



AGI Background

Unregulated future AI is likely to be a disaster, warns experts around the world. There are three kinds of AI: narrow, general, and super. **Artificial Narrow Intelligence** (ANI) including generative AI (or GenAI) is narrow, pre-trained tool that exists today and will continue to improve over the coming years. **Artificial General Intelligence** (AGI) does not exist yet, but is expected to act more like an agent than a simple tool, and can re-write its own code, create novel solutions to novel problems similar to humans (sometimes called human-level AI) and can learn moment by moment to continually get smarter.

AGI could emerge within this decade. **Artificial Super Intelligence** (ASI) sometimes called Artificial Superintelligence emerges from AGI to set its own goals independent from human control.

Some progress on ANI governance:

- US-lead UN General Assembly resolution on AI addresses ANI, not AGI
- The EU AI Act and China's Artificial Intelligence Law address ANI, not AGI
- Bletchley Declaration signed by 28 governments addresses ANI, not AGI

The US-China AI cooperation talks might be exploring AGI governance, but that is not publicly known. Previous White House statements and executive orders address ANI, not AGI.

What is needed now?

- Draft text to authorize a national AGI licensing system and authorize NIST to prepare simulations to test the safety of AGI prior to licensing.
- Use Phase 1 and especially Phase 2 of The Millennium Project's AGI governance study to help think through these issues.
- Consider the national suggestions in The Gladstone Defense in Depth: *An Action Plan to Increase the Safety and Security of Advanced AI*.
- Support a UN General Assembly Resolution to form a group of the willing to draft the text of a UN Convention on AI with two sections: one on ANI and one on AGI.

As The Millennium Project Phase 2 AGI report states, "Governing AGI could be the most complex, difficult management problem humanity has ever faced." Furthermore, failure to solve it before proceeding to create AGI systems would be a fatal mistake for human civilization. No entity has the right to make that mistake.

---- Stuart Russell

Contact: Jerome.Glenn@Millennium-project.org or 1-202-669-4410 for copies of the Phase 1 and 2 AGI reports.

Current Sponsors and Inkinds: 4CF-Strategi Foresight, Poland; AGAHI, Pakistan; Dubai Future Foundation; Egyptian Acad. of Scientific Res. and Tech.; Future of Life Institute, George Washington University; Harvard University; Google; Organization of American States (OAS), Smithsonian Institution; UNESCO, and World Academy of Art and Science.

Board of Directors

Elizabeth Florescu
Jerome C. Glenn, CEO and Acting Chair
Concepción Olavarrieta
Charles Perrotet
Paul Saffo
Ibon Zugasti

Ex Officio

Jose Cordeiro, RIBER
Nicolas Balcom Raleigh, FEN

Planning Committee

Asanga Abeysoonasekera
Veronica Agreda
Rosa Alegria
Mohsen Bshrami
Clem Buzold
Beatriz Bechara de Borge
Mara Di Berardo
Héctor Casameva
Tamar Chachibai
Purnash Chaudhary
Epaminondas Christophilopoulos
José Luis Cordeiro
Martín Kruse
Sherman Cruz
Cornelia Daheim
Francisco Dallmeier
Philippe Destatte
Mara Di Berardo
Pedro Miguel Diegues
Tibomir Divjak
Paul Epping
Jefé Ezzi
Karl Friðriksson
Nadezhda Gaponenko
Lydia Garrido
Blaz Golob
Miguel A. Gutierrez
Sirikka Heinonen
Brock Hinzmann
Arnoldo José de Hoyos
Reyhan Husaynova
Pramod Jaiswal
Zhouying Jin
Ivan Klince
Hayato Kobayashi
Norbert Kolos
Rasgan Maharajh
Kamal Zaki Mahmood
Savvas Mantzarakis
Carlos Marschall
Muhammad Al-Mheiri
Francisco José Mojica
Kate McCallum
Pavel Novacek
Concepción Olavarrieta
Fernando Ortega
Youngsook Park
Charles Perrotet
Christin Pfeiffer
Adrian Pop
Stephanie Pride
David Rejski
Saphia Richou
Stanley G. Rosen
Paul Saffo
Karl Schroeder
Vair Sharan
Mihaly Simai
Nshokov edzo Simbarashe
Yarima Sosa
Rohit Talwar
Mariana Todorova
Paul Werbos
David Wood
Simone Di Zio
Ibon Zugasti

Nodes

Amsterdam, the Netherlands
Arts/Media, Global
Baku, Azerbaijan
Beijing, China
Belgrade, Serbia
Berlin/Essen, Germany
Bogotá, Colombia
Brussels Area, Belgium
Budapest, Hungary
Bucharest, Romania
Buenos Aires, Argentina
C. Europe: Czechia, Poland, Slovakia
Cairo, Egypt
Caracas, Venezuela
Caribbean Node
Colombo, Sri Lanka
Copenhagen, Denmark
Cyber Node Internet
Doha, United Arab Emirates
East Africa: Kenya, Tanzania, Uganda
Helsinki, Finland
Harare, Zimbabwe
Islamabad, Pakistan
Istanbul, Turkey
Kathmandu, Nepal
Kuwait City, Kuwait
Lima, Peru
Lisbon, Portugal
Ljubljana, Slovenia
London, UK
Manila, Philippines
Melbourne, Australia
Mexico City, Mexico
Montevideo, Uruguay
Moscow, Russia
New Delhi, India
Panama City, Panama
Paris, France
Penang, Malaysia
Pretoria/Johannesburg, South Africa
Prishtina, Kosovo
Quebec/Toronto, Canada
Reykjavik, Iceland
Rome/Pescara, Italy
San José, Costa Rica
San Sebastián and Madrid, Spain
Santa Cruz de la Paz, Bolivia
San Salvador, El Salvador
Santiago, Chile
Santo Domingo, Dominican Republic
São Paulo and Curitiba, Brazil
SE Europe: Croatia, Montenegro, Slovenia, Serbia
Seoul, Rep. of Korea
Silicon Valley, USA
Sofia, Bulgaria
Tbilisi, Georgia
Tehran, Iran
Tel Aviv, Israel
Thessaloniki and Athens, Greece
Tokyo, Japan
Tunis, Tunisia
Warsaw, Poland
Wellington, New Zealand