

# Can AI cause the end of the human race?

Marko Baštovanović



# First things first

- The topic is controversial
- Care on getting the facts right
- You will answer if AI can cause the end of the human race

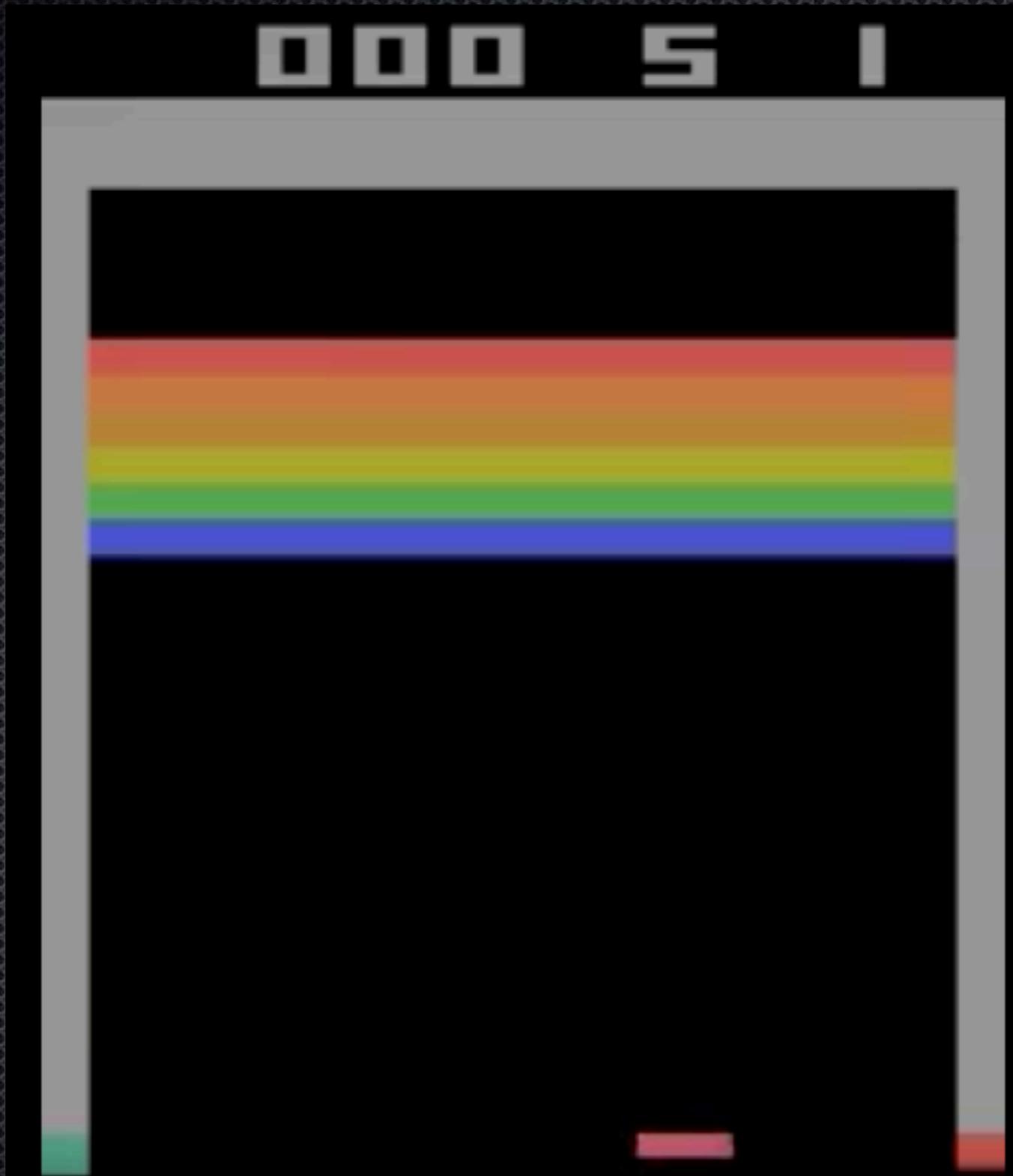
# Current state

- 1997 IBM DeepBlue
- 2016 AlphaGo
- 2017 OpenAI DOTA2



Source: Lee Jin-man / AP

# RL is the way to go in AI

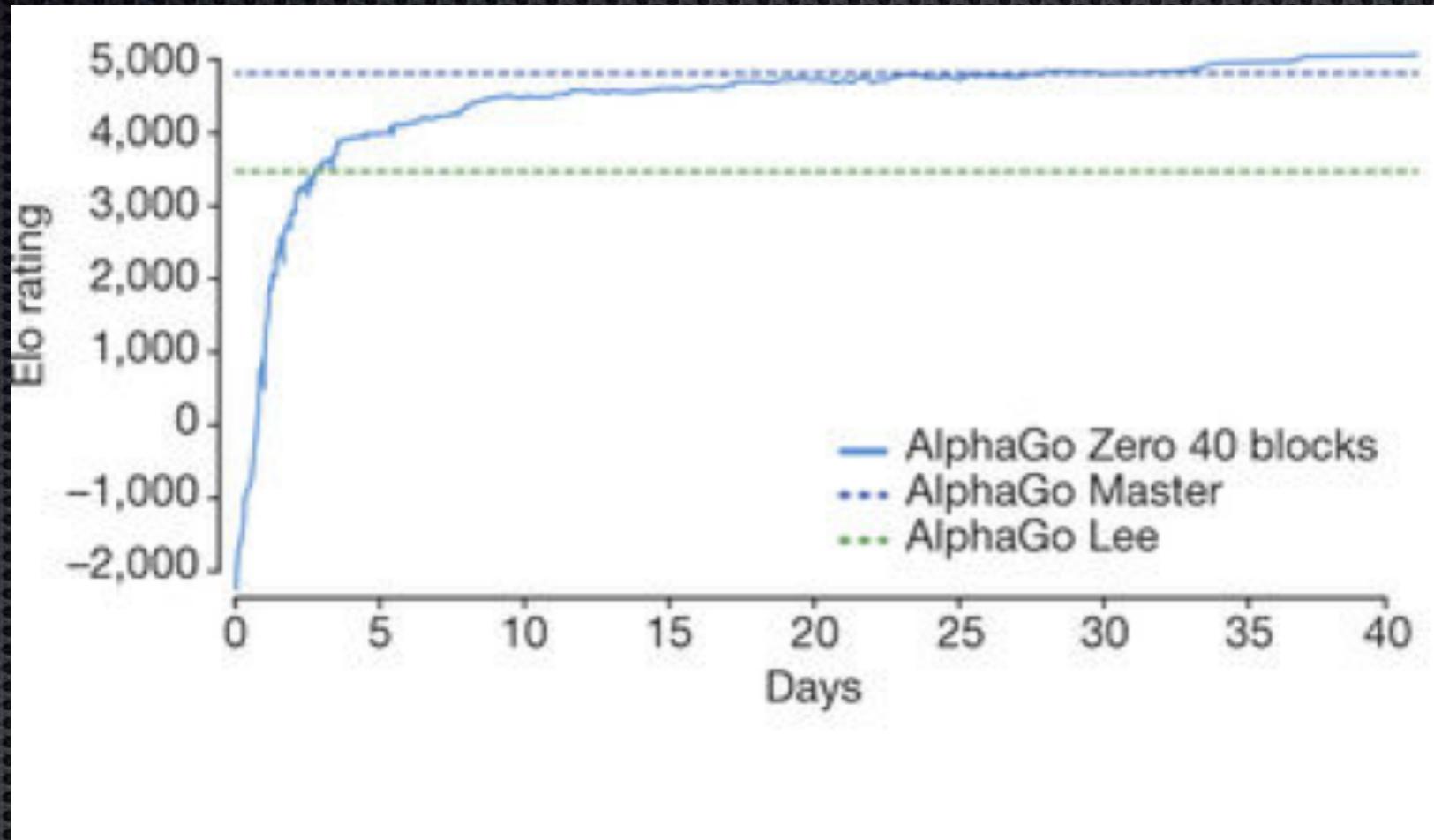


Source: Google DeepMind

# RL is the way to go in AI

Video removed due to GitHub space limitations. See it at:  
<https://www.youtube.com/embed/Q70uIPJW3Gk?start=1&end=55>

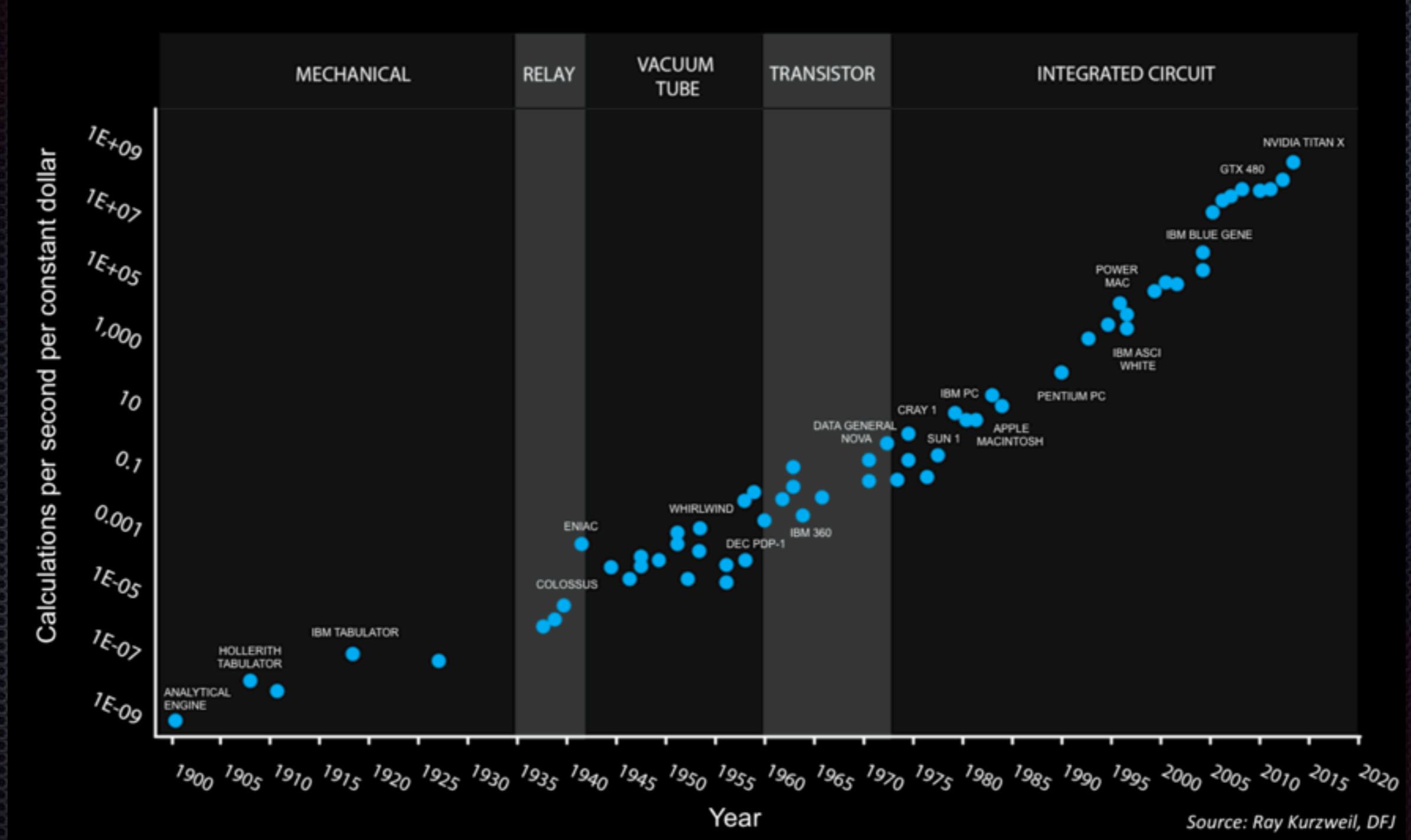
# After a problem is solved



Source: deepmind.com

- Once the human level is reached we soon surpass it by a large margin

# 120 Years of “Moore’s Law”



# Predictions about AI

- Why we don't have a base on Moon
- Wilbur Wright 1901: “I don't think men will fly for a thousand years”
- Watch for: incentives, financials, talent

# Financial incentives

## DeepMind AI Reduces Google Data Centre Cooling Bill by 40%

From smartphone assistants to image recognition and translation, machine learning already helps us in our everyday lives. But it can also help us to tackle some of the world's most challenging physical problems -- such as energy consumption. Large-scale commercial and industrial systems like data centres consume a lot of energy, and while much has been done to **stem the growth of energy use**, there remains a lot more to do given the world's increasing need for computing power.

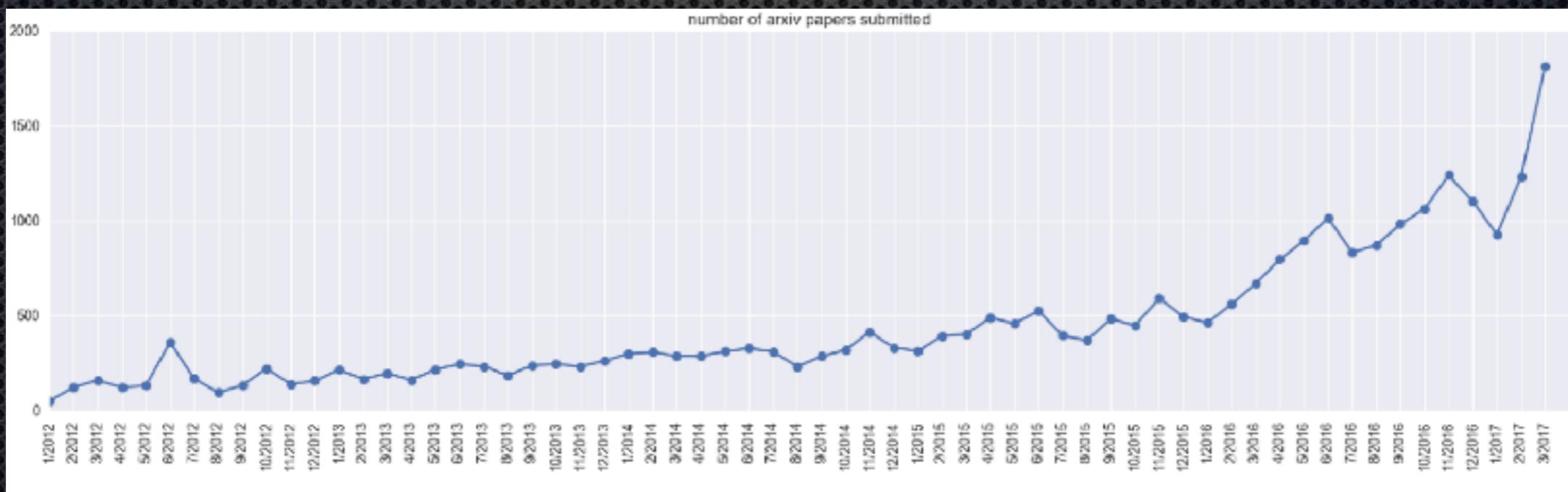
Reducing energy usage has been a major focus for us over the past 10 years: we have built our own **super-efficient servers** at Google, invented **more efficient ways to cool our data centres** and invested heavily in **green energy sources**, with the goal of being powered 100 percent by renewable energy. Compared to five years ago, we now get around 3.5 times the computing power out of the same amount of energy, and we continue to make many improvements each year.

Source: [deepmind.com](https://deepmind.com)

# Where does the talent go

The New York Times

## A.I. Researchers Are Making More Than \$1 Million, Even at a Nonprofit



Number of ML/AI arxiv papers submitted (2012-2017)

Source: Andrej Karpathy

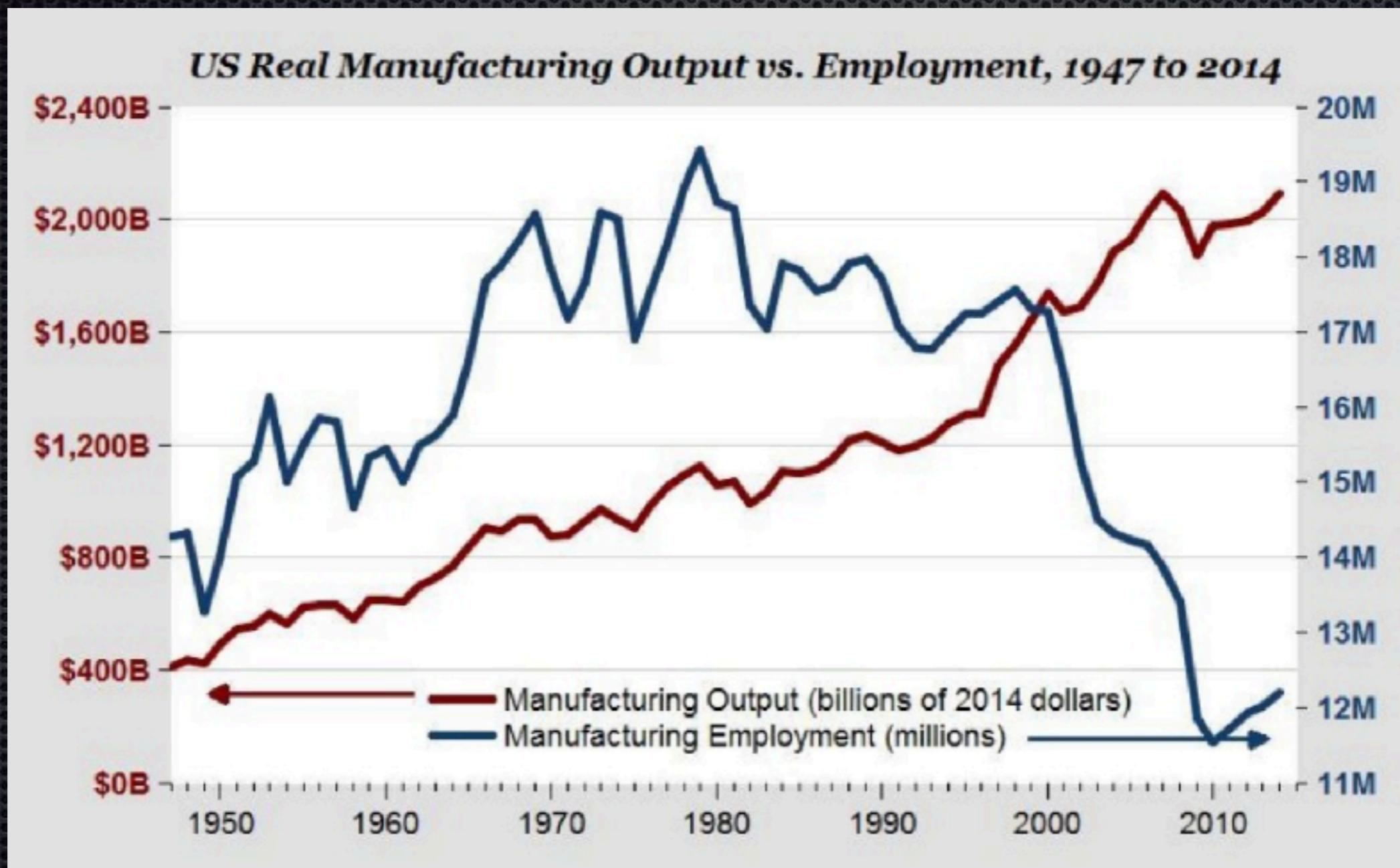
# The good stuff

- Abundance
- Make everybody happy
- Medicine
- Immortality
- Colonise other planets
- and many more



Source: Dimension Films

# The Bad Stuff: Massive unemployment

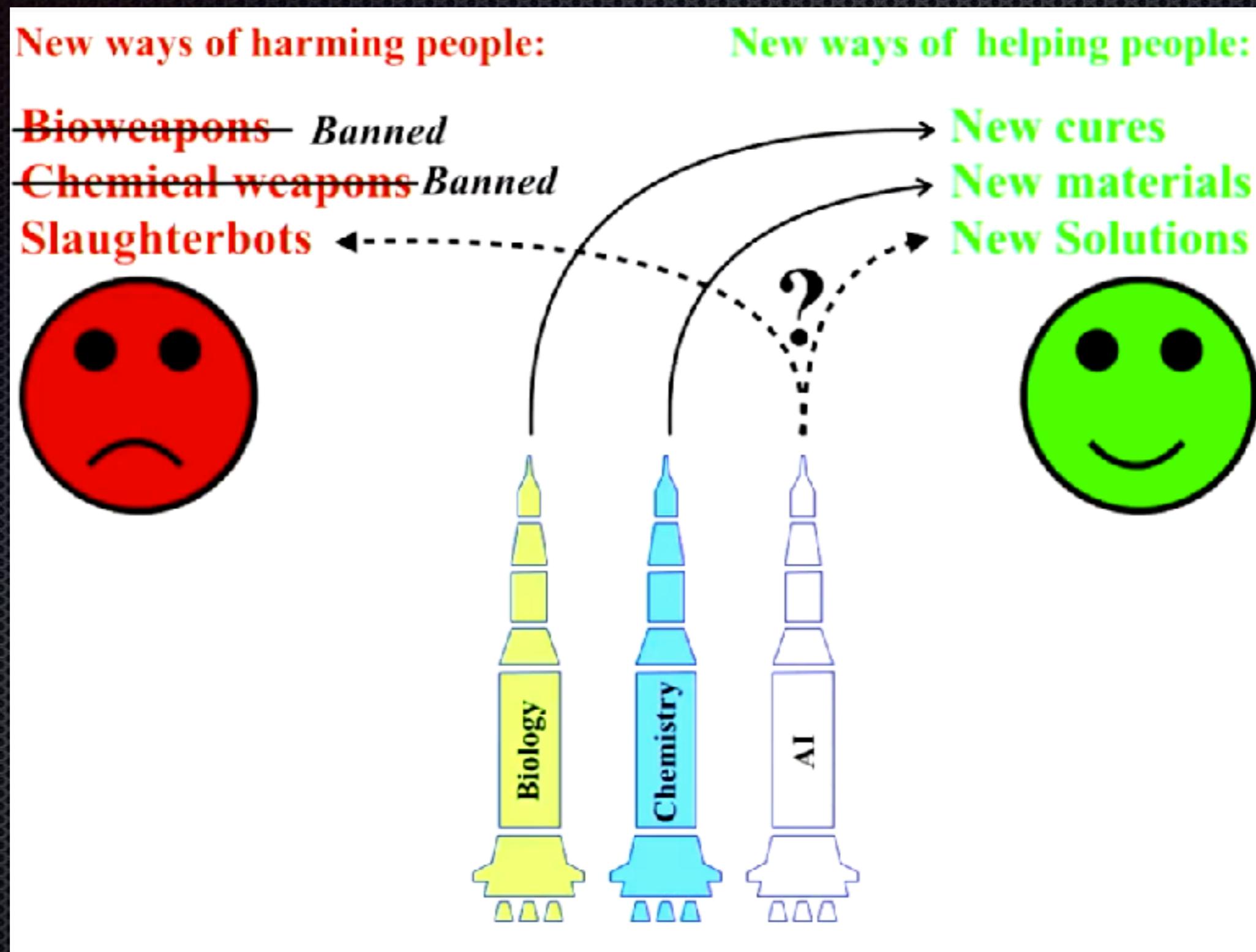


Source: Moshe Vardi

# The Bad Stuff: AI Weaponization

Video removed due to GitHub space limitations. See it at:  
<https://www.youtube.com/embed/9CO6M2Hs0IA?start=15&end=96>

# Ban Lethal AI Weapons



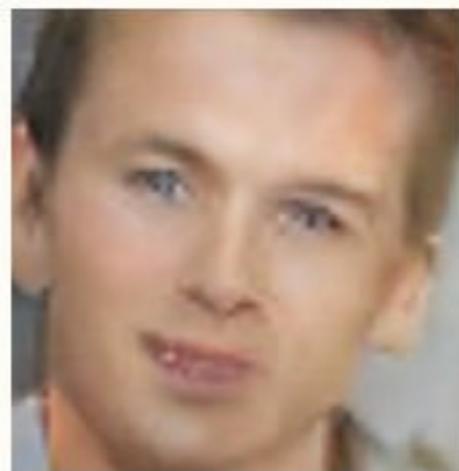
Source: Max Tegmark

# Malicious uses of AI

- ❖ Spear phishing
- ❖ Speech, image and video generation
- ❖ Political instability



2014



2015



2016



2017

Source: Goodfellow et al. (2014)

# Future Progress in AI: A Survey of Expert Opinion

- 170 AI experts
- Superintelligence:  
50% by 2040-2050  
90% by 2075
- 1/3 that outcome is  
“bad” or “extremely bad”



Nick Bostrom

# AI Alignment

King Midas



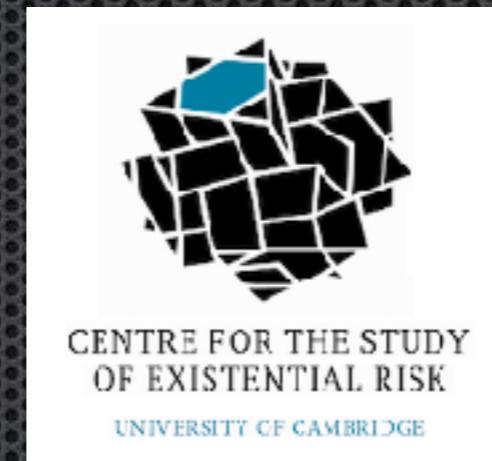
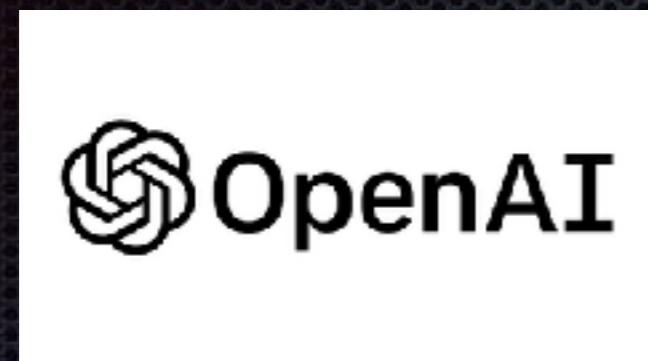
Source: Walter Crane

# Don't panic!



- BENEFICIAL AI 2017 conference 23 principles
- Signed by: Demis Hassabis, Ilya Sutskever, Yann LeCun, Stuart Russell, Peter Norvig, Ray Kurzweil, Stephen Hawking, Elon Musk, Jaan Tallinn, Sam Altman

# But do help if you can



# Thank you!

## Q & A

Marko Baštovanović

