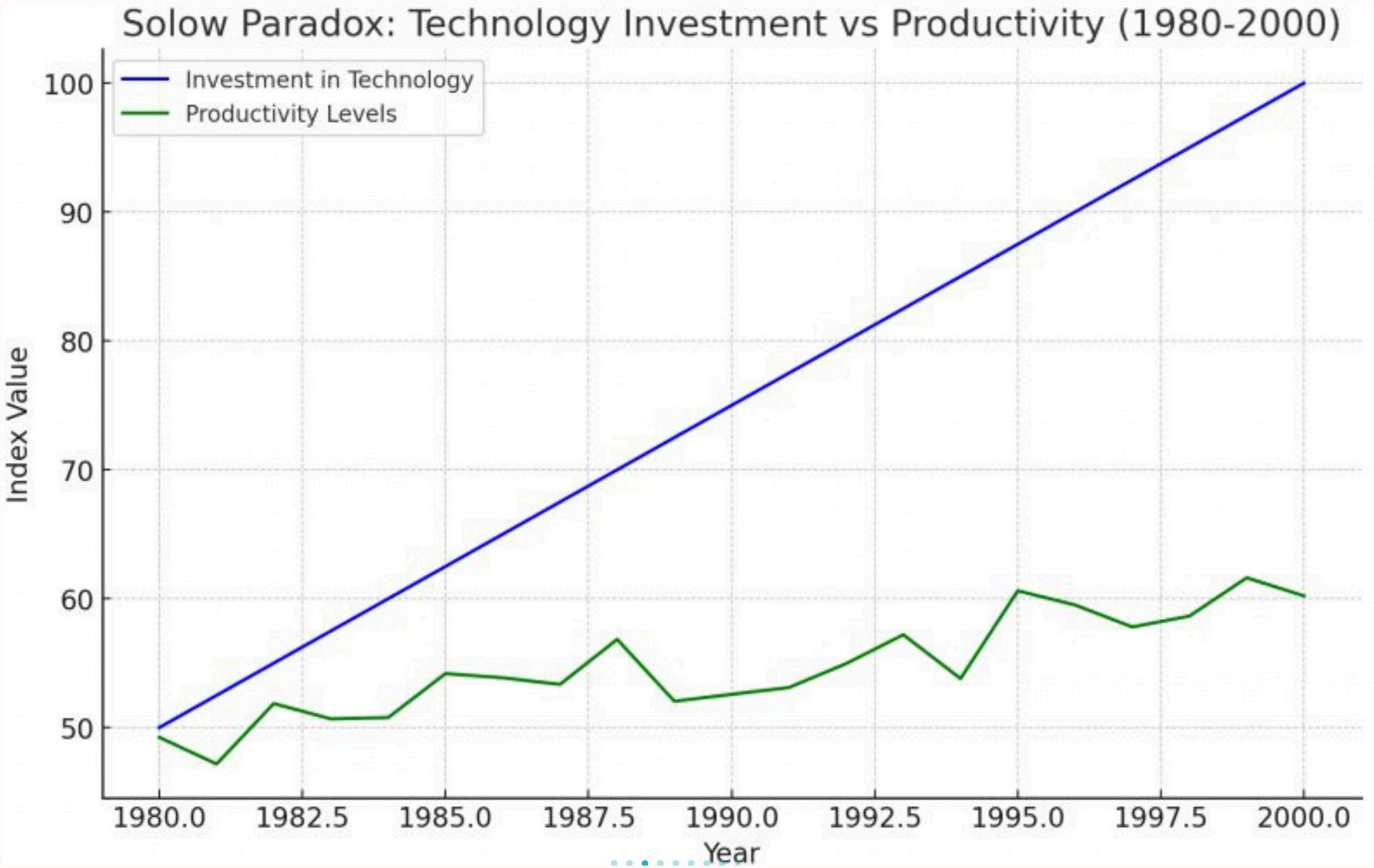
> "You can see the computer age everywhere but in the productivity statistics."
>
> — Robert Solow, Nobel Laureate, 1987

The Lesson: Technology Without Transformation

Simply electrifying a factory designed for steam power delivered minimal benefits. Real productivity gains required completely reimagining factory layouts, workflows, and organizational structures—a process that took decades.

The Application to AI

Paving cow paths—digitizing fundamentally bad processes—yields zero gain. We cannot simply add AI to existing bureaucratic workflows and expect transformation.


Solow Paradox: Technology Investment vs Productivity (1980-2000)
Investment in Technology
Productivity Levels
Index Value
100
90
80
70
60
50
1980.0
1982.5
1985.0
1987.5
1990.0
1992.5
1995.0
1997.5
2000.0
Year

Perez's Four Phases of a Revolution

Irruption (0-10 years)

New technology emerges. Financial capital floods in. Early adopters experiment. Infrastructure begins building.

Frenzy (10-20 years)

Speculation explodes. Infrastructure wildly overbuilt. Wealth inequality surges. We are likely here.

Turning Point

Bubble bursts. Financial crisis erupts. Regulatory reckoning begins. Society demands reform.

Synergy (20-40 years)

Technology spreads broadly. Institutions catch up. 'Golden Ages' emerge. Benefits widely distributed.

The Perez Framework: Five Great Surges

1. Industrial Revolution (1771)

Mechanization, cotton industry, iron works

2. Age of Steam & Railways (1829)

Steam power, railways, telegraph

3. Age of Steel & Electricity (1875)

Steel, heavy engineering, electrical power

4. Age of Oil, Automobiles & Mass Production (1908)

Oil, automobiles, petrochemicals, mass production

5. Age of Information & Telecommunications (1971)

Microelectronics, computers, telecommunications, internet

The Claim: Each revolution follows a remarkably similar pattern of diffusion, crisis, and eventual "Golden Age." The pattern is: Initial innovation → Rapid diffusion → Financial frenzy → Bubble collapse → Regulatory reckoning → Broad deployment → Shared prosperity.

AI (2022?) may be witnessing the sixth great surge. If the pattern holds, understanding where we are in the cycle becomes crucial for policy timing.

Why Bubbles Are a Feature, Not a Bug

The Paradox

Financial bubbles are destructive to individual investors and create social instability. Yet they are also essential to technological progress.

The Function

Irrational exuberance mobilizes capital at unprecedented scale. No rational calculation would justify the infrastructure investments that bubbles produce.

Railways (1840s)

Wildly overbuilt, massive bankruptcies—but we got comprehensive rail infrastructure that powered the next 50 years of growth

Fiber Optics (1990s)

Catastrophic over-deployment, dot-com crash—but we got cheap, abundant bandwidth that enabled the digital economy

AI Data Centers (2020s??)

Being built at unprecedented scale right now. When the music stops, what infrastructure will remain?

Challenges and Possible Policy Responses

Valuations Are Objectively Stretched

The evidence is unambiguous: by virtually every historical measure, equity markets are trading at extreme valuations. These are not subjective assessments but quantifiable metrics that place current market conditions in the 95th percentile of the historical distribution.

Key Valuation Metrics

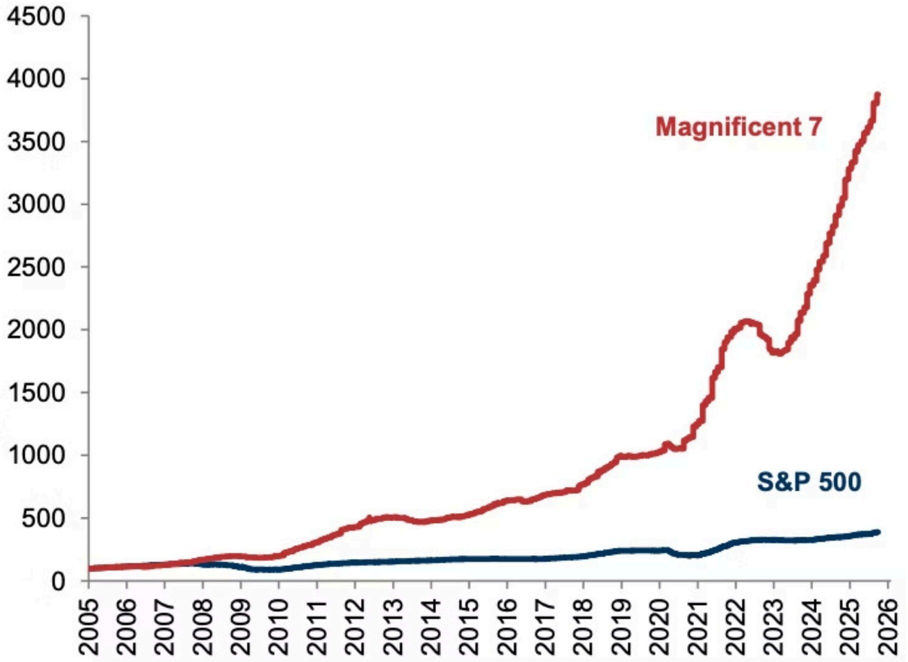
- **Buffett Indicator:** 224.2% (highest ever recorded, surpassing dot-com peak of 130-150%)
- **S&P 500 Trailing P/E:** ~30x vs. 15x historical average
- **Forward P/E:** ~29x, indicating expectations remain elevated
- **Margin of Safety:** <2% vs. Buffett's preferred 20%+ threshold

What This Means

- These metrics demonstrate that stretched valuations are a measurable fact, not a matter of interpretation.
- We're in the extreme tail of valuations—historically a precursor to either exceptional returns or painful corrections.

The critical question is not whether valuations are high, they objectively are—but whether underlying fundamentals can support these levels.

Magnificent 7 and S&P 500, 12m trailing EPS, Jan 2005=100



Source: FactSet, Goldman Sachs GIR.

2025 AI ≠ 1999 Dot-Com

Today's AI-driven market is fundamentally different from the dot-com bubble, with distinct economic models, revenue generation, and competitive dynamics.

1999 Dot-Com Reality

Revenue: Single-digit billions or zero

Profitability: Largely negative, no clear path

Competitive moat: None, low barriers

Network effects: Unproven, theoretical

Adoption: Low (5-15% households)

2025 AI Leaders

Revenue: High ($126B+ for key players)

Profitability: Strong gross margins (>60%)

Competitive moat: Unmatched semiconductor architecture

Proven ROI: Measurable returns on capital expenditure

Adoption: Widespread (80%+ Fortune 500)

AI chip revenue is growing at **~35% year-on-year**, with infrastructure spending reaching $82 billion in Q2 2025 alone. IDC forecasts the market to reach $758 billion by 2029, reflecting **significant capital deployment** and **proven productivity gains**, **not speculative projections**.

The Bull Case: Early-Stage Euphoria Can Be Justified

Despite stretched valuations, a compelling bull case suggests current pricing reflects rational expectations of AI's transformative economic impact. Four pillars support this perspective.

Legitimate Capex ROI

Enterprise AI adoption by Fortune 500 companies yields 5-15% productivity gains and efficient capital deployment, generating real compounding returns.

Winner-Take-Most Economics

High barriers to entry and market concentration, exemplified by NVIDIA's 60%+ gross margins, create compounding advantages for early leaders.

TAM Expansion is Real

The AI market is projected to grow from $391B to $3.5T by 2033, creating entirely new markets with a total revenue potential of $15.7T by 2030.

Early Adopter Advantage

Only 2% of AI deployment is complete, suggesting decades of growth ahead, similar to the Internet's 25-year evolution. Early positioning justifies current valuations.

OpenAI projects revenue exceeding $125 billion by 2029. If these projections are even partially accurate, current valuations reflect rational pricing of transformative potential.

Real Warning Signs to Monitor

A rigorous assessment requires identifying specific, measurable indicators that would signal the transition from justified optimism to speculative excess. These six warning signs provide an evidence-based framework for distinguishing sustainable growth from bubble dynamics.

Valuation Compression Without Earnings Growth

Monitor S&P 500 earnings growth vs. P/E expansion. Risk: earnings growth <3% with P/E ratios >28x indicates disconnected valuations.

AI Capex Returns Deteriorate

Monitor data center ROI (revenue per $1 capex). Risk: Ratio falling below $2:$1 weakens the economic case for AI infrastructure investment.

Funding Concentration in Unprofitable Startups

Monitor VC funding to unprofitable AI startups. Risk: Over 70% of AI funding to pre-revenue companies signals speculative excess.

Margin Compression at Leaders

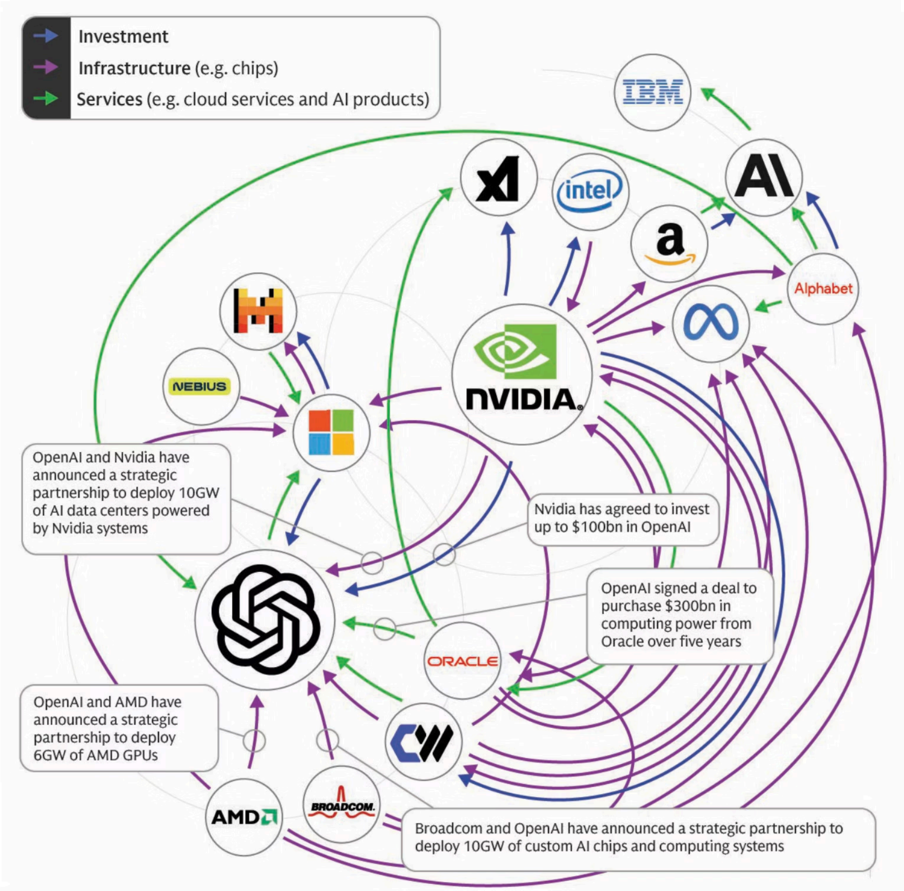
Monitor NVIDIA gross margins (currently 60%+). Risk: Margins below 50% could break the "winner-take-most" thesis.

GDP Growth Divergence

Monitor Buffett Indicator vs. real GDP growth (currently 224% with <2% GDP growth). Risk: If GDP <1.5% and Buffett Indicator >250%, disconnect becomes unsustainable.

Productivity Claims Not Materializing

Monitor labor productivity growth vs. AI spending. Risk: If productivity growth hasn't accelerated to 2-3% CAGR by Q4 2026, the fundamental ROI story is questionable.



Note: Exhibit does not represent an exhaustive view of the AI ecosystem but rather the most prominent AI model companies, infrastructure providers, and hyperscalers.
Source: Company announcements, various news sources, compiled by Goldman Sachs GIR.

The Turning Point

The Crisis

Financial crashes (1847, 1929, 2000) mark the end of the Frenzy phase. The pattern is consistent: speculative excess meets reality, triggering systemic correction and social upheaval.

The Reckoning

Society demands regulation and reform. The excesses of 'Gilded Ages'—extreme inequality, labor exploitation, environmental damage—give way to institutional innovation. This is when unions form, antitrust laws emerge, and social safety nets are constructed.

The Opportunity

This is when policy matters most. Institutional choices made during the Turning Point determine the nature of the next 30 years—whether the Synergy phase delivers broadly shared prosperity or concentrates gains among a narrow elite.

Labor Markets

Jobs, Tasks, and the Dual Effect

The impact of AI on employment is neither simple displacement nor pure augmentation. Understanding the nuanced dynamics of **displacement** versus **reinstatement** is essential for effective policy response.

The Dual Effect of Technology

Displacement Effect

Automation replaces human labor in existing tasks, reducing demand for workers performing those specific activities. This is the immediate, visible impact that dominates public discourse.

Reinstatement Effect

New technology creates entirely new tasks and roles that didn't previously exist, generating fresh demand for human labor in novel areas.

Historical Record

Both effects always occur. The critical questions are timing—how quickly does reinstatement offset displacement—and distribution—do the same workers benefit, or different populations?

The AI Difference: For the first time, automation targets cognitive tasks at scale—analysis, writing, coding, diagnosis. These were the "safe harbor" jobs that absorbed workers displaced from manufacturing and agriculture.

Thinking in Tasks, Not Jobs

The Insight (Acemoglu & Restrepo)

Jobs are bundles of tasks. Technology automates specific tasks, not entire jobs. This distinction is crucial for understanding labor market dynamics and designing effective policy responses.

Example: Policy Analyst

Tasks Automated

- Literature review and research synthesis
- First draft generation
- Data summarization and basic analysis

Tasks Augmented

- Strategic synthesis across domains
- Stakeholder judgment and political timing
- Complex trade-off evaluation

Tasks Created

- Prompt engineering and AI orchestration
- AI output quality assessment
- Algorithmic accountability oversight

The Implication: The job of policy analyst persists, but its composition fundamentally changes. Value shifts from execution speed to judgment quality. Productivity rises, but the skills required transform.

The Displacement Concern: Who Is Most Exposed?

Routine Cognitive Work

Jobs characterized by rule-based, predictable, information-intensive tasks face the highest displacement risk. These include back-office processing, basic legal research, financial analysis, customer service, and significant portions of software development.

The Skill Paradox

AI disproportionately affects the middle and upper-middle portions of the skill distribution—precisely the jobs that absorbed displaced manufacturing workers over the past 40 years. Where do these workers go next?

40%

Knowledge Workers

Estimated task exposure to AI automation

25%

Service Jobs

Portion vulnerable to AI displacement

60%

Routine Tasks

Share automatable in current roles

Geographic & Demographic Concentration: Displacement will not be evenly distributed. Urban centers with high concentrations of knowledge workers may face acute disruption. Education levels, age, and existing skill bases will create winners and losers.

The Compensating Effects

While displacement risks are real, several powerful compensating mechanisms can mitigate job losses—but only if deliberately activated through policy and institutional design.

Productivity Complementarity

AI augments workers, delivering 30-50% productivity gains and potentially increasing demand for enhanced output.

Cost Reduction → Demand Expansion

Cheaper cognitive services make previously unaffordable services accessible, expanding total market size.

New Task Creation

New technologies create new tasks and roles that didn't previously exist.

The Policy Lever: Reinstatement is not automatic. It requires active investment in skills development, institutional adaptation, and labor market policies that create conditions for new task emergence.

The Experience Trap: Cutting the Bottom Rungs

The Mechanism

AI systems excel at routine cognitive tasks—precisely the tasks that junior professionals use to develop expertise. First drafts, basic research, initial analysis—these are learning opportunities, not just work products.

The Problem

If AI writes all first drafts and conducts all preliminary analysis, how do junior lawyers, analysts, doctors, accountants, and engineers develop the tacit knowledge, pattern recognition, and professional judgment that defines expertise?

The Risk

High productivity today may create a crisis of expertise in 5-10 years. We risk creating a generation of professionals who can *direct* AI but lack the foundational knowledge to *evaluate* its output.

Examples Across Professions: Junior lawyers no longer draft routine motions. Entry-level analysts never build financial models from scratch. Medical residents rely on AI diagnostic suggestions without developing clinical intuition. In each case, short-term efficiency may undermine long-term capability development.

Skills and Expertise in the AI Era

What Becomes Valuable?

Tacit Knowledge

Judgment, intuition, and contextual understanding that can't be easily codified or transferred to AI.

Visible Expertise

The ability to explain decisions, understand ethics, and navigate stakeholder dynamics.

Orchestration Skills

Managing AI systems effectively—knowing when to trust, override, and combine tools.

Learning Velocity

The ability to unlearn and relearn rapidly. Adaptability matters more than accumulated knowledge.

Rethinking Professional Development

The Old Model

A linear progression over years:

1. Learn fundamentals through coursework
2. Practice routine tasks repeatedly
3. Gradually build to complexity
4. Eventually develop judgment

This model assumed routine practice was necessary to develop expertise.

The New Model

Accelerated, judgment-focused from day one:

1. Engage with complex problems immediately
2. Use AI to handle routine execution
3. Focus on evaluation and direction
4. Develop judgment through curation, not creation

Juniors learn to assess and guide rather than execute from scratch.

Institutional Response Required

01

Restructure Apprenticeship

Redesign training programs to preserve learning opportunities while leveraging AI augmentation

02

Create AI-Assisted Learning Pathways

Develop explicit curricula for working effectively with AI tools throughout professional development

03

Measure Outcomes Differently

Assess judgment quality and decision-making capability, not just throughput and speed

Trust & Black Box Dilemma

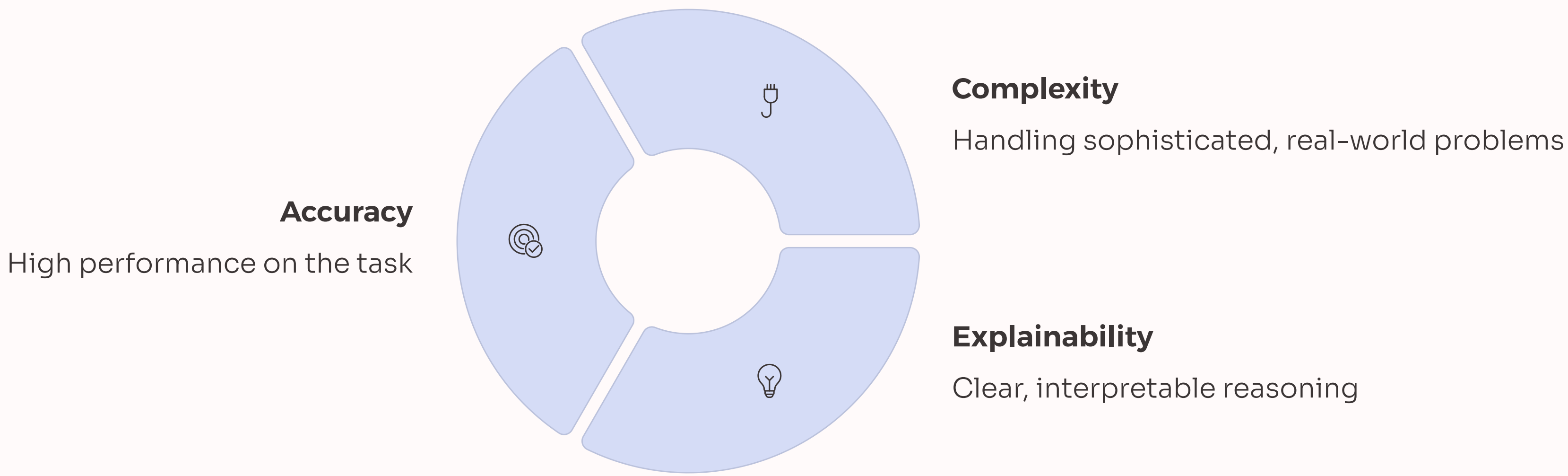
The Fundamental Conflict

Public policy inherently requires accountability, transparency, and explicit reasoning. Citizens have the right to understand why they were denied a benefit, rejected for a license, or subjected to enforcement action.

Yet modern AI systems, particularly the most powerful ones—operate fundamentally through probability and pattern matching, not logical reasoning. They cannot always explain *why* they reached a particular conclusion in terms humans find satisfying.

The Explainability Trilemma

In practice, you can usually achieve only two of three desirable properties:



The Risk: "Computer says no" is not acceptable for denying public benefits, rejecting permits, or making consequential decisions affecting citizens' lives. Yet this is where AI's limitations meet government's obligations most directly.

Operationalizing Trust

We cannot wait for perfect explainability before deploying AI in government. Instead, we must build practical frameworks that enable responsible use despite limitations.

From 'No' to 'Guardrails'

Regulatory Sandboxes

Create safe spaces with explicit error budgets for experimentation. Allow agencies to test AI systems on non-critical decisions with human oversight and clear rollback procedures.

'AI Verify' Testing Frameworks

Implement systematic validation approaches like Singapore's AI Verify—benchmark model behavior against safety, fairness, and robustness standards before deployment in production systems.

Human-Over-The-Loop Architecture

Establish clear rules: for all consequential decisions affecting individual rights or benefits, a human must be accountable for the outcome and have genuine authority to override AI recommendations.

Algorithmic Impact Assessments

Require documentation of potential harms, mitigation strategies, and monitoring approaches before deploying AI systems in high-stakes contexts.

The 95% Pilot Trap

The Reality

It's remarkably easy to build an impressive demonstration. It's incredibly hard to move from pilot to production at scale. This gap—between proof-of-concept and operational deployment—is where most government AI initiatives die.

Why Pilots Fail to Scale

Data Quality

Real-world government data is messy, siloed across agencies, stored in incompatible formats, and often of poor quality. Pilots use cleaned datasets; production faces reality.

Legacy Systems

AI tools don't easily integrate with databases and workflows designed in the 1990s. The technical debt of government IT infrastructure creates massive barriers.

Cultural Resistance

Fear of job displacement, concerns about accountability, skepticism about technology—organizational culture can kill promising initiatives regardless of technical merit.

The Adoption Architecture

Successful AI adoption in government relies less on "AI Strategy" and more on "Implementation Playbooks" focused on robust infrastructure.

Data Governance

Treat data as a strategic asset with clear ownership and quality standards. Essential for AI success.

Procurement Reform

Replace traditional procurement with agile contracting and rapid prototyping to enable AI iteration.

Talent Development

Build broad AI literacy across the civil service, not just a few specialists.

Key Takeaway: The bottleneck is rarely the AI technology itself. The bottleneck is data infrastructure, procurement processes, change management, and organizational culture. Fix these, and adoption becomes possible.

Three Pillars for Adoption Success

Pillar 1: Institutional Readiness

Building foundational capabilities: governance, data, talent, and procurement systems are essential for effective AI adoption.

Pillar 2: Strategic Use Cases

Focus on high-impact, feasible applications that prioritize utility, such as reporting automation and citizen services.

Pillar 3: Ecosystem Collaboration

Coordinate engagement between research, government, and private sectors to accelerate innovation and deployment.

These pillars are **interdependent**; success requires **coordinated progress** across **all three** for **effective AI adoption**.

What Governments Can Do

Fix the Adoption Architecture

- Modernize procurement for agile AI projects.
- Establish whole-of-government data governance.
- Build AI literacy through training programs.
- Create Centers of Excellence for support.

Implement Proactive Labor Market Policy

- Redesign professional development with AI.
- Invest in task reinstatement (new roles).
- Provide transition support for at-risk workers.
- Partner with education for AI-era curriculum.

Enable Whole-of-Government Collaboration

- Build shared AI infrastructure.
- Establish common standards for safety and ethics.
- Coordinate risk management across agencies.
- Create mechanisms for rapid knowledge sharing.

What Organizations Can Do

Redesign Work, Not Just Add AI

Fundamentally rethink workflows to integrate AI. Automate suitable tasks, augment human judgment, and create new roles for AI oversight, focusing on measuring outcomes over mere efficiency.

Protect the Experience Pipeline

Preserve junior staff development by restructuring entry-level roles and implementing "AI-assisted apprenticeships." Focus on professional development outcomes and effective mentorship in an AI-augmented environment.

Build Responsible Adoption Culture

Cultivate a culture that encourages responsible AI use. Train staff on AI capabilities and limitations, foster experimentation with clear error budgets, and establish transparent ethical guidelines.

What Can I Do?

1

Become AI-Literate

Develop a practical understanding of AI's capabilities and limitations. Gain intuition through regular, hands-on experience with AI tools.

2

Invest in Complementary Skills

Focus on human-centric skills AI struggles with: judgment, ethical reasoning, relationship building, and creative problem-solving. These skills will become increasingly valuable.

3

Develop Learning Velocity

Cultivate the ability to quickly unlearn and relearn. Intellectual flexibility and continuous adaptation are more crucial than specific expertise.

4

Be a Bridge

Act as a bridge between technical possibilities and operational realities. Translate AI capabilities for policymakers and policy constraints for technologists.

Practical Steps This Week

Experiment

Spend 30 minutes using AI on a real work task. Draft a briefing note, analyze stakeholder feedback, or stress-test a proposal. Experience both capabilities and limitations firsthand.

Audit Your Role

List your typical tasks. Which involve routine cognitive work that AI could automate? Which require judgment, context, or relationships? Understanding task composition is the first step to adaptation.

Have the Conversation

Talk openly with your team about AI. What are the fears? What are the opportunities? What would be a meaningful quick win? Psychological safety enables productive adaptation.

Champion a Pilot

Identify one low-risk, high-visibility use case in your organization. Something concrete, achievable, and demonstrable. Success breeds momentum for broader adoption.

Stay Informed

The AI landscape changes monthly, not yearly. Follow developments, but focus on implications for your work rather than technical details. Curate sources that provide strategic context.

The AI future is unwritten. It is not determined by the algorithm, but by the institutions in this room—and the individuals who lead them.

Will we repeat the mistakes of globalization?

Or will we build the adoption architecture for shared benefit?