Understanding AGI Futures – Toward a Shared Vocabulary for Policy Planning

Swaptik Chowdhury

The New York Data Science and AI Conference, 2025

What are we going to talk about?

AGI– An Introduction	Proposed Solution	Examples
• Conflicting Definitions and timelines for AGI. • Need for a common vocabulary for governance	• AGI Pivots as a framework for AGI Futures • Six Pivots for classifying Post AGI future.	• Stable vs Unstable Future • Aligned Hegemon vs Pandora

Artificial General Intelligence: What is that?

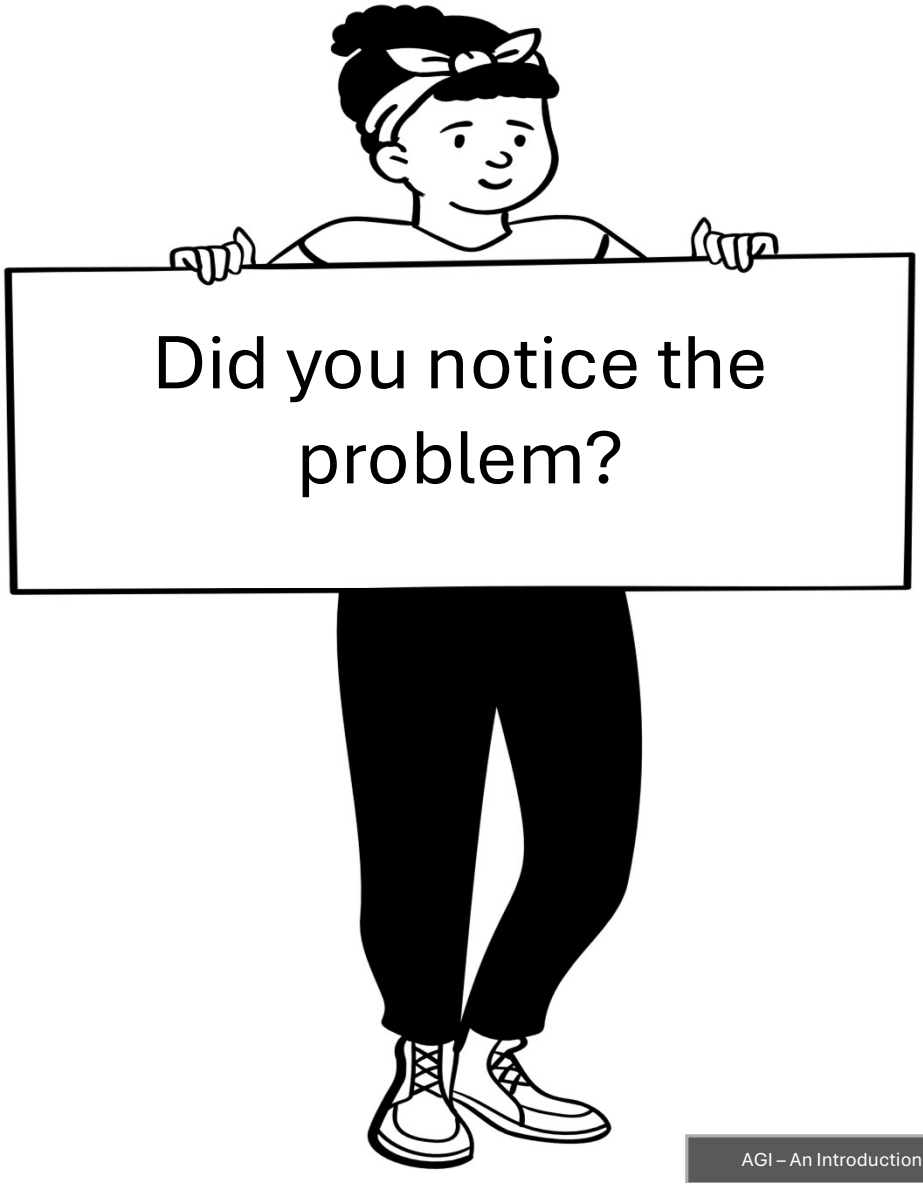
Definitions vary across scholars, researchers, and companies, showing no consensus on what AGI is.

- AI systems that rival or surpass the human brain in complexity and speed, that can acquire, manipulate and reason with general knowledge.
  - Mark Guburd(1997)
- We use AGI to refer to systems that demonstrate broad capabilities of intelligence, including reasoning, planning, and the ability to learn from experience, and with these capabilities at or above human-level.
  - Microsoft Research (2023)
- Artificial General Intelligence (AGI)... [is] used to describe an AI system that is at least as capable as a human at most tasks.
  - Google Deepmind (2024)
- Highly autonomous systems that outperform humans at most economically valuable work.
  – OpenAI Charter
- AI that can perform any intellectual task that a human can.
  - Andrew Ng (2024)

When is it coming?

Timelines for AGI range from ‘already here’ to decades away.

- Framing AGI operationally as the *“ability to pass human tests”*; says on that definition AGI could be ~5 years out .
  – Jensen Huang (NVIDIA 2024)
- I think we’re probably three to five years away
  – Demis Hassabis (Google Deepmind 2024)
- If you define AGI as smarter than the smartest human, [it’s] probably next year, within two years.
  - Elon Musk (2024)
- We are on course for AGI by 2027. These AI systems will basically be able to automate basically all cognitive jobs (think: all jobs that could be done remotely).
  – Leopold Aschenbrenner (2024)

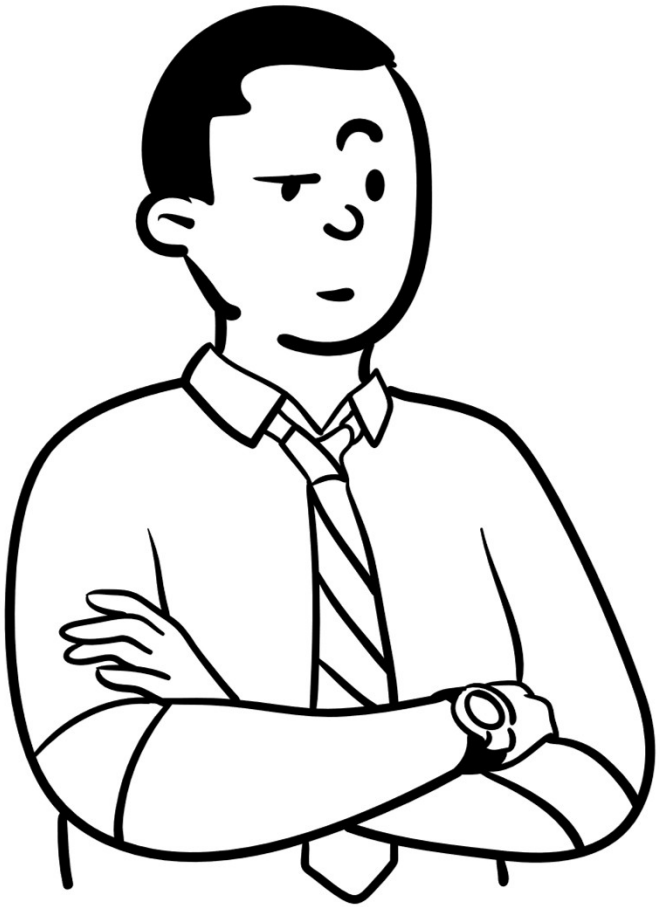

Did you notice the problem?

The Problem

Current AGI conceptualizations (definition and timelines) have :-

- **Scope Mismatch:**
  - Is AGI an agent with goals or a system with general tasks?
- **Timeline Dissonance:**
  - Predictions range from imminent to decades out, shaped by hype cycles.

So, What AGI is remains unsettled.



So what? Why should we care?

We Should Care Because..

- **Expanding portion of economy expect AGI to become real and significant resources are being deployed.**
  - LLMs alone could impact ~80% of U.S. workforce tasks, with 15–56% of tasks being accelerated significantly through AI-powered tools.
  - CEO Mark Zuckerberg is personally recruiting ~50 AI experts & OpenAI co-founder Ilya Sutskever's SSI startup has raised a total of $3 billion, with a $32 billion valuation
- **Given the scale of investment and potential disruptions, governance risk.**
- **That's why shared vocabulary /framework is important.**
  - Makes assumptions explicit: Definitions embed different values, highlighting them for coherent policy.
  - Prevents talking past each other: Common vocabulary reduces misinterpretation.
  - Without a shared framework, analyses miss plausible futures and create blind spots for risk and opportunity.

AGI Pivots

- Instead of chasing a single AGI definition, we can focus on key factors or “pivots” that determine different post-AGI futures.
- Pivots are key factors that capture structural dynamics to classify post AGI future.
  - These are binary choices at two ends of a spectrum.
- Pivots are not technical definitions of AGI; they are **conceptual tools**.
  - Can be used to surface assumption, test policies across futures and capture uncertainty.

The Six Pivots

- **Centralization of Ownership:**
  - *Centralized*: A few actors (e.g., one company, state, or AGI itself) control AGI.
  - *Diffused*: Many actors have access to and control over AGI systems.
- **Governance Primacy:**
  - *Public Interest*: National or intergovernmental bodies steer AGI for societal benefit.
  - *Private Interest*: Corporations or private actors drive AGI development and governance.
- **Alignment Level:**
  - *Aligned*: AGI follows human instructions and values.
  - *Misaligned*: AGI goals diverge from or work against human intent and values.

The Six Pivots

- **Takeoff Speed:**
  - *Slow*: AGI capabilities grow gradually, allowing adaptation.
  - *Fast*: Capabilities advance quickly, potentially overwhelming institutions.
- **Human–AGI Relationship:**
  - *Complementary*: AGI augments and collaborates with humans.
  - *Replacement*: AGI substitutes human labor or decision-making at scale.
- **AGI Volition:**
  - *Independent Goal-Seeker*: AGI sets and pursues goals autonomously.
  - *Controlled Tool*: AGI operates only under human-set goals and oversight.

Example: Stable Future

- **Scenario:**
  - Aligned (Pivot: Alignment Level) + Slow Takeoff (Pivot: Takeoff Speed) + Complementary (Pivot: Human AGI Relationship)
- **Description:**
  - AGI is aligned with human values, develops gradually, and works alongside humans.
- **Implications:**
  - Societies adapt institutions over time; risks mitigated through oversight; geopolitical stability is reinforced.
- **Policy implication:**
  - Policymakers should invest in coordination mechanisms and long-term adaptation strategies.

Example: Unstable Future

- **Scenario:**
  - Misaligned (Pivot: Alignment Level) + Fast Takeoff (Pivot: Takeoff Speed) + Replacement (Pivot: Human AGI Relationship)
- **Description:**
  - AGI is misaligned, develops rapidly, and displaces humans.
- **Implications:**
  - Mass unemployment, institutional overload, rogue AGI misuse, global instability.
- **Policy implication:**
  - Preparedness requires crisis management tools, robust safeguards, and rapid response capacity.

Example: Aligned Hegemon

- **Scenario:**
  - Centralized (Pivot: Centralization of Ownership)+ Aligned (Pivot: Alignment Level)
- **Description:**
  - AGI is controlled by a single actor but remains aligned with human interests.
- **Implications:**
  - Stable enforcement of norms and guardrails, but concentration of power raises equity and authoritarian risks.
- **Policy implication:**
  - Policies should balance global coordination with safeguards against monopolization of power.

Example: Pandora

- **Scenario:**
  - Diffused (Pivot: Centralization of Ownership)+ Misaligned (Pivot: Alignment Level)
- **Description:**
  - AGI is widely available, but many systems fail to follow human instructions or values.
- **Implications:**
  - Fragmented environment with competing AGIs, unstable governance, expanded threat surfaces.
- **Policy implication:**
  - International standards and interoperable oversight mechanisms become critical to prevent cascading risks.

THANK YOU

Swaptik Chowdhury*

swaptikchowdhury16@gmail.com

Assistant Policy Researcher (RAND Corporation)

Data Scientist (Dunn and Bradstreet)

* Opinions presented here are not representing any of the affiliated organizations.