

RISKS PROPOSAL PROTESTS FAO ACT DONATE JOIN Q

List of p(doom) values

p(doom) is the probability of very bad outcomes (e.g. human extinction) as a result of AI. This most often refers to the likelihood of <u>AI taking over</u> from humanity, but different scenarios can also constitute "doom". For example, a large portion of the population dying due to a novel biological weapon created by AI, social collapse due to a <u>large-scale</u> <u>cyber attack</u>, or AI causing a nuclear war. Note that not everyone is using the same definition when talking about their p(doom) values. Most notably the time horizon is often not specified, which makes comparing a bit difficult.

- **Yann LeCun** one of three godfathers of AI, works at Meta (less likely than an asteroid)
 - 10% Vitalik Buterin Ethereum founder
 - **Geoff Hinton** one of three godfathers of AI (wipe out humanity in the next 20 years)

https://pauseai.info/pdoom 1/4

| 9-19.4% | Machine learning researchers (From 2023, depending on the question design, median values: 5-10%) | |
|-------------|--|--|
| 15% | Lina Khan head of FTC | |
| 10-20% | Elon Musk CEO of Tesla, SpaceX, X | |
| 10-25% | Dario Amodei CEO of Anthropic | |
| 20% | Yoshua Bengio one of three godfathers of AI | |
| 5-50% | Emmett Shear | Co-founder of Twitch, former interim CEO of OpenAI |
| 30% | Al Safety Researchers (Mean from 44 AI safety researchers in 2021) | |
| 33% | Scott Alexander Popular Internet blogger at Astral Codex Ten | |
| 35% | Eli Lifland Top competitive forecaster | |
| 40% | Al engineer (Estimate mean value, survey methodology may be flawed) | |
| 40% | Joep Meindertsma Founder of PauseAI (The remaining 60% consists largely of "we can pause".) | |
| 46% | Paul Christiano | former alignment lead at OpenAI, founder of ARC |
| 50 % | Holden Karn | ofsky Executive Director of Open Philanthropy |

https://pauseai.info/pdoom 2/4

Jan Leike Former alignment lead at OpenAI 10-90%

Zvi Mowshowitz AI researcher 60%

Daniel Kokotajlo Forecaster & former OpenAI researcher 70%

Dan Hendrycks Head of Center for AI Safety >80%

Eliezer Yudkowsky Founder of MIRI >99%

99.9999%Roman Yampolskiy AI safety scientist

Do something about it

However high your p(doom) is, you probably agree that we should not allow AI companies to gamble with our future. Join PauseAI to prevent them from doing so.

JOIN PAUSEAI >

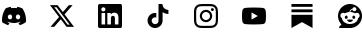














Info **Risks**

Risks overview FAO

Proposal AI Outcomes

Existential risk Learn

Press / Media Psychology of x-risk

AI takeover People

5/28/24, 11:40 PM

Teams

Legal Info Stichting PauseAI

(kvk 92951031)

Take Action

Join PauseAI

How you can help

Local communities

Donate

Merchandise ♂

Protests

Organize a protest

Email builder

Lobby tips

List of p(doom) values

Cybersecurity

Dangerous capabilities

State of the art

Urgency

Other

Auto Theme

Withdraw consent

Edit on GitHub♂

All pages

RSS

License: CC-BY 4.0년

Submit feedback ♂