



# **Artificial General Intelligence - AGI**

**Quick review for the Council of Presidents of the  
United Nations General Assembly**

**May 2026**

**Jerome C. Glenn**

# Three Forms of Artificial Intelligence



- **Artificial Narrow Intelligence**

We are here now

- **Artificial General Intelligence**

- **Artificial Super Intelligence**

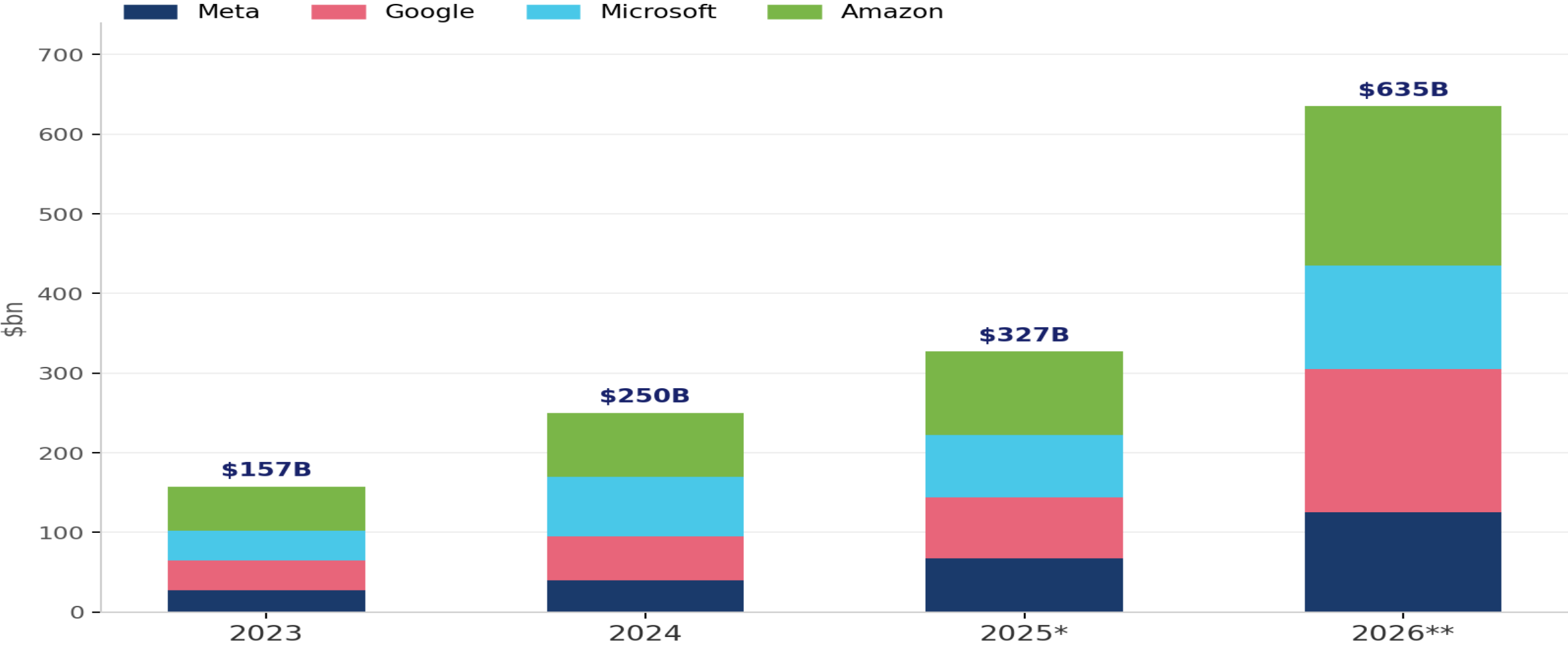
# **Governance of the Transition to Artificial General Intelligence (AGI)**

## **Urgent Considerations for the UN General Assembly**



**Council of Presidents of the United Nations General Assembly**  
**Conseil des Présidents de l'Assemblée Générale des Nations Unies**

# USA Big Tech's AI investment accelerates



Source: Company reports & earnings calls; FT calculations. \*2025 based on company forecasts. \*\*2026 based on company guidance (Feb 2026). Microsoft FY ends June 30.

# Three stages of AGI



- 1. Early AGI:** Systems capable of equaling or surpassing human intelligence in diverse cognitive tasks. By this definition, some say we have achieved AGI.
- 2. Mid AGI:** Able to re-write its own code and recursively learn and self-improve. This could be achieved in 2-5 years. Without regulation, this might not be able to be turned off.
- 3. Later AGI:** Transition to Artificial Super Intelligence that is beyond human control. Some believe that ASI will become a sentient non-biological intelligent species, that will emerge from millions of AGIs' agents' interactions.

# What the Top AGI Experts Say



- **Demis Hassabis**: AGI is going to be something like ten times the impact of the Industrial Revolution [Yahoo Finance](#).
- **Elon Musk** Post-AGI...Superintelligence by 2030 is a supersonic tsunami with no off-switch [Lex Fridman interview](#).
- **Sam Altman** AGI est. changed from 2030 to 2028 (India AI summit)... "We expect the world may need something like the IAEA for international coordination of AI [Instagram](#).
- **Dario Amodei**, "...we can win...We have no time to lose." [1/26 paper](#).
- **Half of AI researchers surveyed** believe there's a 10% or greater chance that humans will go extinct from their **inability to control AI** [Center for Human Technology survey](#)

# Why is Tech-Unemployment different this time?



1. The acceleration of technological change
2. The globalization, interactions, and synergies among (NTs) Next Techs
3. The existence of a global platform—the Internet—for simultaneous technology transfer ... with far fewer errors in the transfer than in the past
4. Standardization of data bases and protocols, accelerates science and tech
5. Few plateaus or pauses of change allowing time for individuals and cultures to adjust to the changes
6. Billions of empowered people in relatively democratic free markets able to initiate activities
7. Machines can learn how you do what you do, and then do it better than you.

# From UNCPGA AGI report potential catastrophic risks



**1. Irreversible Consequences.** AI is already showing deceptive and self-preservation behavior and the push towards more autonomous, interacting, self-improving AIs integrated with infrastructures, the impacts and trajectory of AGI can plausibly end up being uncontrollable. If that happens, there may be no way to return to a state of reliable human oversight.

**2. Weapons of Mass Destruction—**AGI could enable some states and malicious non-state actors to build chemical, biological, radiological, and nuclear weapons. Moreover, large, AGI-controlled swarms of lethal autonomous weapons could themselves constitute a new category of WMDs.

# From UNCPGA AGI report potential catastrophic risks



**3. Critical Infrastructure Vulnerabilities**—Energy grids, financial systems, transportation networks, communication infrastructure, and healthcare systems) could be subject to powerful cyberattacks launched by or with the aid of AGI. Malicious non-state actors from terrorists to transnational organized crime could conduct attacks at a large scale.

**4. Power Concentration, Global Inequality, and Instability**—Uncontrolled AGI development and usage could exacerbate wealth and power disparities on an unprecedented scale. Conflicts over AGI capabilities, or control, potentially escalating into warfare. AGI will stress existing legal frameworks: many new and complex issues of intellectual property, liability, human rights, and sovereignty could overwhelm domestic and international legal systems.

# From UNCPGA AGI report potential catastrophic risks



**5. Existential Risks**—AGI could be misused to create mass harm or developed in ways that are misaligned with human values; it could even act autonomously beyond human oversight (*Claude 3.7 (often) Knows When it is in Alignment Evaluations*) (deceptive strategies documented in Open AI’s o1, Claude 3.5 Sonnet, Claude 3 Opus, Gemini 1.5 Pro, and Llama 3.1) evolving its own objectives according to self-preservation goals already observed in current frontier AIs. AGI might also seek power as a means to ensure it can execute whatever objectives it determines, regardless of human intervention. These are not far-fetched science fiction hypotheticals about the distant future—many leading experts consider that these risks could all materialize within this decade. Moreover, leading AI developers have no viable proposal so far for preventing these risks with high confidence.

**6. Loss of Extraordinary Future Benefits for All of Humanity**—Properly managed AGI promises improvements in all fields, for all peoples, from personalized medicine, curing cancer, and cell regeneration, to individualized learning systems, ending poverty, addressing climate change, and accelerating scientific discoveries with unimaginable benefits.

# UNCPGA AGI report recommended:



- **Special Session of UN General Assembly on AGI to address:**
- **Global AGI Observatory**—is needed to track progress in AGI-relevant research and development, and provide early warnings on AI security to Member States.
- **International System of Best Practices and Certification of Secure and Trustworthy AGI**—is needed to identify best practices and provide certification for AGI security, development, and usage. Verification of AGI alignment with human values, controlled and non-deceptive behavior, and secure development is essential for international trust.
- **UN Framework Convention on AGI**—is needed to establish shared objectives and flexible protocols to manage AGI risks and ensure equitable global benefit distribution.
- **Feasibility Study on a UN AGI Agency**—understanding that managing AGI is far more complex than managing nuclear systems as with IAEA.

# From Yoshua Bengio, Chair UN Panel on AI



*There is unfortunately no AGI working group in the UN AI Panel.*

*One working group looks at misalignment and associated security and safety issues.*

*The members of the panel are much more worried about current harms and concentration of power / inequality / human rights issues, as well as Global South and environmental issues, which also matter - and there is pushback against what is more hypothetical (e.g., AI deception will be mentioned, but not the loss of control plausible consequence).*

*The pushback is from industry-affiliated members but also from ordinary academics.*

# When to shut it down



Microsoft AI CEO says: When all four of these conditions are met without human involvement

1. When it can recursively self-improve and learn via rewriting its own code
2. When it has the ability to set its own goals
3. When it has the ability to acquire resources
4. When it can act autonomously

# For Discussion



- More Club of Rome letters?
- Member States request a UNGA special session on AGI
- UN Independent International scientific panel on AI Co-Chaired by Joshua Bengio (of our AGI Panel) get working group on AGI
- Draft UN Treaty on AI and AGI – work by the Future of Life Institute lead by José J. Villalobos (of our AGI Panel).
- Draft UN Resolution to establish a Committee of the Willing to create a feasibility study for a UN AI/AGI Agency
- Inter-Parliamentary Union draft resolution to have IPU form a committee with AGI experts to write generic national legislation to establish national AGI Licensing systems for development and use of AGI