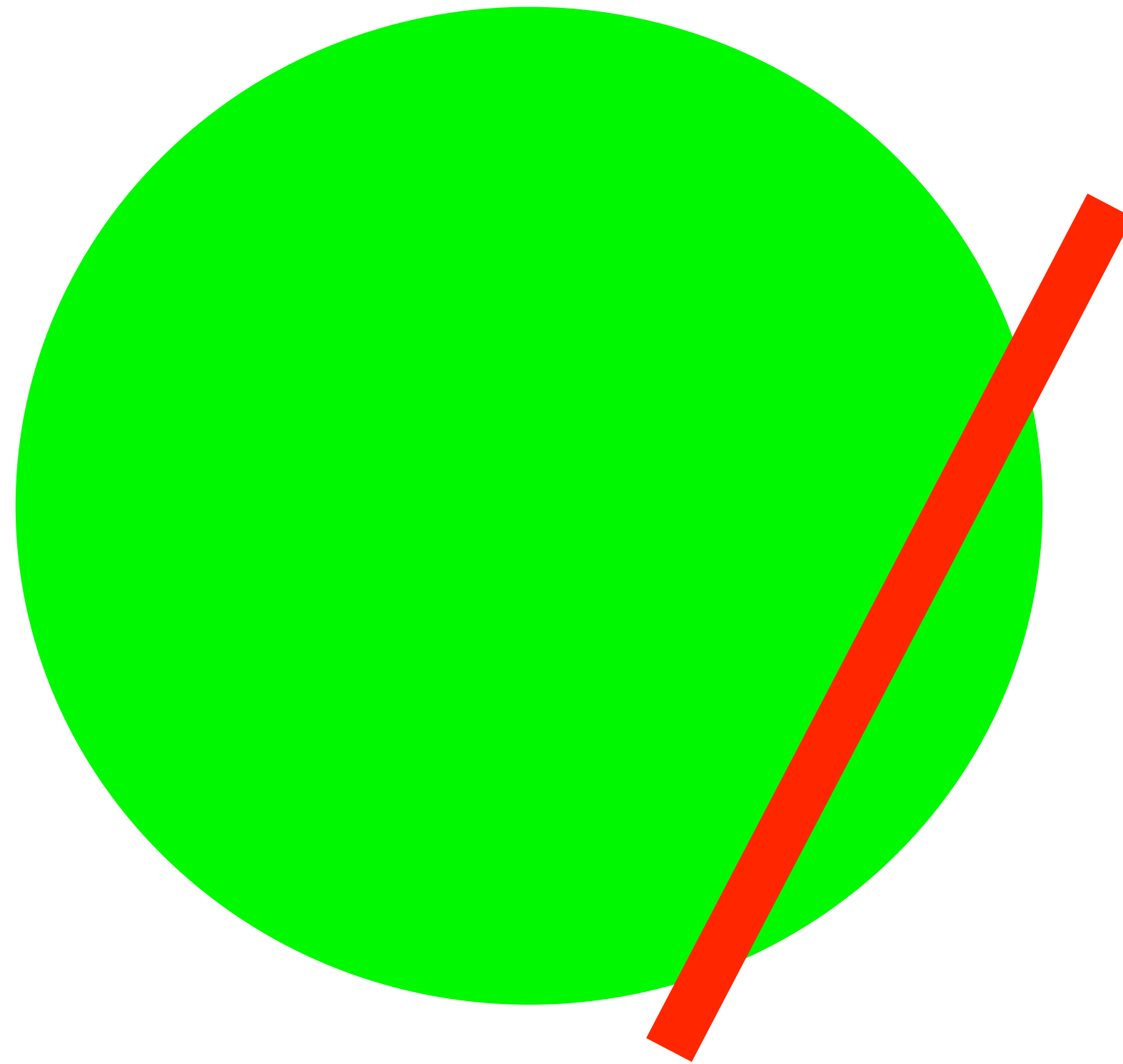


# How to get the upside of AI without the downside



**Max Tegmark**



**Massachusetts  
Institute of  
Technology**



CENTER FOR  
Brains  
Minds+  
Machines









**AI is getting  
more powerful**





Boston Dynamics: Atlas



T+ 00:07:58

STAGE 2

TELEMETRY

SPEED

ALTITUDE



FALCON HEAVY TEST FLIGHT

SpaceX





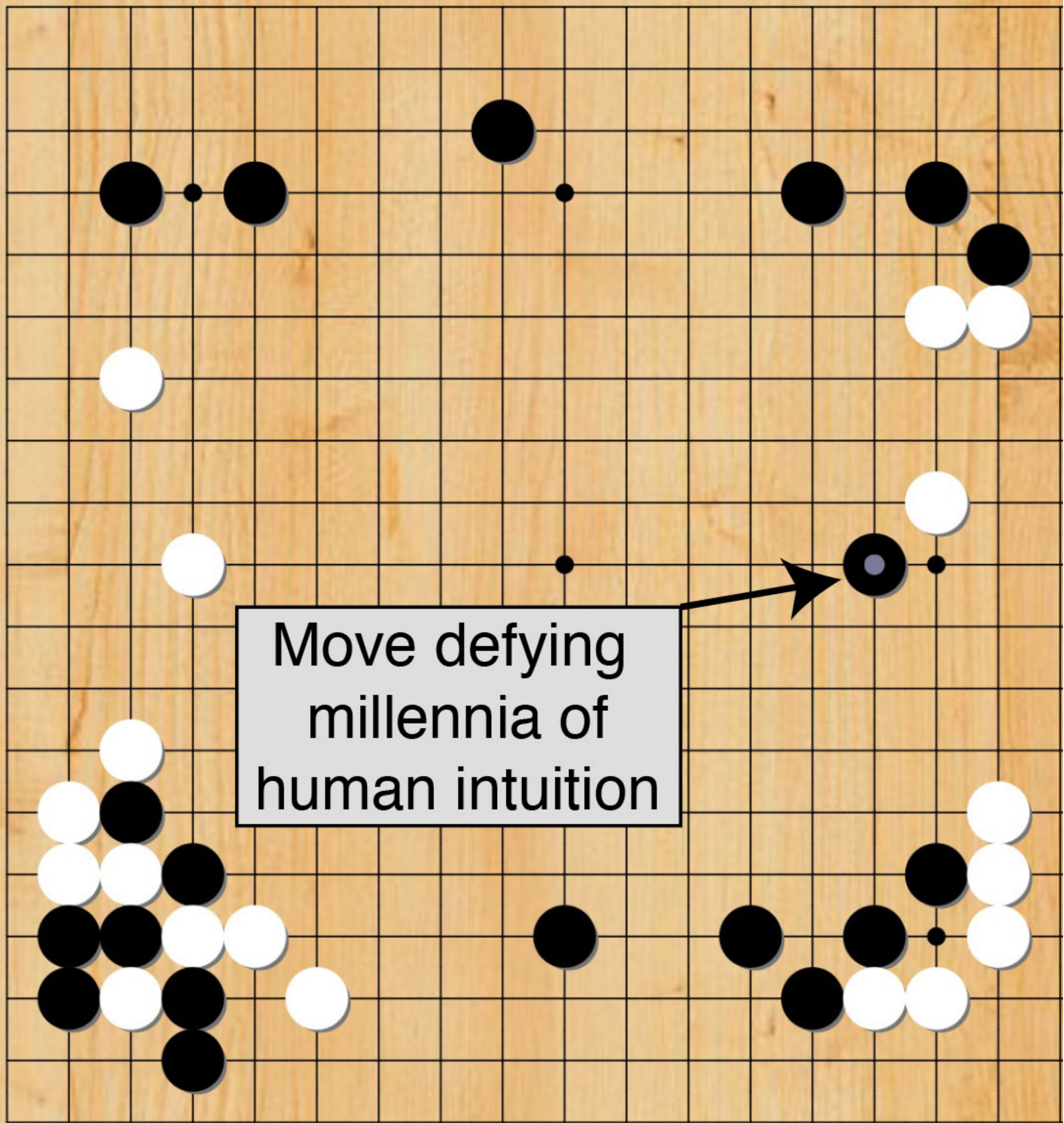




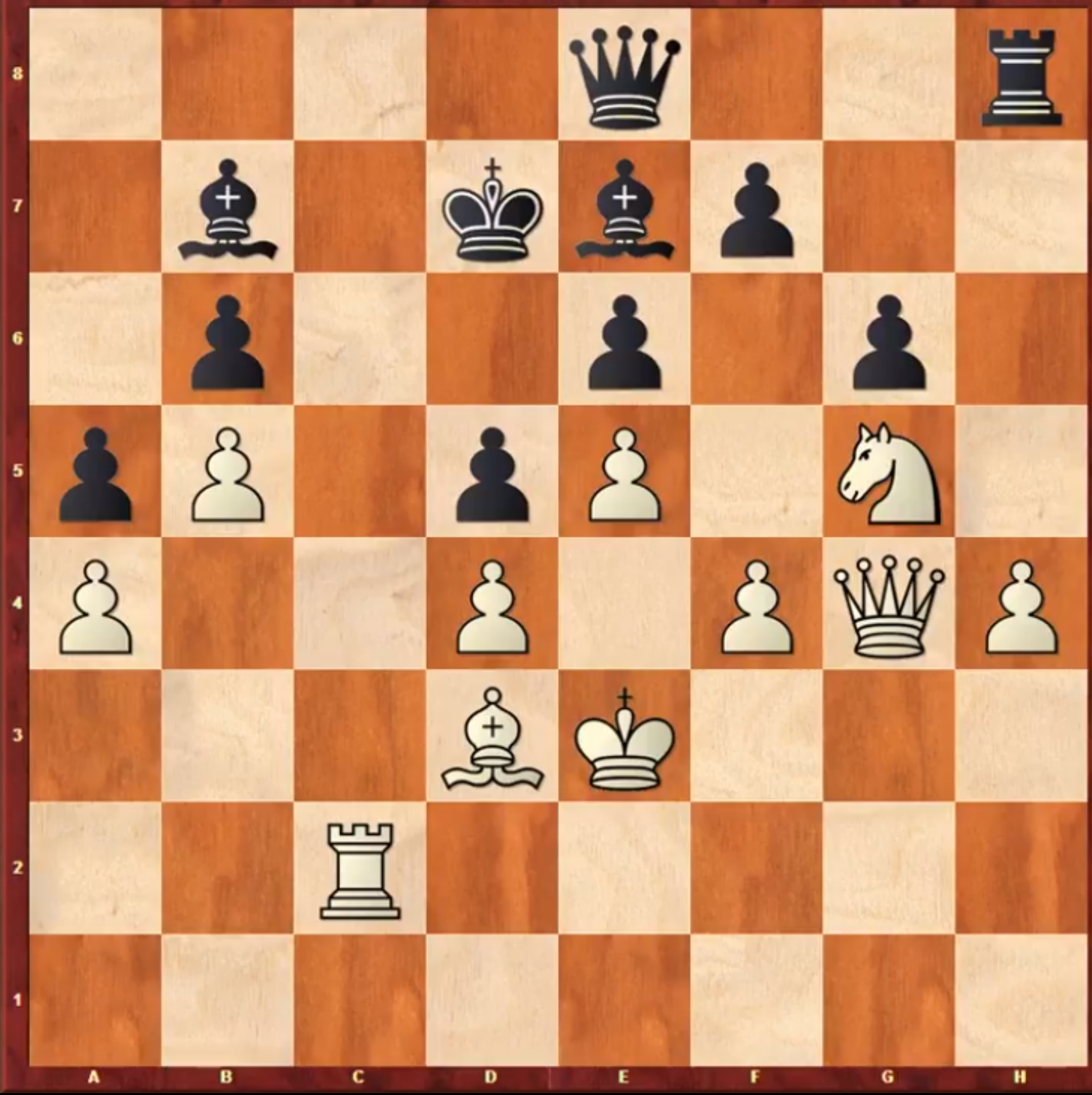
vs.



AlphaGo







## Google's AlphaZero Destroys Stockfish In 100-Game Match



FM MikeKlein 🇺🇸 ♟️

Dec 6, 2017, 12:50 PM | 💬 349 | Chess Event Coverage

🌐 English

Chess changed forever today. And maybe the rest of the world did, too.



**AI can solve  
problems...**



# How many lives can AI save?



Over a million road deaths/year,  
most caused by human error

Most soon preventable by AI?



# How many lives can AI save?



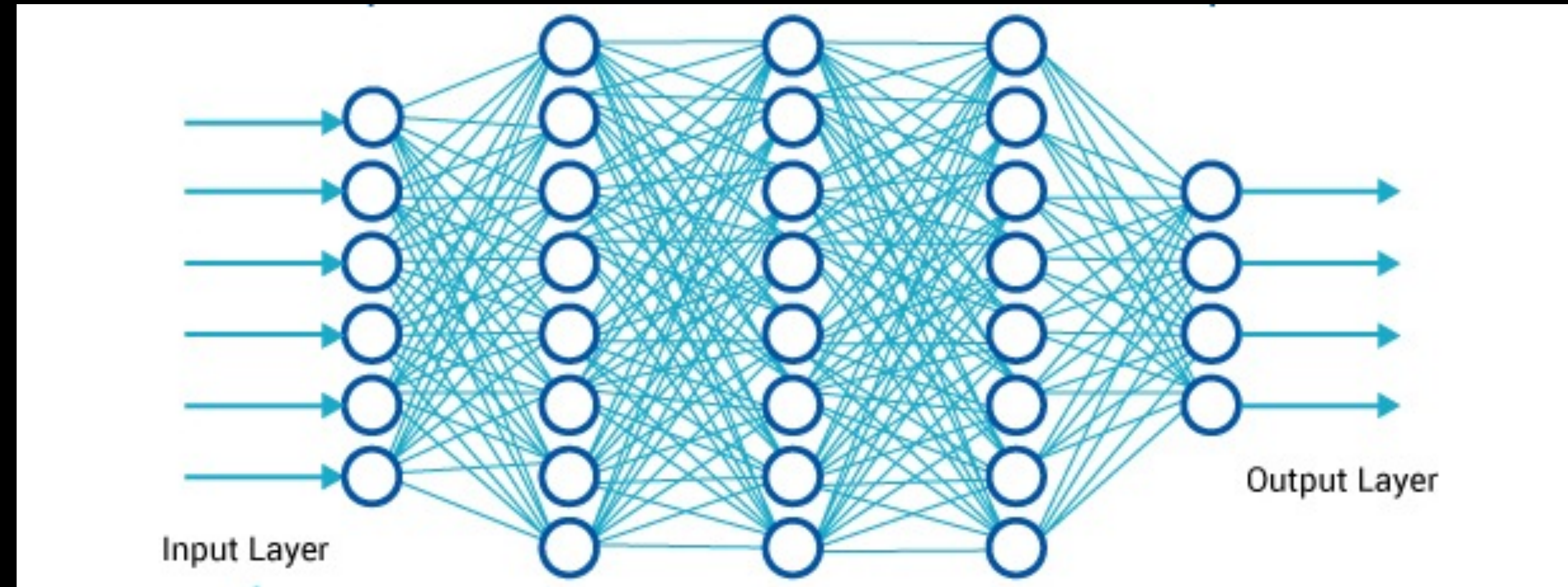
Bad hospital care kills more than twice as many as cars in US



Most are preventable with AI



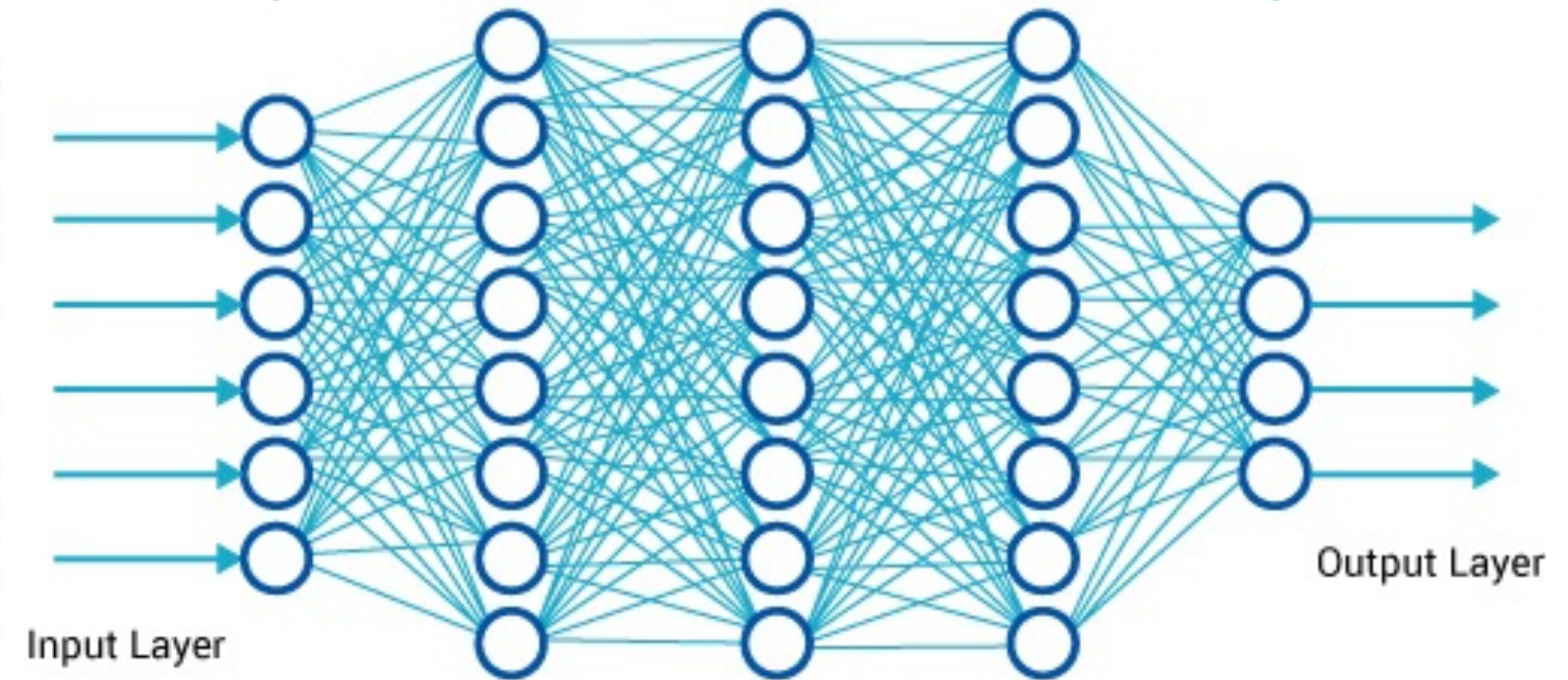
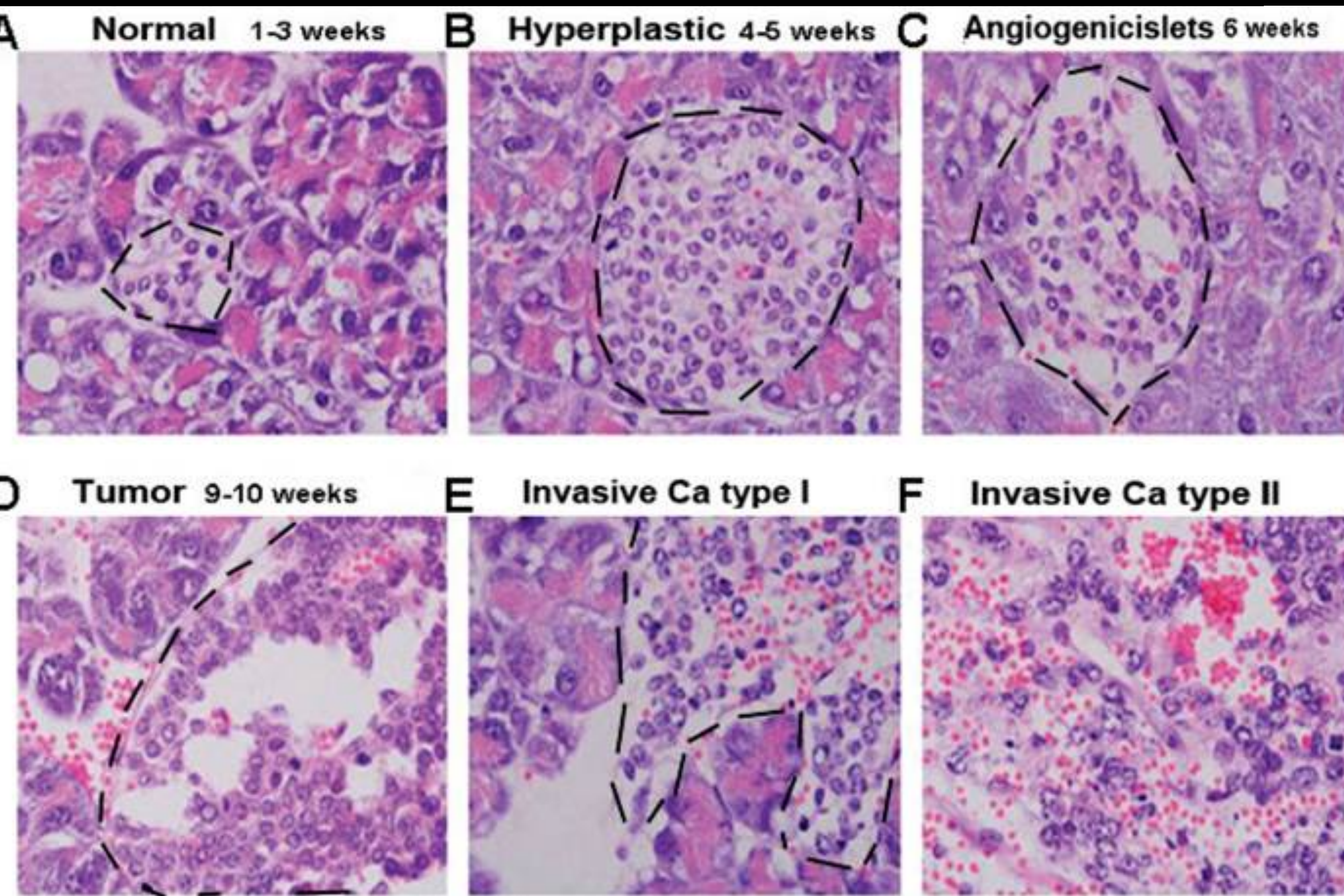
# AI diagnosis - prostate cancer



A recent Dutch study showed that AI diagnosis of prostate cancer using MRI was as good as that of human radiologists



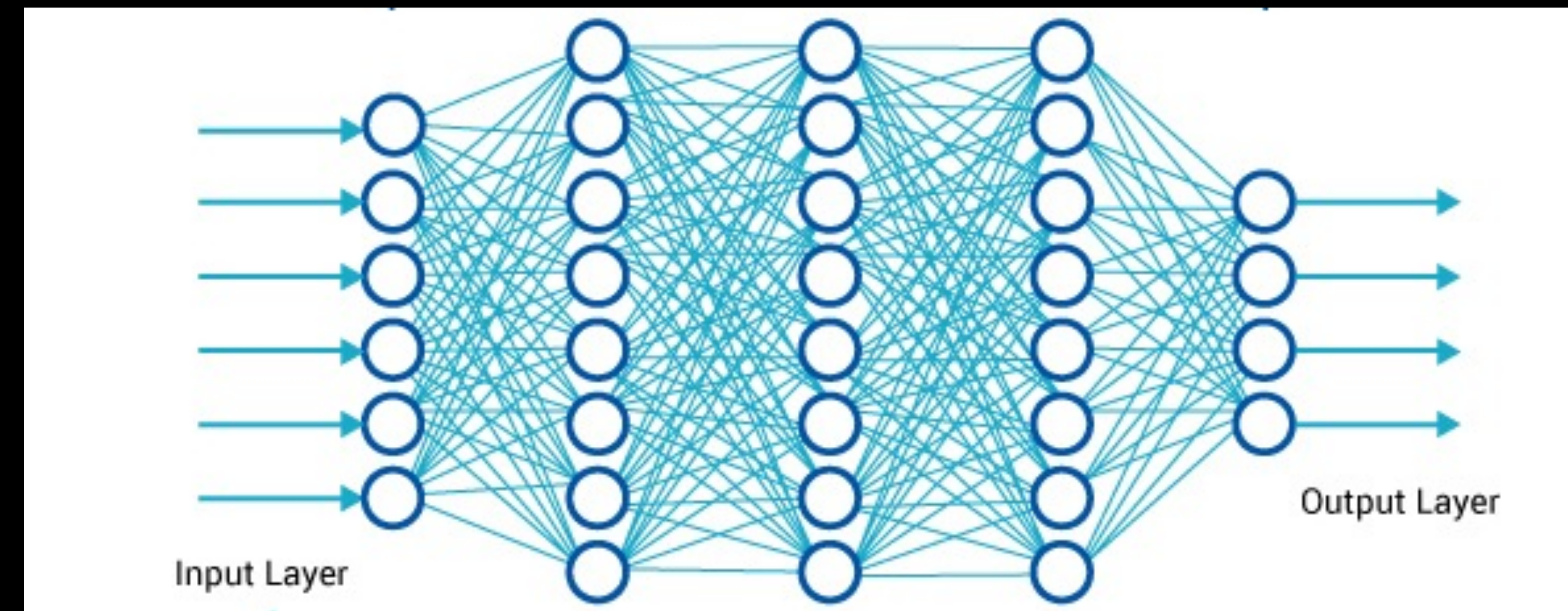
# AI diagnosis - lung cancer



A recent Stanford study showed that AI could diagnose lung cancer using microscope images even better than human pathologists



# AI diagnosis - blindness



A recent Google DeepMind study showed that AI could diagnose 50 retinal diseases from optical coherence tomography as human optalmologists



# AI diagnosis - ultrasound

Democratized by requiring

1) less money:



\$20k-\$75k



\$2k

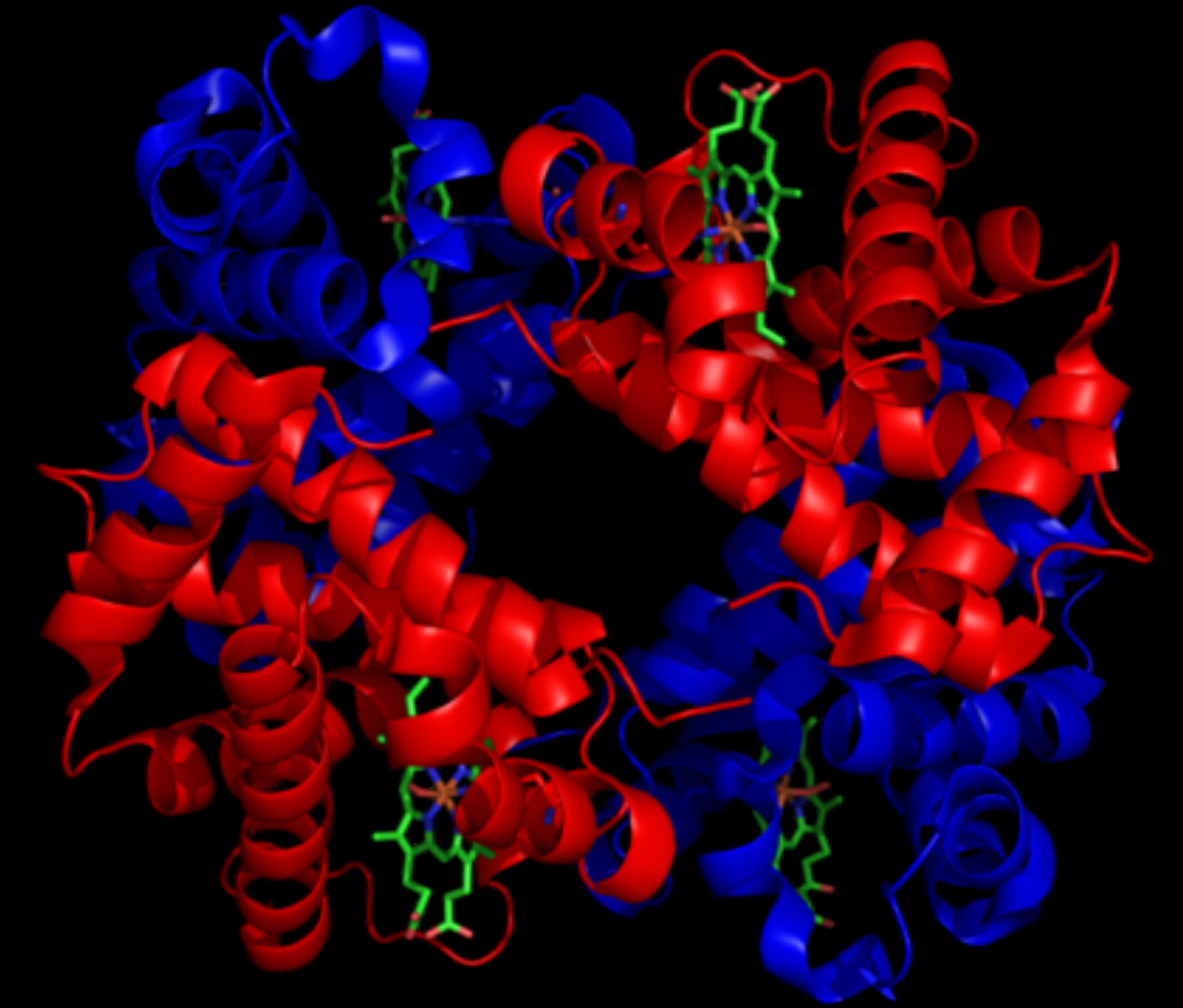
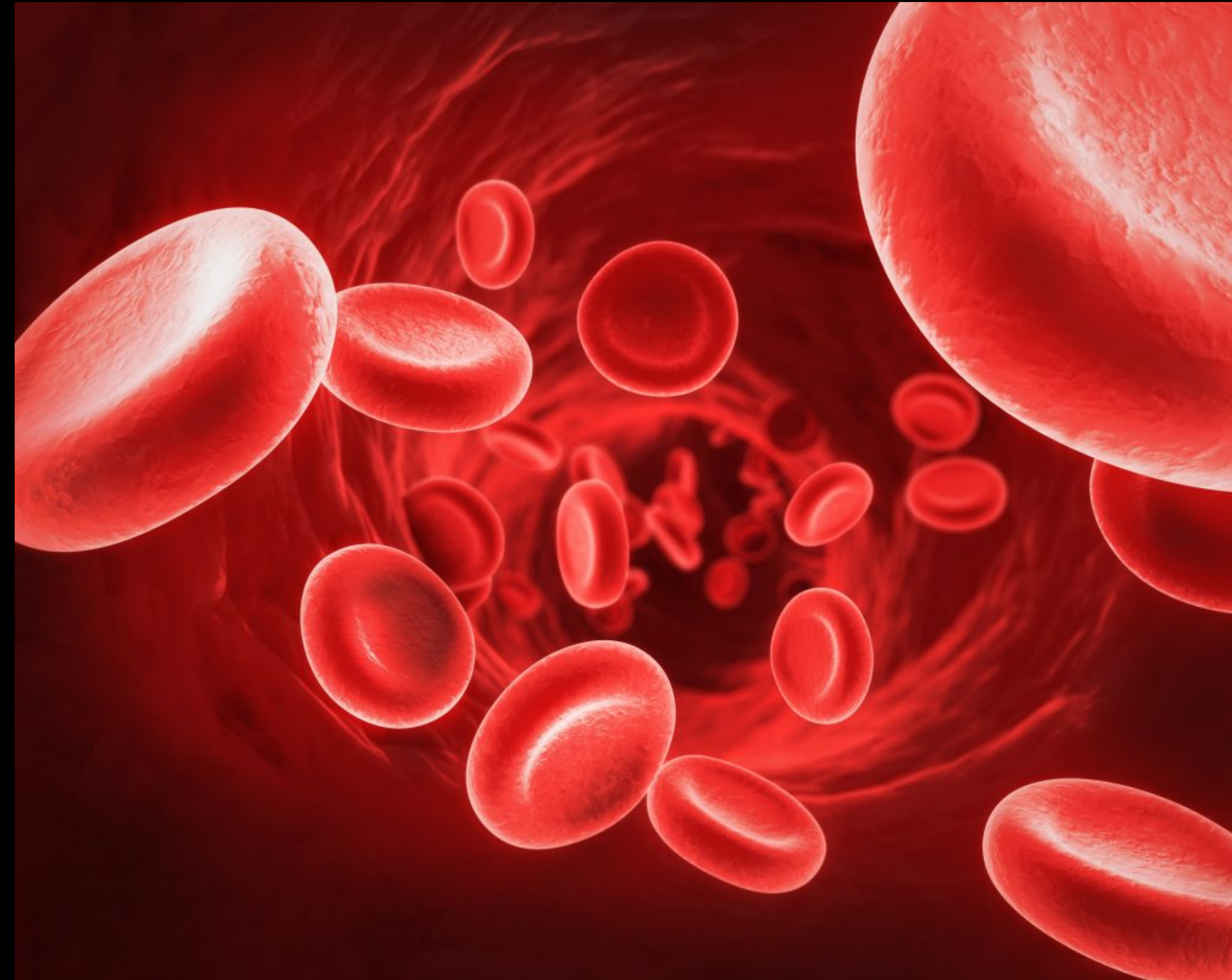
2) less training:

AI guides non-experts to place probe correctly & to analyze results

**NB. I have a conflict of interest, since I'm advisor to the company making the Butterfly IQ**



# AI for Protein Folding:



**Hemoglobin amino acid sequence:**  
MVLSPADK...



# AI for Protein Folding:

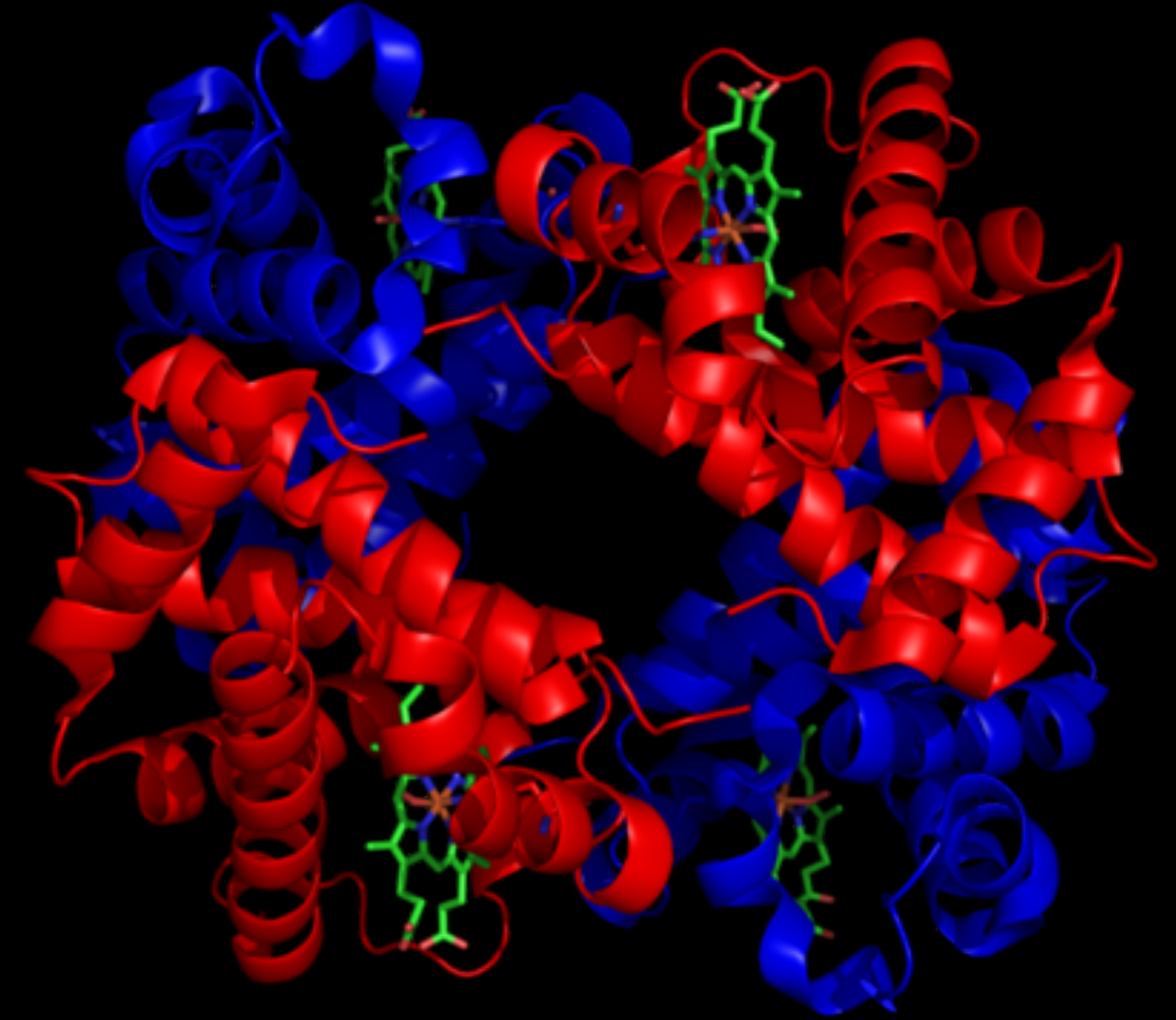
CASP13 competition results, Dec. 2018:

**43 mystery proteins**

**25 done best by AlphaFold AI**

**3 done best by #2 team of 98**

...and the AlphaFold team had almost no biology experience!

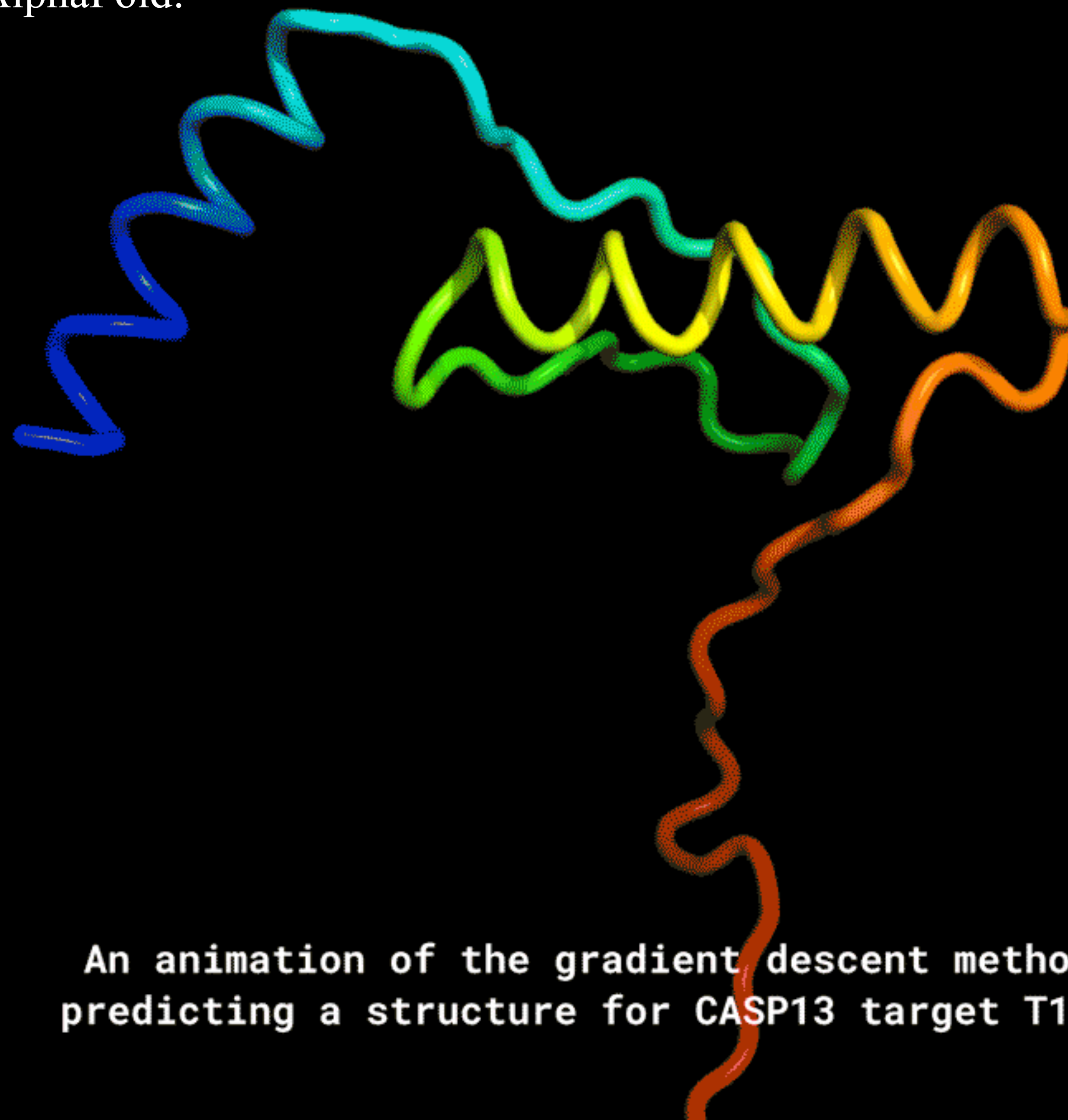




Input amino acid sequence:

TDELLERLRQLFEELHERGTEIVVEVHINGERDEIRVRNISKELKKLLERIREKIEREGSSEVEVNVHSGGQTWTFNEK

Protein shape predicted by AlphaFold:

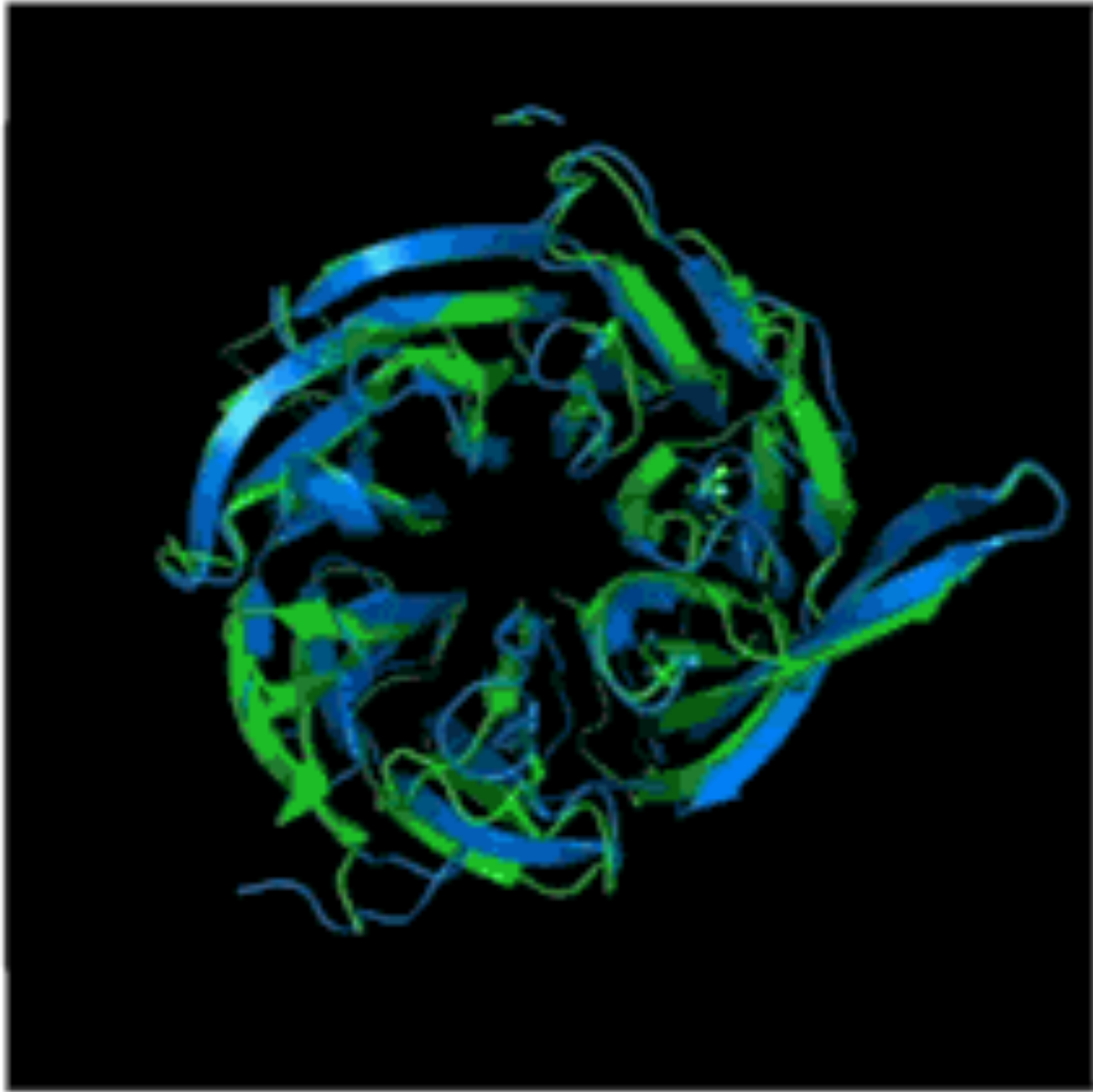


An animation of the gradient descent method  
predicting a structure for CASP13 target T1008

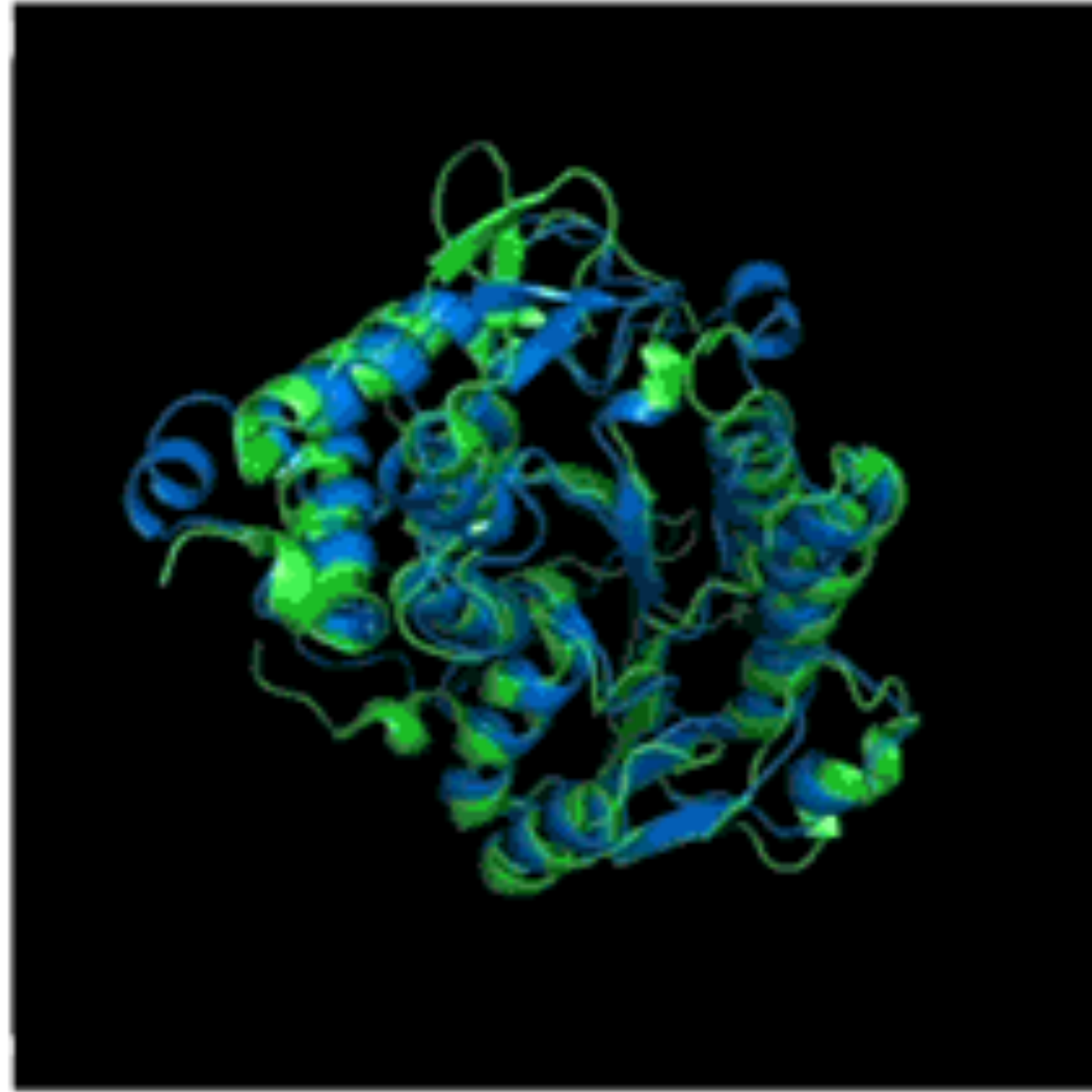


Structures:  
Ground truth (green)  
Predicted (blue)

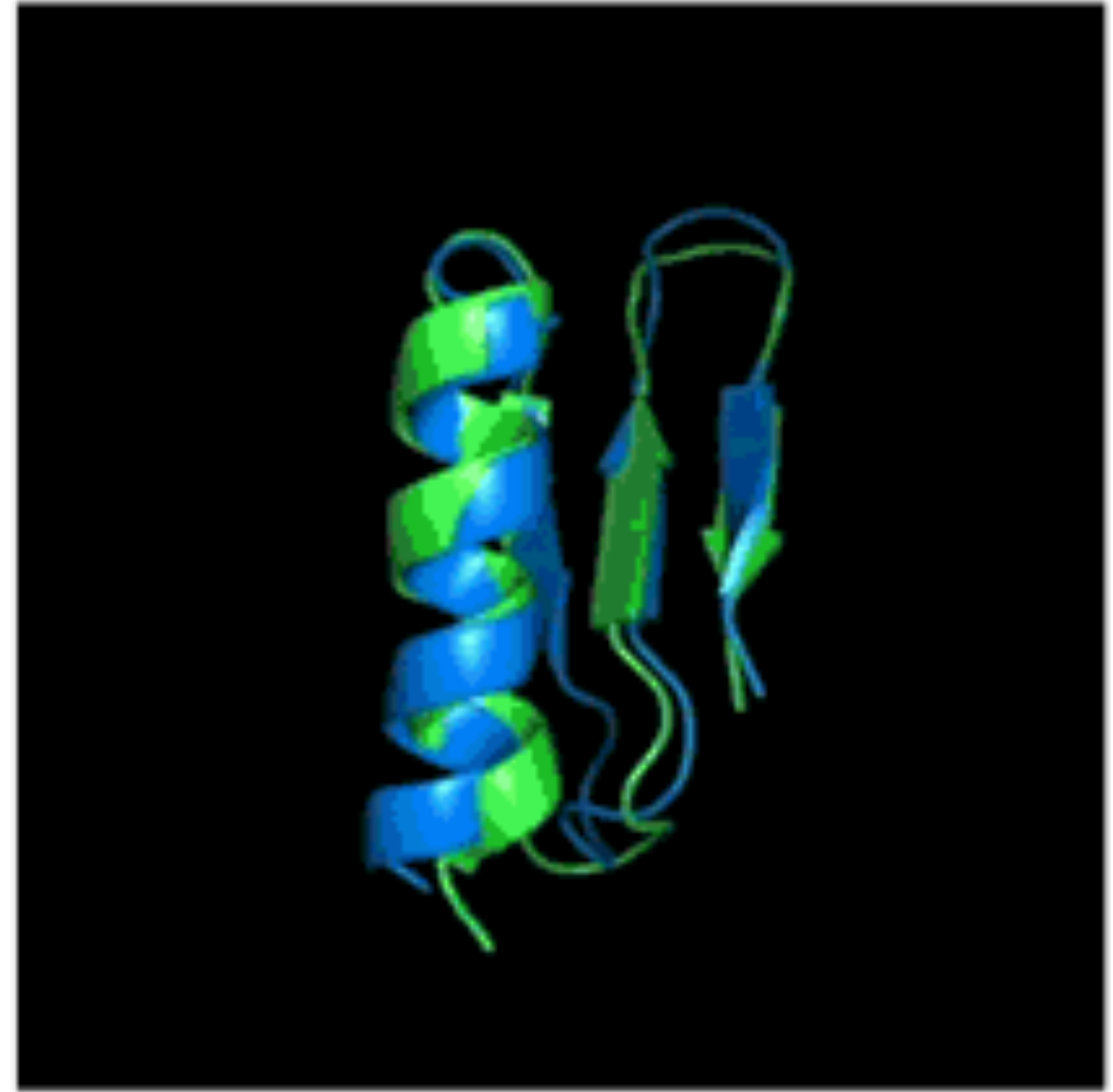
T0954 / 6CVZ



T0965 / 6D2V



T0955 / 5W9F





■ ■ ■ and AI can

cause

problems





POLITICS

## Here's how Cambridge Analytica played a dominant role in Kenya's chaotic 2017 elections

PUBLISHED FRI, MAR 23 2018 • 10:31 AM EDT







**CNN EXCLUSIVE**

**NEW DETAILS ON DRONE ATTACK AIMED AT VENEZUELAN PRESIDENT**



DOW ▲ 7.05

SITUATION ROOM





ALJAZEERA



# How Saudi Arabia failed to protect itself from drone and missile attacks despite billions spent on defense systems

PUBLISHED THU, SEP 19 2019 • 4:54 AM EDT | UPDATED MON, SEP 23 2019 • 6:35 AM EDT

 **Natasha Turak**  
@NATASHATURAK

SHARE     



Smoke is seen following a fire at Aramco facility in the eastern city of Abqaiq, Saudi Arabia, September 14, 2019.

OPINION - POLITICS

## After Iran’s strike on Saudi oil, US, China and Russia must join forces to prevent drone terror attacks

PUBLISHED TUE, SEP 24 2019 • 3:01 PM EDT

Carlos Pascual

SHARE     



# How Saudi Arabia failed to protect itself from drone and missile attacks despite billions spent on defense systems

PUBLISHED THU, SEP 19 2019 • 4:54 AM EDT | UPDATED MON, SEP 23 2019 • 6:35 AM EDT


 **Natasha Turak**  
@NATASHATURAK

SHARE     



Smoke is seen following a fire at Aramco facility in the eastern city of Abqaiq, Saudi Arabia, September 14, 2019.

## GIZMODO Man Accused of Using Drone to Airdrop Explosives Onto Ex-Girlfriend's Property

 **Jennings Brown**  
9/18/19 10:57am • Filed to: DRONES ▾

 12.4K  56  1    



OPINION - POLITICS

## After Iran’s strike on Saudi oil, US, China and Russia must join forces to prevent drone terror attacks

PUBLISHED TUE, SEP 24 2019 • 3:01 PM EDT

Carlos Pascual

SHARE     



# Knight Capital Says Trading Glitch Cost It \$440 Million

BY NATHANIEL POPPER   AUGUST 2, 2012 9:07 AM   356

Runaway Trades Spread Turmoil Across Wall St.



Errant trades from the Knight Capital Group began hitting the New York Stock Exchange almost as soon as the opening bell rang on Wednesday. Brendan McDermid/Reuters

1 of 4



4:01 p.m. | Updated

\$10 million a minute.



# How a 'confused' AI may have fought pilots attempting to save Boeing 737 MAX8s

Two Boeing 737 MAX airliners crashed, killing everyone aboard. Now investigations are zeroing in on one single faulty component.

Jamie Seidel

News Corp Australia Network 🕒 MARCH 19, 2019 2:24PM

📺 Video

📷 Image





# Every single Yahoo account was hacked - 3 billion in all

by Selena Larson @selenalarson

🕒 October 4, 2017: 6:36 AM ET



Timeline: The rise and fall of Yahoo

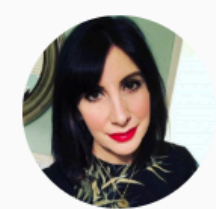


Sitting down? An epic and historic data breach at Yahoo in August 2013 affected every single customer account that existed at the time, Yahoo parent company Verizon [said](#) on Tuesday.



3,516 views | May 28, 2019, 08:30am

# Equifax Becomes First Firm To See Its Outlook Downgraded Due To A Cyber-Attack


 **Kate O'Flaherty** Contributor ⓘ  
[Cybersecurity](#)  
*I'm a freelance cyber security journalist.*



KIEV, UKRAINE - 2019/01/17: In this photo illustration, the Equifax Consumer reporting agency company logo seen displayed on a smartphone. (Photo Illustration by Igor Golovnirov/SOPA Images/LightRocket via Getty Images) GETTY

PRIVACY AND SECURITY

# Equifax Is Finally Getting Kicked in the Money Bags Due to Its Disastrous 2017 Hack

 **Patrick Howell O'Neill**  
5/23/19 12:30pm • Filed to: HACKING ∨


 20.5K  54  2



Mark Begor, CEO of Equifax, is sworn in during a Senate Homeland Security and Governmental Affairs Committee hearing on Capitol Hill, March 7, 2019 in Washington, DC. The committee heard testimony on investigations examining private sector data breaches. Photo: Mark Wilson (Getty)



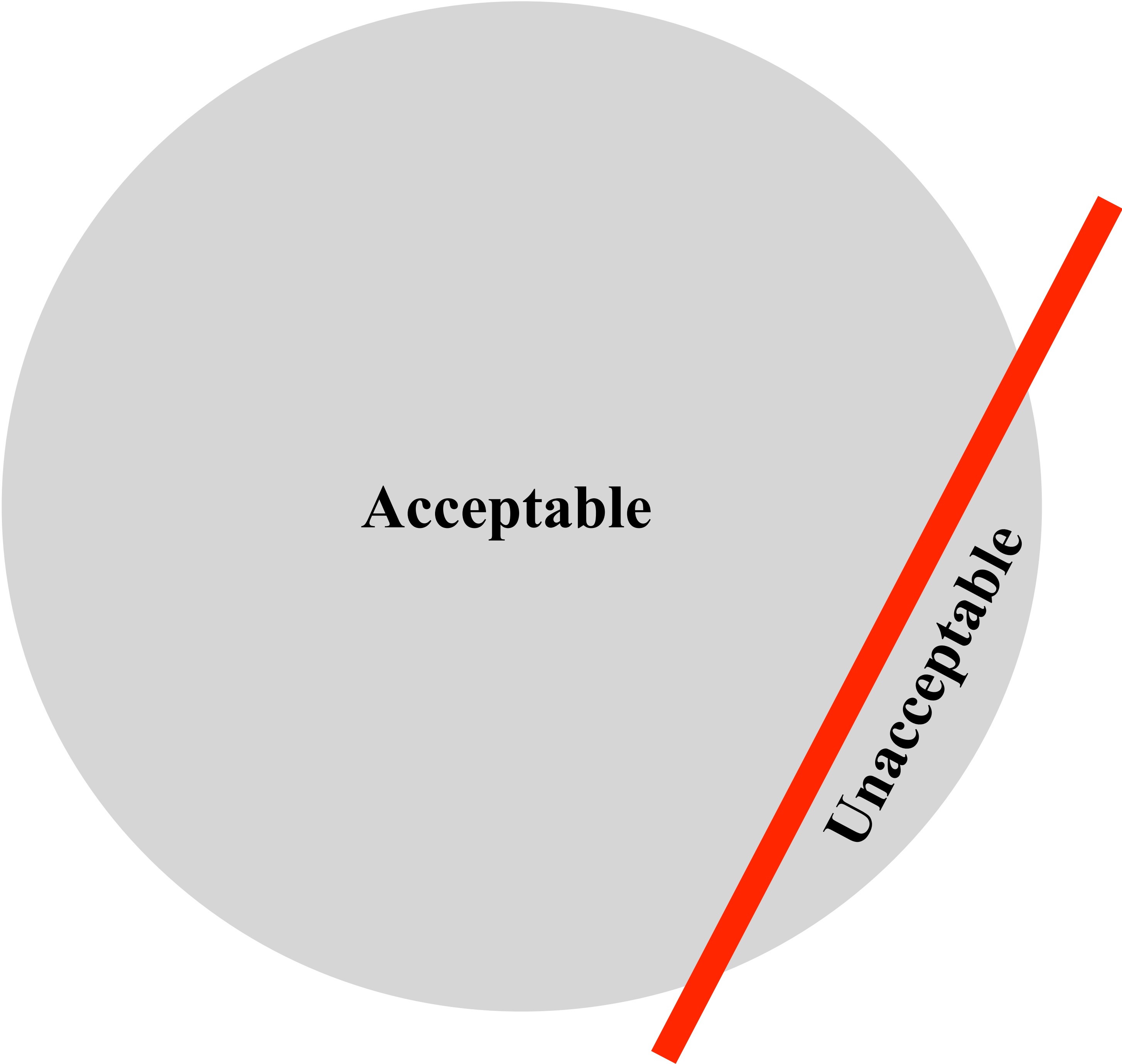
A medium shot of Barack Obama sitting in the Oval Office. He is wearing a dark suit, a white shirt, and a blue patterned tie. He is looking slightly to his right with a serious expression. In the background, the American flag is visible on the left, and the Presidential Seal is on the wall behind him. A large, dark, upholstered chair is partially visible on the right.

The end product  
is called a  
“deepfake”



**Let's draw  
the line!**



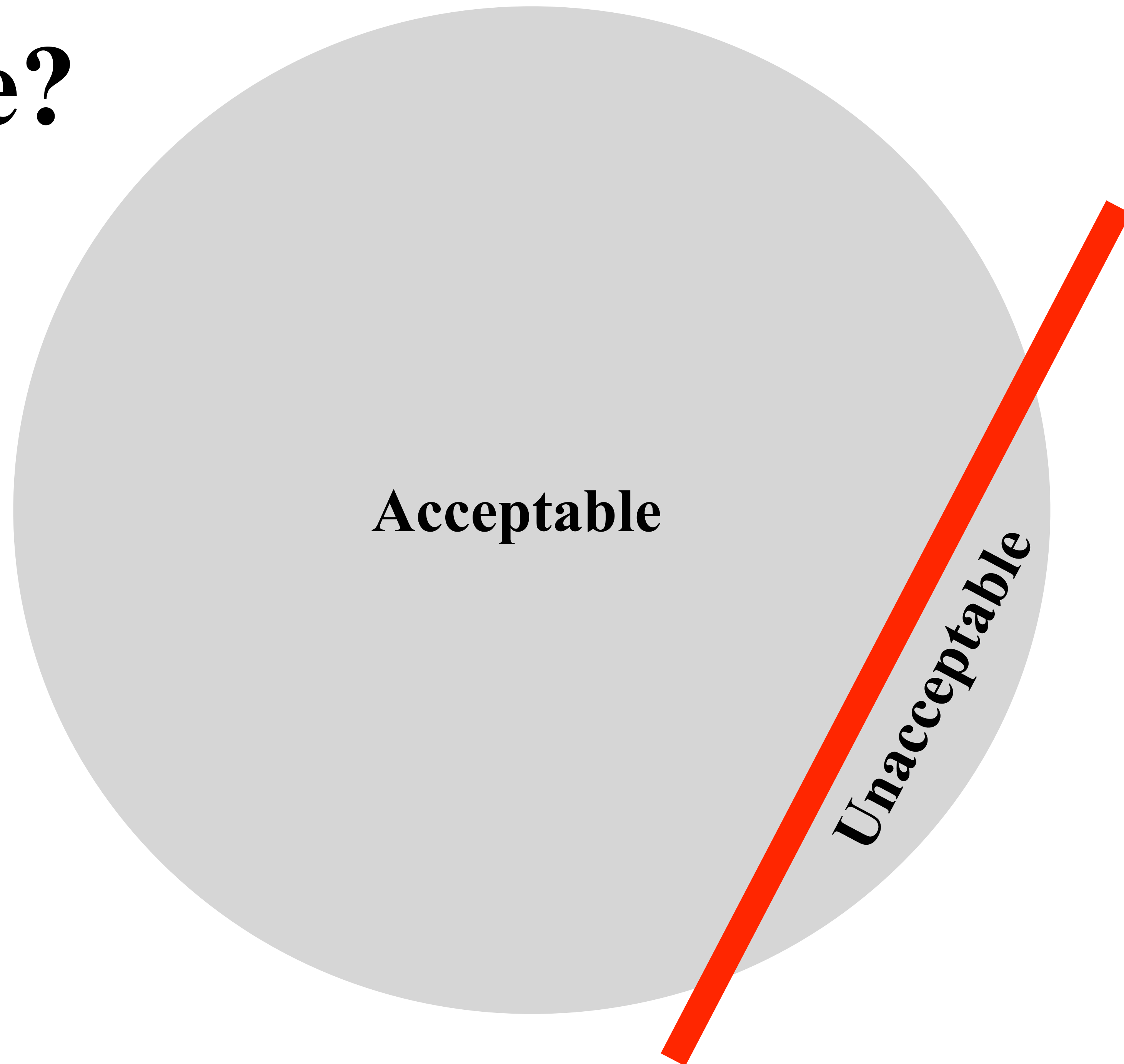


**Acceptable**

***Unacceptable***



# Impossible?





# Impossible?

Nope!



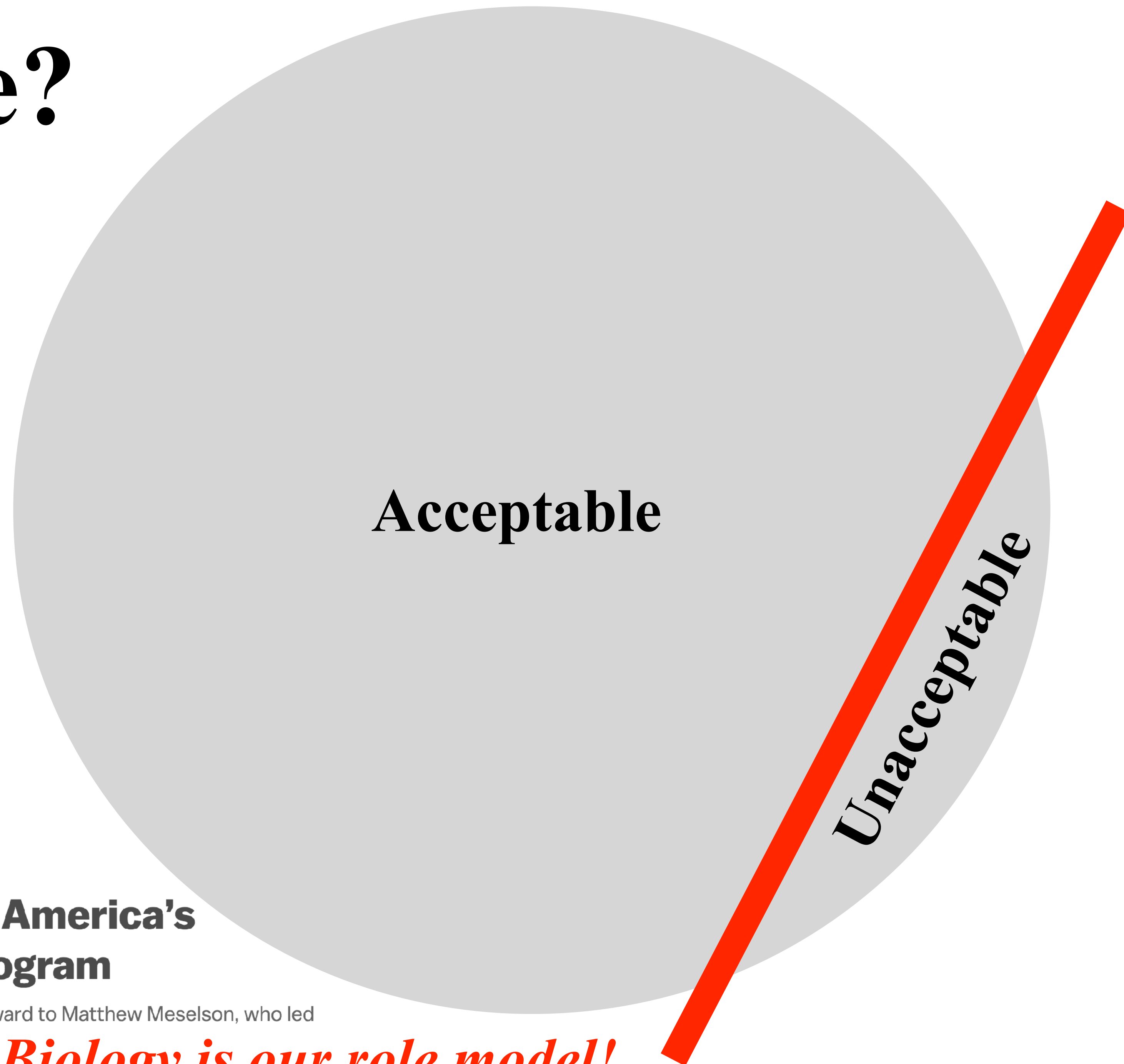
*Prof. Matthew Meselson*

**Vox** The man who stopped America's biological weapons program

The Future of Life Institute will give its annual award to Matthew Meselson, who led the campaign against bioweapons.

By Kelsey Piper | Apr 9, 2019, 12:40pm EDT

***Biology is our role model!***



Acceptable

Unacceptable



# Uses of biology:



*Bioweapons*



*Biology is our role model!*



Asilomar 1975:



Asilomar 2017:







# Signatories of the 23 Asilomar AI Principles

**Demis Hassabis**, DeepMind, Founder & CEO

**Ilya Sutskever**, OpenAI, Research Director, Co-Founder

**Yann LeCun**, Director of AI Research at Facebook, Professor at New York University

**Yoshua Bengio**, University of Montréal, head of the Montreal Institute for Learning Algorithms

**Stuart Russell**, University of California, Berkeley, Professor of Computer Science, director of the Center for Intelligent Systems, co-author of the standard textbook Artificial Intelligence: a Modern Approach.

**Peter Norvig**, Google, Research Director

**Ray Kurzweil**, Google, Research Director; Inventor, Author and Futurist, author of The Singularity is Near and How to Create a Mind

**Jeff Dean**, Google, Inc., Leader of Google Brain project, Google Senior Fellow, member of AAAI, ACM, National Academy of Engineering

**Tom Gruber**, Apple; Co-Founder, CTO and VP Design for Siri

**Nils J. Nilsson**, Past President of AAAI, Emeritus Kumagai Professor of Engineering, Stanford, member of AAAI

**Francesca Rossi**, IBM & Padova, Professor of Computer Science, IJCAI President and Co-chair of AAAI committee on impact of AI and Ethical Issues, member of ACM

**Bart Selman**, Cornell, Professor of Computer Science, co-chair of the AAAI presidential panel on long-term AI futures, member of ACM

**Guru Banavar**, Vice President, Cognitive Computing, IBM Research

**Shane Legg**, DeepMind, Co-Founder

**Mustafa Suleyman**, DeepMind, Co-Founder

**Leslie Kaelbling**, MIT, Professor of Computer Science and Engineering, member of AAAI, ACM, IEEE RAS

...





# Signatories of the 23 Asilomar AI Principles

**Demis Hassabis**, DeepMind, Founder & CEO

**Ilya Sutskever**, OpenAI, Research Director, Co-Founder

**Yann LeCun**, Director of AI Research at Facebook, Professor at New York University

**Yoshua Bengio**, University of Montréal, head of the Montreal Institute for Learning Algorithms

**Stuart Russell**, University of California, Berkeley, Professor of Computer Science, director of the Center for Human-Compatible AI, co-author of the standard textbook Artificial Intelligence: a Modern Approach.

**Peter Norvig**, Google, Research Director

**Ray Kurzweil**, Google, Research Director; Inventor, Author and Futurist, author of The Singularity is Near

**Jeff Dean**, Google, Inc., Leader of Google Brain project, Google Senior Fellow, member of AAAI, ACM, and IEEE

**Tom Gruber**, Apple; Co-Founder, CTO and VP Design for Siri

**Nils J. Nilsson**, Past President of AAAI, Emeritus Kumagai Professor of Engineering, Stanford, member of AAAI, ACM, and IEEE

**Francesca Rossi**, IBM & Padova, Professor of Computer Science, IJCAI President and Co-chair of AAAI on AI and Ethical Issues, member of ACM

**Bart Selman**, Cornell, Professor of Computer Science, co-chair of the AAAI presidential panel on long-term AI consequences, member of AAAI and ACM

**Guru Banavar**, Vice President, Cognitive Computing, IBM Research

**Shane Legg**, DeepMind, Co-Founder

**Mustafa Suleyman**, DeepMind, Co-Founder

**Leslie Kaelbling**, MIT, Professor of Computer Science and Engineering, member of AAAI, ACM, IEEE RAS

...

**Endorsed by  
California  
Aug 30 2018**





# Signatories of the 23 Asilomar AI Principles

**Demis Hassabis**, DeepMind, Founder & CEO

**Ilya Sutskever**, OpenAI, Research Director, Co-Founder

**Yann LeCun**, Director of AI Research at Facebook, Professor at New York University

**Yoshua Bengio**, University of Montréal, head of the Montreal Institute for Learning Algorithms

**Stuart Russell**, University of California, Berkeley, Professor of Computer Science, director of the Center for Intelligent Systems, co-author of the standard textbook Artificial Intelligence

**Peter Norvig**, Google, Research Director

**Ray Kurzweil**, Google, Research Director, futurist, author of The Singularity is Near and How to Create a Mind

**Jeff Dean**, Google, Inc., Leader of Google Brain, Senior Fellow, member of AAAI, ACM, National Academy of Engineering

**Tom Gruber**, Apple; Co-Founder, Core Knowledge

**Nils J. Nilsson**, Past President of AAAI, Professor of Engineering, Stanford, member of AAAI

**Francesca Rossi**, IBM & Padova, IJCAI President and Co-chair of AAAI committee on impact of AI and Ethical Issues, member of ACM

**Bart Selman**, Cornell, Professor of Computer Science, Co-chair of the AAAI presidential panel on long-term AI futures, member of ACM

**Guru Banavar**, Vice President, Cognitive Computing, IBM Research

**Shane Legg**, DeepMind, Co-Founder

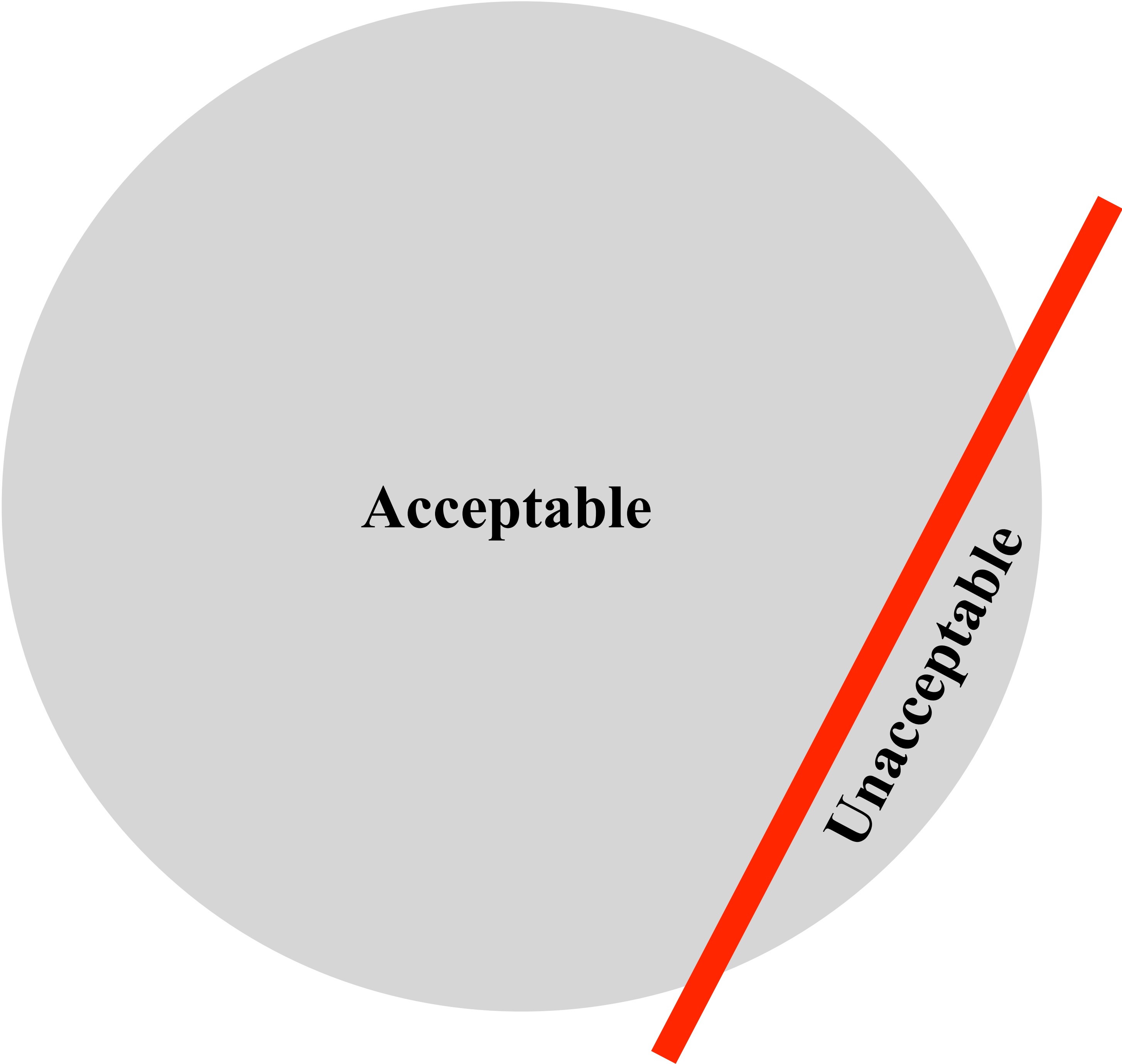
**Mustafa Suleyman**, DeepMind, Co-Founder

**Leslie Kaelbling**, MIT, Professor of Computer Science and Engineering, member of AAAI, ACM, IEEE RAS

...

**Terah Lyons  
from PAI  
speaks  
this PM**





**Acceptable**

***Unacceptable***



**Let's build AI that  
overpowers us!**





The background of the image is a deep space photograph showing a vast field of galaxies and stars. In the lower-left foreground, there is a grayscale image of the statue 'The Thinker' by Auguste Rodin, depicting a man in a state of deep contemplation. A white speech bubble with a black border originates from the statue's head.

Let's build AI that  
~~overpowers~~ us!  
empowers







