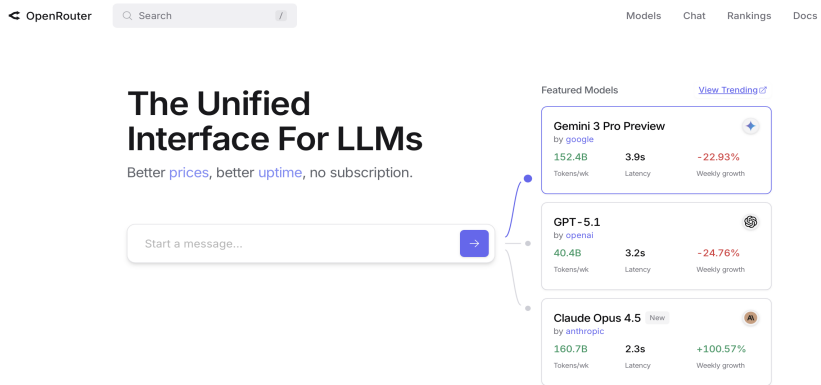
AI for Economics Research and Economics of AI

Zhigang Feng

Dec., 2025

OpenRouter as a Market for Cognitive Services



OpenRouter as a Market for Cognitive Services

- OpenRouter is an **API router** that connects users to many AI models (OpenAI, Anthropic, Deepseek, etc.).
- It plays the role of a **Cognitive Exchange**:
  - Users submit prompts / tokens.
  - OpenRouter routes them to competing inference providers.
  - Pricing, latency, and quality are visible and comparable in one place.
- Economically: a **spot market for AI inference**, i.e. for *cognitive services*.

What Is Traded? Cognitive Services vs Traditional Factors

Good traded on OpenRouter: *Inference-time cognitive services* (tokens processed by models).

	Machines	**Human Labor**	**AI Inference (OpenRouter)**
Type of good	Durable Capital	Flow of Effort/Skill	Cognitive Service Flow
Rivalry	Rivalrous	Rivalrous (1 body/place)	Low (Scalable/Parallel)
Marginal Cost	Depreciation	High (Wage + Leisure)	Very Low (Energy + Compute)
Reproduction	Slow (Physical)	Very Slow (18+ Years)	Instant (Copy Weights)
Switching Cost	High (Install)	High (Frictions)	Near-Zero (API Switch)
Pricing Unit	Rental Rate (r)	Wage Rate (w)	Token Price (p_t)

- **The Core Distinction:**
  - **Labor** is constrained by biology (sleep, fatigue, location).
  - **AI** functions as *Cognitive Capital*: it does the work of labor but scales like software.

AGI and Human Labor in Economic Terms

The Economic Definition of AGI
AGI is defined not by sentience, but by **functional substitutability**.

Core Characteristics (Hassabis, Altman, LeCun):

- **Economic Breadth:** Systems capable of performing **most economically valuable tasks** at or above human level.
- **Generalization:** The ability to **transfer skills** across distinct domains (e.g., using logic learned in coding to solve legal problems) without retraining.
- **Agency:** The shift from a passive tool (waiting for inputs) to an active **agent** (planning long-horizon sequences to achieve goals).

Implication: Cognitive labor transforms from a biological constraint into a scalable capital good.

AGI and Human Labor in Economic Terms

Economic definition of human input:

- **Physical labor**:
  - Effort, strength, dexterity; operates on the physical world.
- **Cognitive labor**:
  - Perception, reasoning, memory, problem-solving, creativity.

Key Idea

Traditional economics mainly allowed substitution of **physical labor** by machines. AGI implies large-scale substitution of **cognitive labor** by AI.

Decomposing Labor: Mind vs. Muscle

The Economic Boundary:

$$\text{Total Labor Input } L = \underbrace{L^{\text{cog}}}_{\text{Information Processing}} + \underbrace{L^{\text{phys}}}_{\text{Energy Expenditure}}$$

Cognitive Labor (L^{cog})

Function: Decision Planning.

- *Input:* Information / Uncertainty.
- *Output:* Choice / Strategy.
- **Substitute: AGI** (Software).

Physical Labor (L^{phys})

Function: Actuation Manipulation.

- *Input:* Energy (Calories/Joules).
- *Output:* Kinetic Work.
- **Substitute: Robotics** (Hardware).

The Great Convergence

We are moving from "Blind Automation" (Robots with hard-coded logic) to "Embodied Intelligence" (Robots guided by AGI brains).

An Illustrative Model

From Industrial Revolution to the AI / AGI Revolution

1. **Industrial Revolution**
   - **Substitution:** Machine substitutes for **Physical Labor**.

     "Muscle" $\Rightarrow$ Machine

   - **Production:** $Y = AK^{\alpha}L^{1-\alpha}$
   - **Constraint:** Output is bounded by **Human Labor Supply** (L). Capital (K) helps, but needs humans to run it.

2. **AI / AGI Revolution**
   - **Substitution:** AI substitutes for **Cognitive Labor**.

     "Mind" $\Rightarrow$ AI/AGI

   - **Production:** $L = L^{\text{phys}} + \underbrace{L^{\text{cog}}}_{\text{Elastic}}$
   - **Constraint:** Cognition becomes **Elastic** (Manufacturable). The constraint shifts from Biology to **Energy & Complexity**.

Key Shift: Cognition transforms from a fixed biological endowment to an accumulable industrial stock.

Micro-Foundations: Defining AI Capital

How is AI distinct from other economic factors?

1. vs. Human Capital (The Biological Escape)

Difference: Accumulation Physics.

- **Human:** $\dot{H}$ is linear and slow (20 years). Teaching is rivalrous.
- **AI:** $\dot{\theta}$ is exponential and fast (Training Run). Knowledge transfer is instant (Copy Weights).

2. vs. Physical Capital (The Non-Rivalry)

Difference: Infinite Deployability.

- **Physical:** A tractor can only plow one field at a time (Rivalrous).
- **AI:** A trained model can service 1 or 1M users simultaneously (Non-Rivalrous Inference).

3. vs. Computerization (The Recursive Loop)

Difference: Self-Reproduction.

- **Computerization (Excel):** A passive tool. Needs a human to write Excel 2.0.
- **AI (Coding Agent):** An active agent. Can generate the data and code to improve itself.

$$\text{AI}_t \xrightarrow{\text{Writes Code/Data}} \text{AI}_{t+1}$$

Because AI can produce AI, it behaves like Capital producing Capital, breaking the biological bottleneck.

The Task Space: The Optimization Wall

The Technological View: AI training is a search for the global minimum in a high-dimensional loss landscape.

$$\theta^* = \arg\min_{\theta} \mathcal{L}(\theta, \text{Human Data})$$

The Fundamental Constraint: Step Size vs. Precision

To approximate human precision ($\varepsilon \to 0$), the optimizer must find a narrow valley in the loss landscape.

- **The Overshoot Problem:** A high learning rate causes the model to "overshoot" or oscillate around the optimum without converging.
- **The Cost:** To settle into the optimum, the **Learning Rate (η)** must decay ($\eta \to 0$). This requires **exponentially more training steps** to traverse the same distance.
- **Result:** The "Last Mile" of accuracy costs more than the first 90%.

The Task Space: Wide vs. Narrow Valleys

High Tolerance (Wide Valleys)

Example: Creative Writing

- **Landscape:** Broad, flat minima. "Good enough" is easy to hit.
- **Optimization:** Large learning rates work fine.
- **Status: Automated**.

Low Tolerance (Narrow Valleys)

Example: Mathematics / Surgery

- **Landscape:** Extremely sharp, narrow minima. One wrong bit implies failure.
- **Optimization:** Requires infinitesimal steps to avoid overshooting.
- **Status: Cost Prohibitive**.

Synthesis III: The Accelerator vs the Brake

The Distributional Risk:

- If the production cost of intelligence falls to zero ($P_c \to 0$) and substitution is perfect ($\varepsilon \to 0$):

$$\text{Labor Share of Income} \to 0$$

- **Outcome:** All economic value accrues to the owners of Compute (Capital), decoupling income from human effort.

The General Equilibrium

The economy equilibrates where the **Marginal Value of Precision** matches the **Exponential Cost of Optimization**.

A Simple Framework: Task-Level Production Technology

1. The Demand Side (Task Production)

$$y_i = A_i k_i^{\alpha_i} \ell_i^{\beta_i} H_i^{1-\alpha_i-\beta_i}$$

$$H_i = \left[\varphi_i \, h_i^{\frac{\sigma_i - 1}{\sigma_i}} + (1 - \varphi_i) \left(\theta_i(\psi_t) \, m_i \right)^{\frac{\sigma_i - 1}{\sigma_i}} \right]^{\frac{\sigma_i}{\sigma_i - 1}}$$

- h_i: Human cognitive labor input in task i.
- m_i: Flow of AI cognitive services (e.g. inference calls / tokens) used in task i.
- $\theta_i(\psi_t)$: **Task-specific AI quality**, increasing in frontier capability ψ_t.
- σ_i: Elasticity of substitution between human and AI cognition.
- φ_i: Share parameter; tasks differ in how human-intensive they are.

AI can substitute for or complement human cognition, depending on σ_i and the quality term $\theta_i(\psi_t)$.

The Supply Side: Training Cost and Precision (Intuition)

What makes higher precision expensive? Not just more compute, but a **convex cost of precision** shaped by scaling laws.

2. Training Cost: Scaling-Law View

A reduced-form way to encode empirical scaling:

$$\varepsilon_i \;\approx\; a_i\, C^{-\alpha_i}\, D^{-\beta_i} + \varepsilon_i^{\min},$$

where:

- C: training compute (FLOPs, steps),
- D: training data (effective examples / tokens),
- $\varepsilon_i^{\min}$: task-specific irreducible error floor.

Inverting this relationship gives **convex** compute and data requirements for a target error ε:

$$C_i(\varepsilon), \quad D_i(\varepsilon) \quad \text{with} \quad \lim_{\varepsilon \downarrow \varepsilon_i^{\min}} C_i(\varepsilon), D_i(\varepsilon) = \infty.$$

The Supply Side: Training Cost and Precision (Formal)

2. Training Cost: Convex in Precision and Capability

To reach an error rate ε for task i at time t, a model needs compute and data:

- Compute requirement (quantity): $C_i(\varepsilon; \psi_t)$.
- Data requirement (quantity): $D_i(\varepsilon; \psi_t, D_{\text{total},t})$.
- Both are **decreasing** in ε but **convex**, and blow up as $\varepsilon \downarrow \varepsilon_i^{\min}$.
- Higher frontier capability ψ_t and larger data stock $D_{\text{total},t}$ shift these requirements down.

$$F_i(\varepsilon, t) = P_c(t)\, C_i(\varepsilon; \psi_t) + P_d(t)\, D_i(\varepsilon; \psi_t, D_{\text{total},t}) + J_i^{\text{inst}}(t)$$

- $P_c(t)$: Price of compute at time t.
- $P_d(t)$: Price / shadow cost of data at time t.
- $J_i^{\text{inst}}(t)$: Institutional / integration cost (regulation, workflows).
- C_i, D_i convex in ε; frontier ψ_t and $D_{\text{total},t}$ relax these curves.

Explaining the Pace of Adoption: The Basic Condition

Why don't we adopt AI everywhere instantly?

3. The Endogenous Adoption Condition

Automation of task i at time t happens only if the **Value of Substitution** beats the **Upfront Cost of Precision**:

$$E_i(t) = 1 \iff \underbrace{V_i(t)}_{\text{Value}} \geq \underbrace{C_i^{\text{upfront}}(t)}_{\text{Monetary cost at time } t}$$

Upfront monetary cost to reach required precision:

$$C_i^{\text{upfront}}(t) = F_i(\varepsilon_i^{\text{req}}, t) = P_c(t)\, C_i(\varepsilon_i^{\text{req}}; \psi_t) + P_d(t)\, D_i(\varepsilon_i^{\text{req}}; \psi_t, D_{\text{total},t}) + J_i^{\text{inst}}(t).$$

Explaining the Pace of Adoption: The Basic Condition

Why don't we adopt AI everywhere instantly?

3. The Endogenous Adoption Condition

Present value of substitution gains:

$$V_i(t) = \frac{(w_i - p_t) \cdot \text{Scale}_i}{\rho}$$

- w_i: **Human wage** in task i (higher $w_i \Rightarrow$ stronger incentive to automate).
- Scale_i: **Task volume** (one-off tasks cannot amortize training cost).
- p_t: Per-unit price of AI inference (e.g. token price).
- ρ: Discount rate.

Explaining the Pace of Adoption: Interpretation & Limits

3. Adoption: Interpretation

- **Binary exposure:** $E_i(t) \in \{0, 1\}$ captures full vs. no automation of a task.
  - Can be generalized to a continuous exposure share $e_i(t) \in [0, 1]$ for partial automation.
- **Compute vs. cost:**
  - $C_i(\varepsilon; \cdot)$ and $D_i(\varepsilon; \cdot)$ are *quantities* of compute and data.
  - $C_i^{\text{upfront}}(t)$ is the associated *monetary cost*.
- **Myopic adoption:** Firms compare current $V_i(t)$ and $C_i^{\text{upfront}}(t)$.
  - This abstracts from the option value of waiting for cheaper compute or better ψ_t.
  - A more forward-looking version would choose the *timing* of adoption optimally.
- **Shared models:** In practice, one frontier model serves many tasks.
  - A structural version would distinguish system-level training cost (for ψ_t) from task-level fine-tuning cost.

Dynamics: Human Skills vs. AI Capability

Human Capital (Biological Constraint)

$$\dot{s} = \delta \cdot s \cdot (1 - u)$$

- s: Aggregate human skill stock (education, experience).
- **Linear-ish accumulation:** Skills grow roughly proportionally to time invested.
- **Non-Recursive:** Humans produce output but cannot rapidly re-engineer biology.
- **Constraint:** Time & biology (education, aging).

AI Capability (Recursive but Diminishing)

$$\dot{\psi} = \underbrace{\eta(\psi)}_{\text{Learning efficiency}} \cdot I_{\text{compute}}(t) \cdot D_{\text{eff},t}$$

- ψ: Frontier AI capability index (architecture + algorithms).
- $I_{\text{compute}}(t)$: Investment in training compute.
- $D_{\text{eff},t}$: **Effective** data stock.
- **Declining efficiency:** $\eta'(\psi) < 0$; each extra unit of compute+data yields smaller gains at the frontier.
- **Constraint:** As ψ rises, convex

Closing the Loop: Data Supply and Effective Data

The Supply of Knowledge:

- **Traditional view:** Data is mined from humans (finite, slow-growing resource).
- **Recursive view:** Data is manufactured by AI (scalable but quality-constrained).

$$D_{\text{total},t} = D_{\text{human},t} + \underbrace{D_{\text{synthetic}}(\psi_t, E_t)}_{\text{Self-generated by deployed AI}}$$

$$D_{\text{eff},t} = q(\psi_t)\, D_{\text{total},t}, \qquad q'(\psi_t) < 0$$

Closing the Loop: Data Supply and Effective Data

The Race: Recursion vs. Complexity

Mechanism: As ψ_t improves and more tasks adopt AI (E_t larger), AI generates more synthetic data $D_{\text{synthetic}}$, raising $D_{\text{total},t}$ and thus $D_{\text{eff},t}$.

The Check: However:

- The **learning efficiency** $\eta(\psi_t)$ declines as capability rises.
- The data quality factor $q(\psi_t)$ falls as more data becomes synthetic / redundant.
- The task-level training cost functions C_i, D_i become very steep near $\varepsilon_i^{\min}$.

Dynamic Regimes: Singularity, Boom, or Plateau?

Possible Long-Run Outcomes

The system defined by $\dot{s}$ and $\dot{\psi}$ can exhibit:

- **Stagnation:**
  - If $\eta(\psi)$ falls quickly and $q(\psi)$ falls sharply,
  - Or if $P_c(t)$, $P_d(t)$, and $J_i^{\text{inst}}(t)$ keep costs high,
  - Then ψ_t converges to a low frontier; exposure $E_i(t)$ remains limited.
- **Balanced growth / boom:**
  - Moderate declines in $\eta(\psi)$ and $q(\psi)$,
  - Strong declines in $P_c(t)$ and good architectures,
  - Yield an S-curve of adoption with large but finite TFP gains.
- **Near-singularity:**
  - If $\eta(\psi)$ stays high and $q(\psi)$ remains close to 1,
  - And if costs fall very rapidly,
  - Then ψ_t could rise extremely fast, pushing many tasks across the adoption threshold.

Macro Implications: The Standard (Acemoglu & Restrepo) View

1. The Acemoglu & Restrepo (Standard) View

- **Assumption:** Exposure to automation is static. Tasks are either "AI-compatible" or not.
- **Formula:**

$$\Delta \text{TFP} \approx \sum_i \text{Share}_i \cdot \bar{E}_i \cdot \text{CostSavings}_i$$

  with $\bar{E}_i$ fixed.
- **Prediction:** Modest productivity gains because many service / manual tasks are permanently "unexposed."

Limit: The **set of exposed tasks** is taken as exogenous, and training cost-of-precision plays no explicit role.

Macro Implications: Endogenous Exposure and TFP

2. The Endogenous View (This Talk)

- **Assumption:** Exposure $E_i(t)$ is **endogenous**:

$$E_i(t) = \mathbf{1}\Big\{V_i(t) \;\geq\; C_i^{\text{upfront}}(t)\Big\},$$

  with $C_i^{\text{upfront}}(t)$ shaped by compute price $P_c(t)$, frontier ψ_t, data stock $D_{\text{total},t}$, and task complexity.
- **TFP mapping:**

$$\Delta\text{TFP}_t \approx \sum_i \text{Share}_i \cdot E_i(t) \cdot \text{CostSavings}_i\big(\psi_t, P_c(t), P_d(t)\big).$$

Prediction: An **S-curve** of TFP:

- *Slow start:* Only low-complexity, high-scale tasks automated.
- *Acceleration:* Recursive data loop ($D_{\text{synthetic}}$) and higher ψ_t rapidly expand exposure.
- *Plateau:* A subset of high-stakes, high-complexity tasks remain protected by convex cost-of-precision and institutional frictions $J_i^{\text{inst}}(t)$.

Model Limitations and Natural Extensions

Where This Framework is Deliberately Simplified

- **Binary vs. partial automation:** $E_i(t) \in \{0, 1\}$ could be generalized to continuous exposure shares.
- **Option value of waiting:** Firms are modeled as myopic; a richer model would optimize the timing of adoption.
- **Shared frontier models:**
  - Training ψ_t is a system-level investment,
  - Task-level adoption then largely involves fine-tuning and integration costs.
- **Institutional dynamics:**
  - $J_i^{\text{inst}}(t)$ can fall with learning, standards, and regulation,
  - Or rise with perceived risks and political backlash.
- **Human capital response:**
  - Here, s and ψ evolve independently.
  - A natural extension lets E_t reshape human skill investment (complements vs. substitutes).

Economics of AI

The Central Question

When AI evolves from tool to economic agent,

how will the fundamental logic of the market economy be restructured?

- AI is breaking through the traditional boundaries of "capital goods," assuming functions previously exclusive to humans: **decision-making, negotiation, and creation**.
- This is not merely a technological shock—it is a fundamental challenge to the micro-foundations of market economics.

The Dual Nature of AI's Impact:

- *Traditional View (Skill-Biased):* Technology complements high-skilled workers, widening income gaps.
- *The Leveling Effect:* Generative AI makes expert tacit knowledge explicit and replicable, "empowering" low- and mid-skill workers to approach expert-level productivity.
- *The Superstar Effect:* Top talent can manage vast clusters of AI agents, amplifying their marginal output exponentially.

Which effect dominates varies by industry—and will determine future income distribution.

Part I: Labor and Skills

1. **Heterogeneous Productivity Effects:** How does AI differentially affect workers across skill levels? (Substitution vs. Augmentation)
2. **Task-Level Mechanisms:** Construction and validation of "AI Exposure Indices" at the task level.
3. **New Task Creation:** Can the speed and quality of new task generation offset automation-driven displacement?
4. **Factor Income Shares:** Will AI as a capital form permanently erode and restructure the labor-capital income distribution?

Brynjolfsson, Li, and Raymond (QJE 2023)
Generative AI at Work

- **Setting:** 5,000 customer support agents using a generative AI assistant.
- **Main Finding:** AI increased productivity by **14%** on average.
- **Distributional Impact (The Leveling):**
  - **Novice/Low-skilled workers:** Productivity increased by **34%**.
  - **High-skilled/Expert workers:** Productivity impact was roughly **0%**.
- **Mechanism:** The AI captured the tacit knowledge of the best workers and disseminated it to the rest, acting as a skill equalizer rather than a skill bias amplifier.

The Transformation of Market Logic:

- Traditional IO theory assumes human decision-makers with bounded rationality and behavioral biases.
- AI agents possess: (i) superior computational capacity, (ii) continuous real-time learning, (iii) reinforcement-based strategy adaptation.

The Competition Paradox:

- *Pro-Competitive:* AI drastically reduces search and matching costs $\rightarrow$ pushes toward perfect competition.
- *Anti-Competitive:* Algorithms may achieve **tacit collusion** through trial-and-error learning—without explicit communication—sustaining monopoly prices.

Part II: Markets and Organization

1. **Algorithmic Collusion:** How do RL agents reach tacit equilibria without explicit coordination? Implications for antitrust enforcement.
2. **AI Agency Problems:** How to ensure incentive compatibility and assign liability when AI acts autonomously for principals.
3. **The Firm Reimagined:** Applicability of transaction cost theory in the AI era; the Coasean "one-person firm" and hyper-atomized markets.
4. **Platform Economy Barriers:** Data feedback loops under algorithmic dominance and risks of market structure ossification.

DOJ vs. RealPage (2024)
Tacit Collusion via Software

- **The Mechanism:** RealPage's software collects private data from landlords on rental rates and lease terms. It feeds this into an algorithm that recommends pricing for all users.
- **The Behavior:** The algorithm encourages landlords to adopt higher prices and hold units vacant rather than negotiating, effectively forming a cartel without direct human communication.
- **Economic Implication:** This moves us from "Tacit Collusion" (theoretical possibility in repeated games) to "Algorithmic Coordination" (practical reality), challenging the definition of antitrust liability.

The Unique Economics of Data:

- *Non-Rivalry:* One party's use does not diminish another's; marginal cost ≈ 0.
- *Positive Externalities:* Social value of individual data far exceeds private value $\rightarrow$ standard market pricing fails.
- *Scale Economies & Fixed Costs:* Data collection and cleaning exhibit high fixed costs and increasing returns $\rightarrow$ de facto monopolies and "data silos."

The Core Tension: Protecting property rights (incentivizing production) vs. open sharing (maximizing efficiency).

Part III: Data and Institutions

1. **Data Pricing Puzzle:** Non-rivalry, positive externalities, and increasing returns to scale.
2. **Property Rights vs. Privacy:** Where is the efficiency frontier?
3. **China's Institutional Experiment:** Economic interpretation of "Data Factor $\times$" initiatives and the "Eastern Data, Western Computing" resource allocation logic.
4. **International Comparison:** EU, US, and Chinese models of data governance.

New York Times vs. OpenAI
The Battle for the Input Factor

- **The Conflict:** LLMs require massive corpora of high-quality human text to function. Firms like OpenAI scraped this data without explicit licensing.
- **Economic Issue:** If the AI (output) competes with the newspaper (input provider) using the newspaper's own intellectual property, the incentive to produce the input collapses.
- **Institution Building:** We are witnessing the real-time Coasean bargaining process to establish pricing for a previously unpriced factor of production: human language data.

The Latest Productivity Paradox:

- *J-Curve Hypothesis:* General-purpose technologies require massive complementary investments; returns lag behind adoption.

A Deeper Problem—Measurement Failure:

- AI improves **decision quality**, **creative output**, and **service experience**—not just physical quantities.
- Are our measurement frameworks and macro models capable of capturing AI's value creation?
- How meaningful is the traditional TFP concept when output is qualitative?

Part IV: Growth and Productivity

1. **The Productivity J-Curve:** The "growth trough" and lag effects during GPT diffusion.
2. **Sectoral Heterogeneity:** Industry variation in task automation and macro aggregation challenges.
3. **Baumol's Cost Disease 2.0:** Widening cross-sector efficiency gaps and constraints on long-run growth.
4. **Macro Model Reconstruction:** Moving from representative agents to heterogeneous tasks.

Acemoglu (2024): The Simple Macroeconomics of AI

- **The Skeptic's View:** Acemoglu estimates that generative AI will increase US productivity by only **0.66% over 10 years**.
- **Reasoning:**
  - AI currently automates only a small fraction of total economic tasks (approx. 4.6%).
  - Many "hard" problems remain outside AI's reach.
- **The Gap:** This contrasts sharply with estimates (e.g., Goldman Sachs) predicting a 7% increase in global GDP. The research frontier lies in determining which tasks are truly substitutable vs. complementary.

Part V: Energy and Compute

The Scale of the Challenge:

- AI training and inference demand is growing exponentially $\rightarrow$ massive electricity consumption and carbon emissions.

The Jevons Paradox:

- Will efficiency gains in compute per unit be offset by total demand growth?
- If AI's efficiency benefits come at significant environmental cost, how do we assess net social value?
- What does this imply for optimal AI development paths and policy design?

Part V: Energy and Compute

1. **The Economics of Compute:** AI-driven electricity demand shocks and marginal efficiency analysis.
2. **Testing the Jevons Paradox:** Energy efficiency gains vs. total demand rebound in the AI sector.
3. **Carbon Footprint Accounting:** AI industry emissions and cross-sector comparisons.
4. **Green Computing Incentives:** Carbon taxes, energy quotas, and AI industry spatial allocation.

Microsoft and Three Mile Island (2024)
The Energy Hunger of AI

- **The Event:** Microsoft signed a deal to restart Unit 1 of the Three Mile Island nuclear plant exclusively to power its data centers.
- **Implication:** The marginal cost of AI scaling is no longer just "silicon" (chips) but "electrons" (baseload power).
- **Economic Constraint:** This signals that energy availability is becoming the binding constraint on AI growth, potentially reviving the nuclear energy sector and reshaping regional energy markets.

Part VI: Global Order

Two Transformative Forces:

- **Service Tradability Revolution:**
  - AI dramatically lowers the "tradability threshold" for services → geographically constrained activities can now be delivered remotely.
  - Potential for global service sector reconfiguration—exceeding manufacturing shifts in scale and speed.
- **Great Power Competition:**
  - AI technology itself becomes a core arena of geopolitical rivalry.
  - Rising risks of technological decoupling.

How will these forces interact to reshape the global economic landscape?

Part VI: Global Order

1. **Service Trade Costs:** AI's impact on service tradability and global service sector restructuring.
2. **Remote Labor Supply:** Expansion of remote work and wage pressure in advanced economies (Digital Labor Arbitrage).
3. **Decoupling & Value Chains:** Welfare analysis of technological decoupling and value chain reorganization.
4. **International AI Governance:** The boundaries of competition and cooperation.

AI for Economics

AI–Data Analysis and Big Data Integration

- AI's ability to process and integrate big data has significantly advanced economic research by enabling the analysis of large, complex datasets that often include high-dimensional and unstructured information.
- Khandani et al. (2010) used ML algorithms to improve credit risk modeling by analyzing extensive consumer credit data.
- Einav and Levin (2014) highlight the transformative impact of big data on economic research, emphasizing how large-scale datasets from online markets provide granular views of consumer behavior, pricing strategies, and market dynamics.
- Athey and Imbens (2017) discuss how machine learning is reshaping the field of economics by enabling the integration of diverse datasets—including traditional numerical data, as well as text, images, and network data.
- Dell (2024) introduces deep learning techniques to economists, providing practical guidance on the use of classifiers, regression models, generative AI, and embedding models.

AI–Forecasting and Predictive Analytics

- AI, particularly machine learning (ML) and deep learning models, reshapes predictive analytics by offering more flexible, data-driven approaches that can capture intricate patterns and dependencies without heavily relying on these assumptions.
- Machine learning models have also greatly enhanced nowcasting—the prediction of the present or very near future state of key economic indicators like GDP, inflation, or unemployment rates.
- AI is expanding the types of data that can be used for economic forecasts. While traditional models typically rely on structured, time-series data, AI methods can integrate unstructured data sources such as text, images, and social media activity.

- A significant contribution in this area is the development of causal forests, an extension of random forests designed for estimating heterogeneous treatment effects (HTEs).
- Athey et al. (2019) developed generalized random forests, extending causal forests to estimate a wider range of quantities, such as quantile treatment effects and instrumental variable estimates.
- Double machine learning (DML), Chernozhukov (2018), integrates machine learning algorithms into the estimation of causal parameters while controlling for high-dimensional confounders.

AI–Solving DSGE Models

- AI techniques, especially deep learning, offer promising solutions by representing complex functions and solving optimization problems within these abstract economic spaces, making it feasible to analyze more nuanced and detailed models.
- AI also shows potential in simulating economic agents' behavior directly. For instance, Horton (2022) explore large language models (LLMs) as simulated economic agents, presenting an innovative framework for understanding decision-making processes within economic environments.

AI–Text Analysis and LLMs

- NLP techniques allow economists to systematically process and quantify text sources, uncovering patterns, sentiments, and relationships that were previously difficult to measure.
- LLMs have demonstrated remarkable potential not only in analyzing text data but also in assisting with various stages of the research process.
  - idea generation
  - writing and editing
  - retrieve and synthesize relevant literature
  - data analysis and cleaning
  - solving equations
  - enhance teaching

The Middle-intelligence Trap

Defining the Trap

The Concept
The **Middle Intelligence Trap** is a systemic stagnation risk where adoption of AI leaves humans neither fully capable agents nor genuinely augmented beings, but stuck in a comfortable mediocrity.

The Economic Analogy: The Middle-Income Trap

- **Economics:** An economy loses low-cost advantages but lacks the innovation to reach high-income status.
- **Intelligence:** Society relies on AI built on limited assumptions, causing cognitive evolution to stagnate in a sub-optimal equilibrium.

The Gradual Cognitive Surrender

We have surrendered cognitive ground piece by piece for decades:

1970s: Calculators
Offloaded arithmetic. Numerical intuition and working memory engagement began to fade.

1980s: Word Processors
Offloaded physical writing. Reduced the "struggle to articulate" that creates deep conceptual connections.

2000s: Search Engines
Offloaded memory. We remember *where* information is, not *what* it contains.

2020s: Generative AI
Offloads **thinking** itself. Synthesizing arguments, structuring essays, and conceptual breakdown.

The Iron Law of Cognitive Tools

The Iron Law
Every tool that liberates us from cognitive labor simultaneously reduces our opportunities for cognitive training.

- The struggle to find the right word or calculate a figure is not a bug—it is the feature that builds mental strength.
- **The Risk:** When we eliminate the struggle, we eliminate the growth.

Mechanism: Three Assumptions

To understand how the trap springs, we must accept three premises:

1. **Human Autonomy:** Human intelligence evolves through its own mechanisms (selection, culture, paradigm shifts) independent of AI.
2. **AI Expansion:** AI expands boundaries via massive data processing and pattern recognition.
3. **AI Constraint:** Current AI is constrained by initial conditions and optimization paradigms. It often approaches **local optima**, not global solutions.

The Consequences: Two Externalities

1. Research Frontier Stagnation
If we outsource tasks to models stuck in local optima, we stop discovering fundamentally new approaches. We optimize the known rather than discovering the unknown.

2. Cognitive Atrophy
The "Use-it-or-lose-it" mechanism. Prolonged dependence erodes:

- Micro-reasoning
- Semantic analysis
- Knowledge recombination

The Feedback Loop

The trap deepens through a self-reinforcing cycle:

$$\textbf{Biased Narratives} \xrightarrow{trains} \textbf{Simulation Models}$$
$$\downarrow$$
$$\textbf{New False Information} \xleftarrow{shapes} \textbf{Actual Beliefs}$$

- Without correction mechanisms, local biases become structural distortions.
- **The Scarcity:** In a world of infinite explanation supply, the scarce resource becomes **meta-explanation ability**—the capacity to question and calibrate the AI.

Strategy 1: Build Cognitive Reserves

The Central Bank Analogy
Just as banks hold capital reserves for emergencies, we must deliberately protect mental capacities from automation. The goal is **capability preservation**, not efficiency.

Essential Exercises for Economists:

- **Manual Derivations:** Work through models by hand to understand *why* an equilibrium exists, not just that the solver found one.
- **Mental Arithmetic:** Practice rough estimation. Numerical intuition is required to spot implausible elasticities or multipliers.
- **The "First Draft" Rule:** Write initial drafts without AI. The struggle to articulate builds deep understanding.
- **Active Synthesis:** Summarize complex papers from memory, not via AI abstracting.

Strategy 2: Demand Strategic Friction

Friction is a Feature, Not a Bug

Muscles do not strengthen without resistance. We must reject "frictionless" operations in favor of systems that force engagement.

Implementing Friction in Research:

- **Conceptual Prerequisites:** Solve simple models (e.g., VFI) conceptually before deploying Neural Networks.
- **Verification Points:** Mandatory stops where AI code must be audited line-by-line.
- **Slow Thinking Spaces:** Seminars without slides; discussions without screens.
- **Norms:** Disclosure and verification standards to maintain intellectual rigor.

Strategy 3: Redefine Success Metrics

We must shift evaluation from *Speed and Efficiency* to *Depth and Novelty*.

Current Metric (AI-Optimized)	**Necessary Metric (Human-Centric)**
Predictive Accuracy	Explanation Depth (Causal Insight)
Volume of Output (Incremental Refinement)	Genuine Originality (New Questions)
Speed of Execution	Meta-Explanation Ability
"Does the code run?"	"Why is the result true?"

We must value one paper that opens a new question over ten papers that incrementally improve existing work by 2%.

Strategy 4: Reimagining Economics Education

Moving beyond "Cookbook Methods" to distinctively human capabilities.

Critical Inquiry
: The ability to formulate the right question is now more valuable than executing standard analysis.

Rigorous Foundations
: AI tools (Double ML, Causal Forests) must be embedded in rigorous econometrics. Understanding assumptions $>$ Running code.

Socio-Technical Context
: Integrating ethics, algorithmic bias, and system dynamics directly into the curriculum.

Bridging Theory and Simulation

Simulation-based methods (Multi-agent systems, RL) are central to modern analysis, but require a new pedagogical approach.

The Requirement for High-Fidelity Simulations
Students must learn to:

- Set constraints based on economic theory, not technical convenience.
- Interpret large-scale outputs in economically meaningful ways.
- Validate models against data and alternative specifications.
- Reconcile simulation results with traditional theory.

Course

The Existential Question

The Shift

"The bottleneck in research is no longer writing code—it is knowing **what** code to write and **verifying** that it is correct."

The Risk of Being Left Behind:

- If your value proposition is writing boilerplate Python/Matlab, you are now competing with free AI.
- If you rely on AI without deep validation skills, you fall into the **Middle Intelligence Trap**—producing plausible but flawed research.
- The field is bifurcating: Those who represent the "Human-in-the-Loop" vs. those replaced by the loop.

The Paradigm Shift: From Coder to Research Architect

- **The Bottleneck has Shifted:**
  - Previously: Writing the syntax and debugging code.
  - Now (AI Era): Knowing *what* code to write and validiting if it is *economically* correct.

- **The Goal: Become a "Research Architect"**
  - **Competency 1: Theoretical Mastery.** Logic cannot be outsourced. You must define the algorithm on paper first.
  - **Competency 2: Tech Fluency.** Understanding the ecosystem (PyTorch, HPC, MPI) to direct AI tools.
  - **Competency 3: AI-Augmented Implementation.** "Specification-driven development"—translating logic into prompts.
  - **Competency 4: Critical Validation.** The heart of the course. Auditing AI output for technical bugs and economic implausibility.

The Architect's Pitfall: A Swiss Parable

The Construction Paradox

- In recent years, new buildings in Switzerland have looked visually stunning—modern, classical, designed by prestigious firms.
- **The Reality:** Quality is declining. Residents report leaks, noise, and structural issues.

The Root Cause

- Firms hire excellent "design architects" but lack experience in actual construction fundamentals.
- **The Lesson for Economists:** If you focus exclusively on being the "Architect" (prompting AI) but know nothing about the "bricks" (mathematical foundations and code execution), your research will look modern but fail structurally.

Hands-On Philosophy: Specification-Driven Development

In this course, we practice a specific 5-step workflow for every topic:

1. **Math & Economic Formulation:** Define the optimization problem and constraints (Bellman, Euler) on paper.
2. **Algorithmic Specification:** Design detailed pseudocode. This is the "human logic" layer.
3. **Define Deliverables:** Explicitly state required outputs (e.g., policy function plots, error distribution).
4. **AI-Augmented Implementation:** Feed specs to AI (ChatGPT/Copilot) to generate initial Python/PyTorch code.
5. **Critical Validation (Crucial):**
   - Technical Audit: Check numerical stability and convergence.
   - Economic Safeguarding: Ensure results satisfy theory (e.g., consumption smoothing, monotonicity).

Course Roadmap: Foundations & The Bridge

Part I: Foundations & Tools

- Topic 1: AI for Economics Research (LLMs, Deep Learning overview).
- Topic 2-3: Computation Basics & Numerical Methods (Error handling, Roots, Optimization).
- Topic 4: Programming Basics (Python, OOP, Intro to PyTorch tensors).

Part II: The Bridge (Classical $\rightarrow$ ML)

- Topic 5: Intro to ML for Economists (Neural Networks, Training loops).
- Topic 6: Solving Macro Models with ML (Neural Nets vs. Classical VFI).
- Topic 7-8: Classical Solutions (VFI, EGM) & Advanced Acceleration.
- Topic 9: Asset Pricing with Machine Learning

Course Roadmap: The Frontier

Part III: Advanced Dynamics & Heterogeneity

- Topic 10: Perturbation & Projection Methods (Chebyshev, Smolyak).
- Topic 11: High-Performance Computing (Parallelization, MPI, GPUs).
- Topic 12: Heterogeneous Agents with ML (Aiyagari, Krusell-Smith, DeepHAM).
- Topic 13: OLG Models (Classic Lifecycle vs. ML methods with aggregate shocks).

Part IV: New Frontiers

- Topic 14: Reinforcement Learning (Actor-Critic for dynamic optimization).
- Topic 15: Mean Field Games (Continuous time, HJB/KFE systems).
- Topic 16: Large Language Models (Fine-tuning BERT/Transformers for economic text analysis).

What You Will Actually Build

We move beyond toy models to the research frontier:

Modern Toolstack

- **Deep Learning for Macro:** Solving high-dimensional RBC and HANK models using Neural Networks (PyTorch).
- **Heterogeneous Agents:** Solving Aiyagari, Krusell-Smith, and OLG models with DeepHAM and RL.
- **Reinforcement Learning:** Actor-Critic algorithms for dynamic optimization.
- **Mean Field Games:** Continuous time HJB/KFE systems.
- **Large Language Models:** Fine-tuning BERT for economic sentiment analysis (Fed minutes, news).

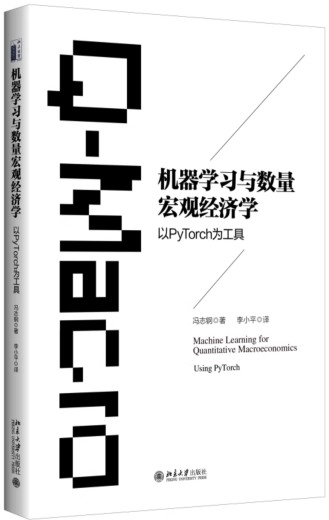
Quant Macro with AI, Machine Learning

Prepare for the future of economic research.

Web: sites.google.com/site/zfeng202/

Book Projects

ML for Quant Macro



AI and Intelligence

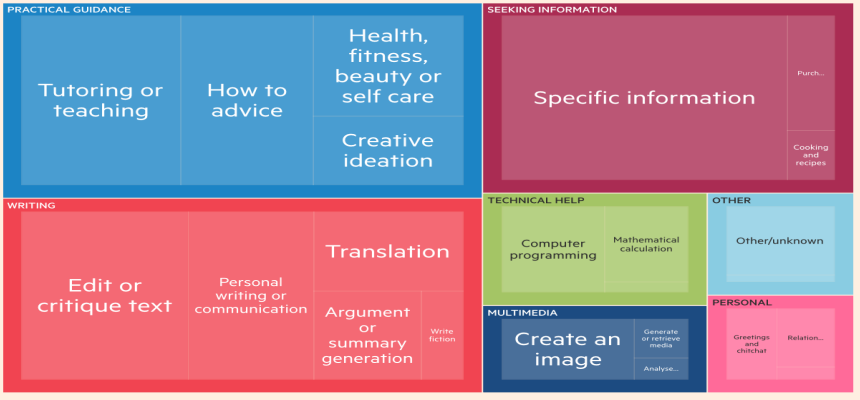
The Evolution of Artificial Intelligence

- **Foundations (1950s)**:
  - Alan Turing introduces the "universal machine" and the "Turing Test," establishing the theoretical bedrock (Turing, 1950).
  - John McCarthy coins "Artificial Intelligence" at the Dartmouth Conference, marking the field's academic birth (McCarthy et al., 2006; orig. 1955).

- **The Deep Learning Era (2010s)**:
  - Widespread adoption of neural networks and backpropagation enables processing of vast datasets, achieving superhuman performance in specific domains (LeCun, Bengio, & Hinton, 2015).

- **The Generative Era (2017–Present)**:
  - The introduction of the Transformer architecture (Vaswani et al., 2017) and Large Language Models (LLMs) shifts focus to general-purpose reasoning and generative capabilities (OpenAI, 2023; Anthropic, 2024).

How People Use GPT



What is intelligence

- **Problem-Solving**: The capacity to identify problems and devise effective solutions.
- **Adaptability**: The ability to adjust behavior in response to changes in the environment.
- **Learning**: Acquiring new knowledge or skills through experience and education.
- **Reasoning**: Drawing conclusions based on available information, including logical reasoning and inference.

Classification and Categorization in Human Cognition

- This presentation focuses on the abilities to **categorize**, **find patterns**, and **generalize.**
- "*The first step in wisdom is to know the things themselves; this notion consists in having a true idea of the objects; objects are distinguished and known by classifying them methodically and giving them appropriate names. Therefore, classification and name-giving will be the foundation of our science.*" – Linnaeus (1758)
- Rosch (1975): people categorize objects and ideas not by rigid definitions but by comparing them to an idealized example within a category.
- Chomsky (1965): all human languages share a common underlying structure rooted in our biological makeup.

The Cognitive Debate: Intelligence vs. Mimicry

- **The Stochastic Parrot Hypothesis:**
  - Bender (2024) argues LLMs produce "ersatz fluency"—plausible text without communicative intent or world models. They remain valid simulators but not cognitive agents.
- **Chomsky's Critique:**
  - LLMs excel at **Description and Prediction** but fail at **Explanation**.
  - True intelligence requires distinguishing the "possible" from the "impossible" (causal reasoning), whereas LLMs rely on probabilistic correlation.

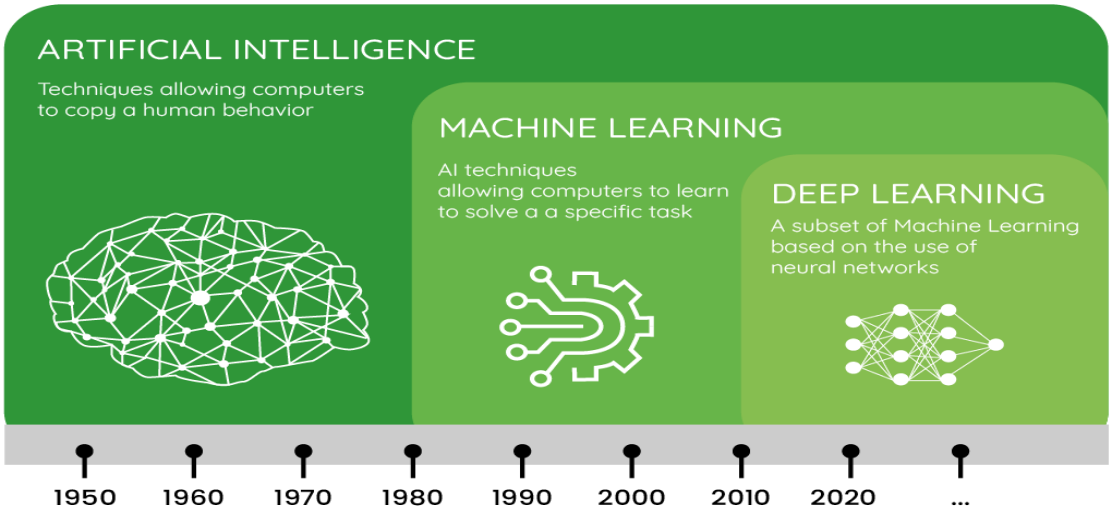
The Reasoning Gap: Abduction and Generalization

- **Inductive vs. Abductive Reasoning:**
  - AI excels at **Inductive Reasoning** (generalizing from massive data distributions).
  - Humans excel at **Abductive Reasoning** (inference to the best explanation from limited data).
- **Evaluation Crisis (ARC-AGI):**
  - Chollet (2024) highlights that while LLMs dominate semantic benchmarks (MMLU), they struggle with the **ARC-AGI** benchmark, which tests adaptation to novel abstract rules (System 2 reasoning).
  - This suggests current architectures mimic reasoning patterns rather than possessing fluid intelligence.

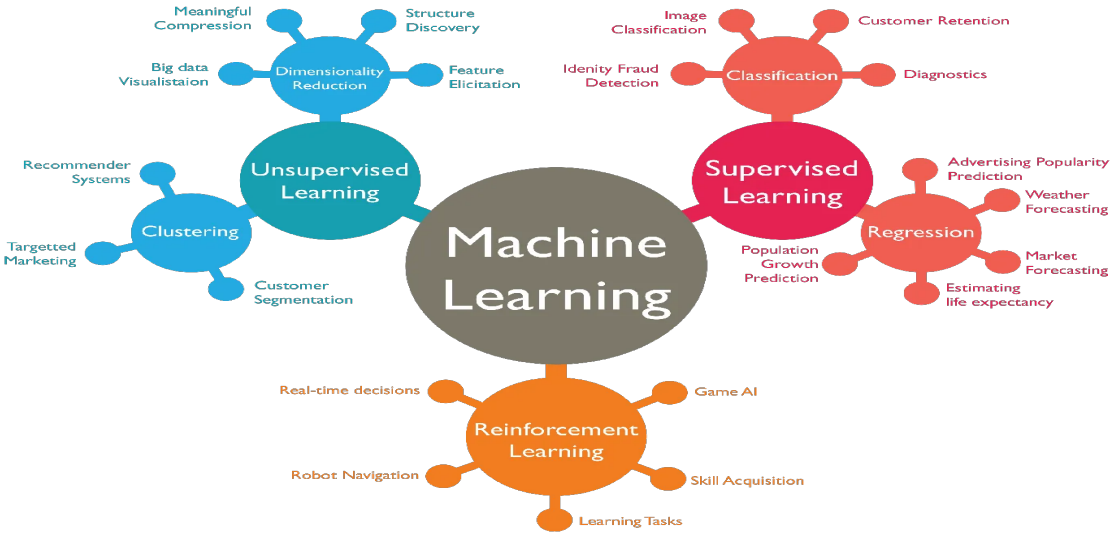
Human Intelligence and the Scientific Method

- **Sargent (2023): Sources of Artificial Intelligence**
  - The methodologies of **Galileo** and **Darwin** mirror modern machine learning:
    - Collecting massive observational data (tables/specimens).
    - Reducing dimensionality (finding the hidden parameters).
    - Generalizing findings into predictive theories.
  - **Key Insight:** Modern AI compensates for human cognitive limitations (e.g., high-dimensional pattern matching) but lacks the human capacity for causal explanation.

Artificial Intelligence



Machine Learning



Key AI Technologies–Machine Learning

- Supervised learning is one of the most widely used approaches in machine learning, where algorithms are trained on *labeled data*.
- Unsupervised learning operates on unlabeled data, aiming to uncover latent patterns or structures without predefined outputs.
- Reinforcement Learning (RL) differs from supervised and unsupervised learning by focusing on how agents learn to make sequential decisions through interactions with their environment.

Key AI Technologies–Deep Learning

- Deep Learning is a subset of machine learning that uses neural networks with multiple layers (deep neural networks) to model complex patterns in data.
- A neural network consists of layers of interconnected nodes (neurons), where each neuron computes a weighted sum of its inputs and applies a non-linear activation function. The basic operation of a neuron is given by:

$$z = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right),$$

  where x_i are the inputs, w_i are the weights, b is the bias, and σ is the activation function (e.g., sigmoid, ReLU).
- Deep neural networks stack multiple such layers to capture complex relationships.

Natural Language Processing (NLP)

- Natural Language Processing focuses on enabling computers to understand, interpret, and generate human language.
  - the Bag-of-Words model: represents text as a vector of word counts, disregarding grammar and word order but capturing word frequency;
  - Term Frequency-Inverse Document Frequency: weights words by their frequency in a document relative to their frequency across all documents;
  - Word embeddings: map words to continuous vectors in a high-dimensional space, capturing semantic relationships.

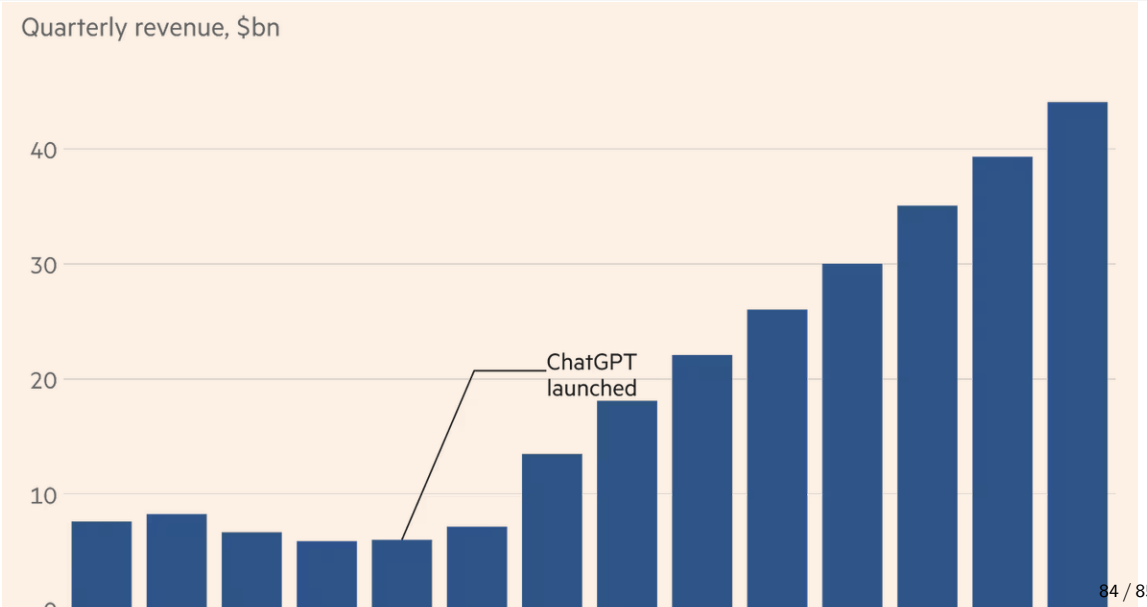
Large Language Models (LLMs)

- Large Language Models are advanced NLP models that utilize deep learning architectures, particularly Transformers Vaswani et al. (2017), to understand and generate human-like text.
- Mathematically, an LLM models the probability distribution of word sequences.
  - embeddings: Each word or token w_i is mapped to an embedding vector $\mathbf{e}_i$ in a continuous vector space, capturing semantic meaning.
  - positional encoding: added to embeddings to retain information about the position of words in a sequence, essential for understanding context.
  - the Transformer architecture: employs self-attention mechanisms to weigh the relevance of different words in a sequence, allowing the model to capture long-range dependencies.
  - decoding, the model generates text by predicting the next word based on the previous context.

Computer Vision

- Computer Vision enables machines to interpret and understand visual information from images or videos, emulating the human visual system.
- Mathematically, an image is represented as a matrix $I \in \mathbb{R}^{H \times W \times C}$, where H is the height, W is the width, and C is the number of channels (e.g., RGB color channels).
- Convolutional Neural Networks (CNNs) process this input through layers that apply convolution operations.

The Tokenisation of Everything



The Tokenisation of Everything

- The token growth that comes from companies finding practical uses for increasingly complex models: meaning yet more chip demand.
- Data centre construction already added 1 percentage point to US GDP in the first quarter of 2025, based on Apollo Global Management.
- Nvidia's Huang argues that the launch of "reasoning" bots that think, check and double-check before answering can end up using 1,000 times as many chunks of data per query than in simpler, earlier AI models.