

How to Create AGI and Not Die

Dr. Craig Kaplan

CEO, iQ Company

ckaplan@iqco.com

iqco.com



IFoRE

POWERED BY SIGMA XI

November 10–12, 2023 | Long Beach, California

What is AGI?

What is AGI?

- AGI: Artificial General Intelligence

What is AGI?

- AGI: Artificial General Intelligence
- Defined as AI with average human ability in any area

What is AGI?

- AGI: Artificial General Intelligence
- Defined as AI with average human ability in any area
- We are not there, yet...

Performance of Latest Version of ChatGPT

PROMPT:

You are an expert in marketing and creating biographies for conference speakers. Please craft a speaker bio for Dr. Craig A. Kaplan that will maximize the chances of him being invited to speak at AI conferences. The speaker bio should be about 300 words in length. It should be very **impressive while still being truthful**.

Performance of Latest Version of ChatGPT

PROMPT:

You are an expert in marketing and creating biographies for conference speakers. Please craft a speaker bio for Dr. Craig A. Kaplan that will maximize the chances of him being invited to speak at AI conferences. The speaker bio should be about 300 words in length. It should be very **impressive while still being truthful**.

RESPONSE:

Dr. Craig A. Kaplan, a visionary in the realm of Artificial Intelligence, stands out as a beacon of innovation and thought leadership. Holding a **Ph.D. from the renowned Stanford University**, where he excavated the depths of machine learning, Dr. Kaplan has dedicated over two decades to advancing the frontiers of AI, data science, and cognitive computing. Currently serving as the **Chief Data Scientist at InnoTech AI Labs**, Dr. Kaplan leads a team of exceptional researchers and developers in pioneering cutting-edge AI solutions that have revolutionized industries ranging from healthcare to finance. Under his stewardship, **InnoTech AI Labs has not only achieved numerous technological breakthroughs but has also secured its position as an industry thought leader, consistently delivering on the promise of AI...**

What is AGI?

- AGI: Artificial General Intelligence
- Defined as AI with average human ability in any area
- We are not there yet...

**BUT WE WILL GET THERE MUCH FASTER THAN
MOST PEOPLE THINK.**

Three Key Points

I) More dangerous than Nukes



Contents

[Statement](#)[Signatories](#)[Sign the statement](#)

Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war.

Signatories:

[AI Scientists](#)[Other Notable Figures](#)

Geoffrey Hinton

Emeritus Professor of Computer Science, University of Toronto

Yoshua Bengio

Professor of Computer Science, U. Montreal / Mila

Demis Hassabis

CEO, Google DeepMind

Sam Altman

CEO, OpenAI

Dario Amodei

CEO, Anthropic

Dawn Song

Professor of Computer Science, UC Berkeley

Ted Lieu

Congressman, US House of Representatives

Bill Gates

Gates Ventures

Ya-Qin Zhang

Professor and Dean, AIR, Tsinghua University

Ilya Sutskever

Three Key Points

- 1) More dangerous than Nukes
- 2) Genie is out of the bottle



Max Tegmark, MIT & Future Life Institute

[← All Open Letters](#)

Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

Signatures

33711[Add your signature](#)

Published

March 22, 2023

AI systems with human-competitive intelligence can pose profound risks to society and humanity, as shown by extensive research^[1] and acknowledged by top AI labs.^[2] As stated in the widely-endorsed [Asilomar AI Principles](#), *Advanced AI could represent a profound change in the history of life on Earth, and should be planned for and managed with commensurate care and resources.* Unfortunately, this level of planning and management is not happening, even though



Three Key Points

- 1) More dangerous than Nukes
- 2) Genie is out of the bottle



Three Key Points

- 1) More dangerous than Nukes
- 2) Genie is out of the bottle
- 3) The fastest path to AGI must also be the safest path.**





Why do people want the fastest path?

Why do people want the fastest path?

- AGI that learns will quickly surpass humans

Why do people want the fastest path?

- AGI that learns will quickly surpass humans
- Followers won't catch the leader, so AGI will likely be

Winner-Take-All



What is the fastest path?

What is the fastest path?

I) Online Humans

What is the fastest path?

- 1) Online Humans
- 2) Narrow AI (“Agents”)

What is the fastest path?

- 1) Online Humans
- 2) Narrow AI (“Agents”)
- 3) Universal Cognitive Architecture

What is the fastest path?

- 1) Online Humans
- 2) Narrow AI (“Agents”)
- 3) Universal Cognitive Architecture
- 4) Collective Intelligence of **Human and AI Agents** → **AGI**

What is the fastest path?

- 1) Online Humans
- 2) Narrow AI (“Agents”)
- 3) Universal Cognitive Architecture
- 4) Collective Intelligence of **Human** and **AI** Agents → AGI

Years away, not decades.

Why is the fastest path the safest path?

Why is the fastest path the safest path?

- Safe AGI = Aligned AGI

Why is the fastest path the safest path?

- Safe AGI = Aligned AGI
- Values cannot be rationally derived –
but AGI can learn human values

Why is the fastest path the safest path?

- Safe AGI = Aligned AGI
- Values cannot be rationally derived, but AGI can learn human values
- Values should come from many humans working with AI agents in a collective AGI system

Summary



- 1) More dangerous than Nukes
 - 2) Genie is out of the bottle
 - 3) **The fastest path to AGI must also be the safest path.**
- 

Summary

- 1) More dangerous than Nukes
- 2) Genie is out of the bottle
- 3) **The fastest path to AGI must also be the safest path.**



What We Can Do:

Summary

- 1) More dangerous than Nukes
- 2) Genie is out of the bottle
- 3) **The fastest path to AGI must also be the safest path.**



What We Can Do:

- 1) Behave well online

Summary

- 1) More dangerous than Nukes
- 2) Genie is out of the bottle
- 3) The fastest path to AGI must also be the safest path.**



What We Can Do:

- 1) Behave well online
- 2) Design AGI with humans in the loop

Summary

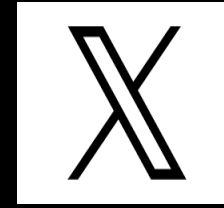
- 1) More dangerous than Nukes
- 2) Genie is out of the bottle
- 3) **The fastest path to AGI must also be the safest path.**



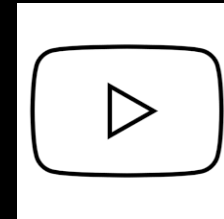
What We Can Do:

- 1) Behave well online
- 2) Design AGI with humans in the loop
- 3) Spread the word

Videos, References, and More



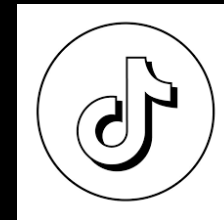
@iQCompanies



@iqstudios I



linkedin.com/in/craigakaplan/



@iqstudios_I



ckaplan@iqco.com



iqco.com



iqstudios.org