# The impact of AI on research now and in the next 5 years

**13th ICA Rectors and Deans Forum**
**19th October 2023**

Alistair Nolan, Senior Policy Analyst, Directorate for Science and Technology Policy, OECD

# Artificial Intelligence in Science

## CHALLENGES, OPPORTUNITIES AND THE FUTURE OF RESEARCH

OECD

# Today’s presentation

**Is science getting harder ?**

**AI in science today and tomorrow**

**Impacts of AI in science so far**

**Public policy and universities**

# Is science getting harder ?

# And recent attention to the productivity of research spurred by

the NATIONAL BUREAU of ECONOMIC RESEARCH

## Are Ideas Getting Harder to Find?

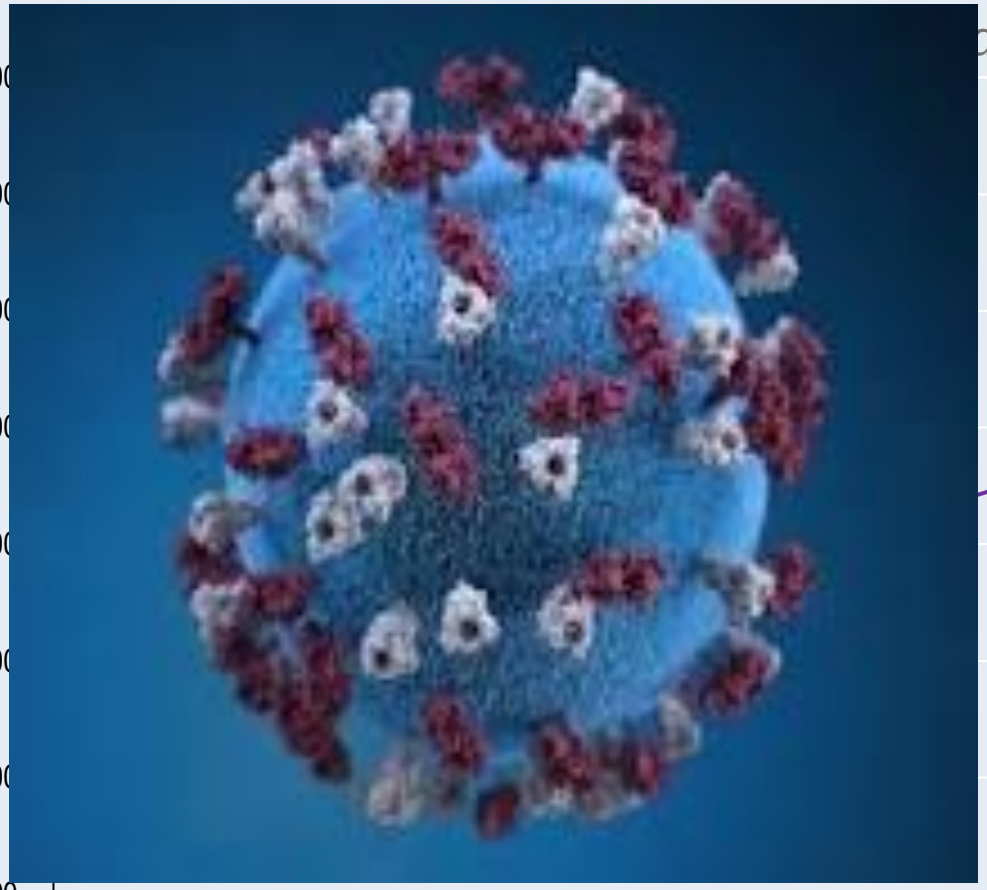
Nicholas Bloom, Charles I. Jones, John Van Reenen, Michael Webb

**NBER Working Paper No. 23782**
**Issued in September 2017**
**NBER Program(s):** Economic Fluctuations and Growth, Productivity, Innovation, and Entrepreneurship

In many growth models, economic growth arises from people creating ideas, and the long-run growth rate is the product of two terms: the effective number of researchers and their research productivity. We present a wide range of evidence from various industries, products, and firms showing that research effort is rising substantially while research productivity is declining sharply. A good example is Moore's Law. The number of researchers required today to achieve the famous doubling every two years of the density of computer chips is more than 18 times larger than the number required in the early 1970s. Across a broad range of case studies at various levels of (dis)aggregation, we find that ideas — and in particular the exponential growth they imply — are getting harder and harder to find. Exponential growth results from the large increases in research effort that offset its declining productivity.

# Information overload

(annual number of scientific publications, 1996-2018)



**100,000 scientific papers on coronavirus in year 1 of the pandemic**

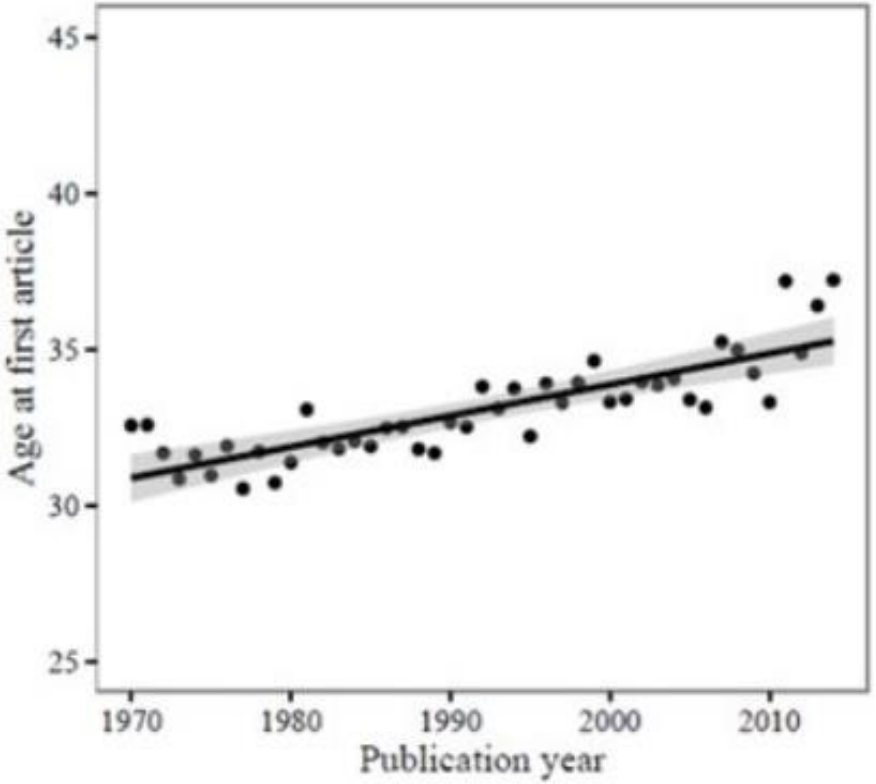
4,000,000
3,500,000
3,000,000
2,500,000
2,000,000
1,500,000
1,000,000
500,000
0

1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018

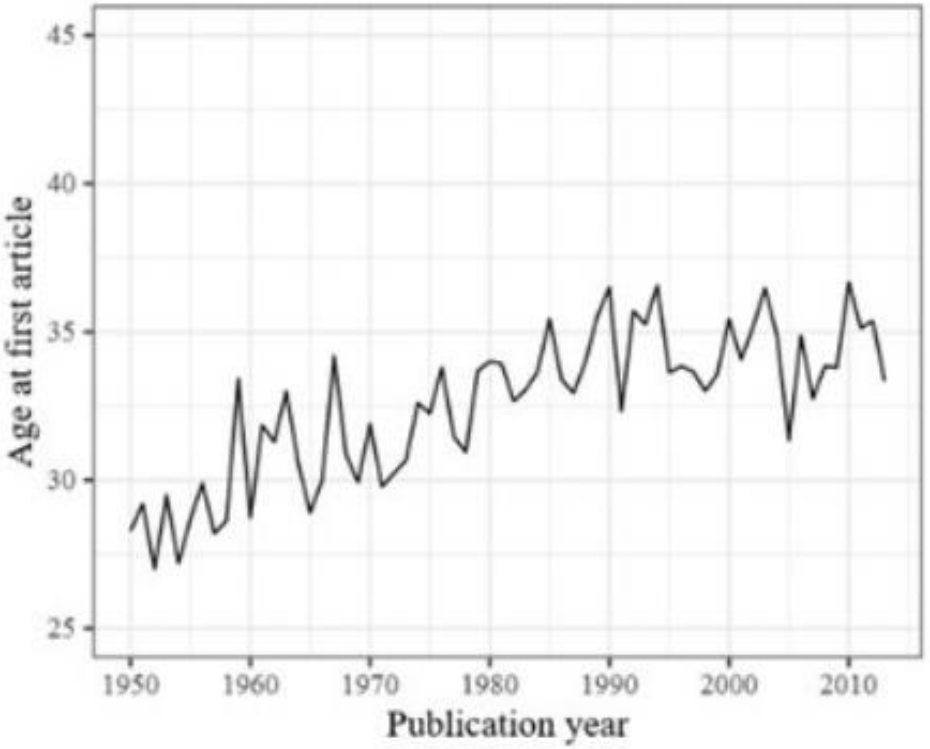
Scientific publications

# Creating a knowledge burden ?

Age at first solo economics article




Age at first solo (top) mathematics article




Schweitzer and Brendel (2020) Brendel and Schweitzer (2019)

# Discovery getting harder ?

$$F = m \times a$$

**1686**

$$\ln \frac{K_2}{K_1} = \frac{-\Delta H^{\emptyset}}{R} \left( \frac{1}{T_2} - \frac{1}{T_1} \right)$$

**1884**

$$\left(1 - e^{-2\Delta}\right) r^{D-3} = \frac{2\kappa}{D-2} \int_0^r \rho(r') r'^{D-2} \, dr' = \frac{2 \cdot 8\pi (D-3) G}{(D-2)} \frac{M}{\Omega_{D-2}} \Rightarrow$$

$$\Rightarrow \frac{1}{3} \frac{4 \left[ anti\log \frac{\int_0^{\infty} \frac{\cos \pi t x w'}{\cosh \pi x} e^{-\pi x^2 w'} dx}{e^{-\frac{\pi^2}{4} w'} \varphi_{w'}(itw')} \right] \cdot \frac{\sqrt{142}}{t^2 w'}}{\log \left[ \sqrt{\left( \frac{10 + 11\sqrt{2}}{4} \right)} + \sqrt{\left( \frac{10 + 7\sqrt{2}}{4} \right)} \right]} . \quad (2.93c)$$

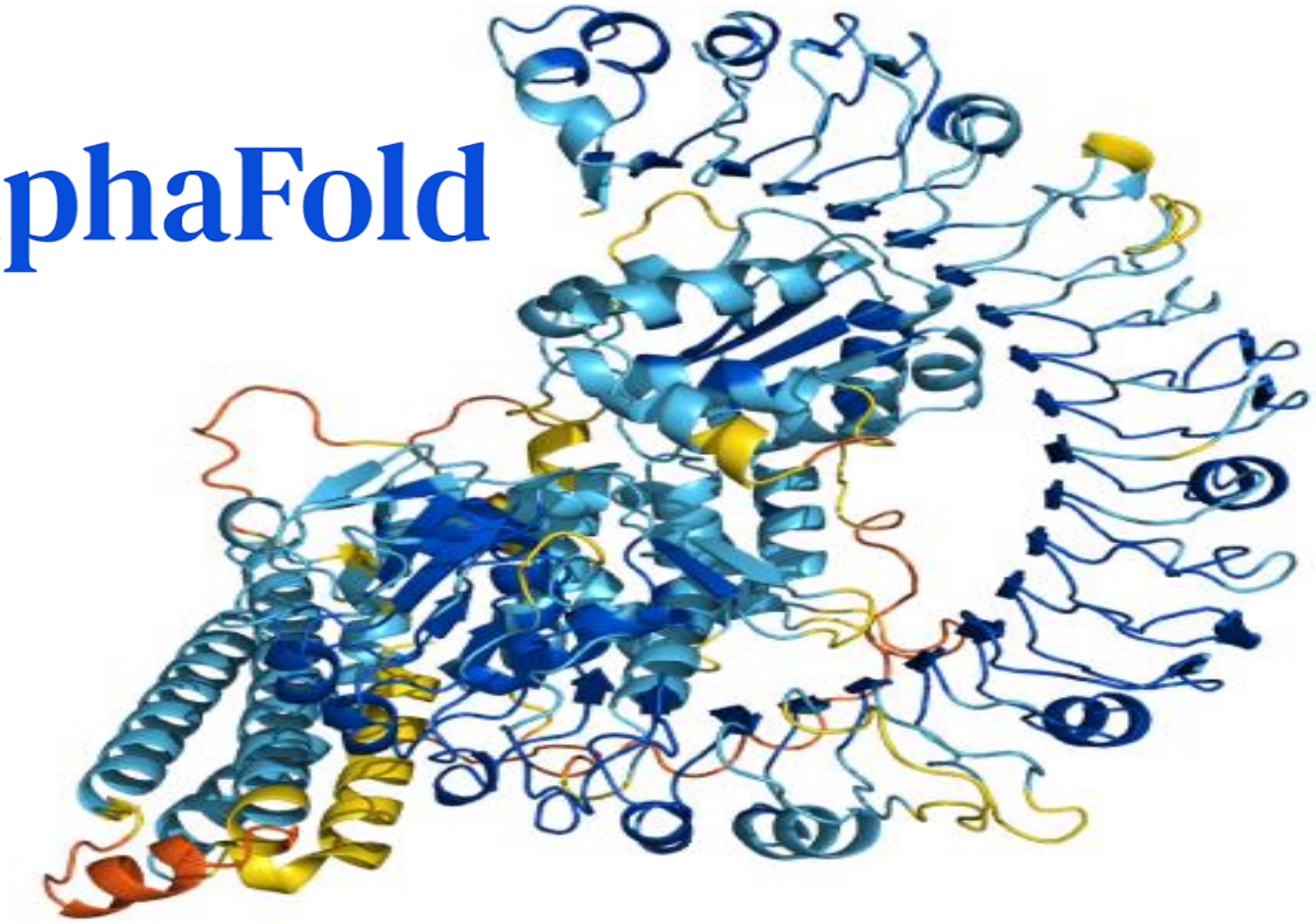
**1973**

Thanks to Prof. Hugh Cartwright, Oxford University

# AI : Coming to scientific knowledge in new ways

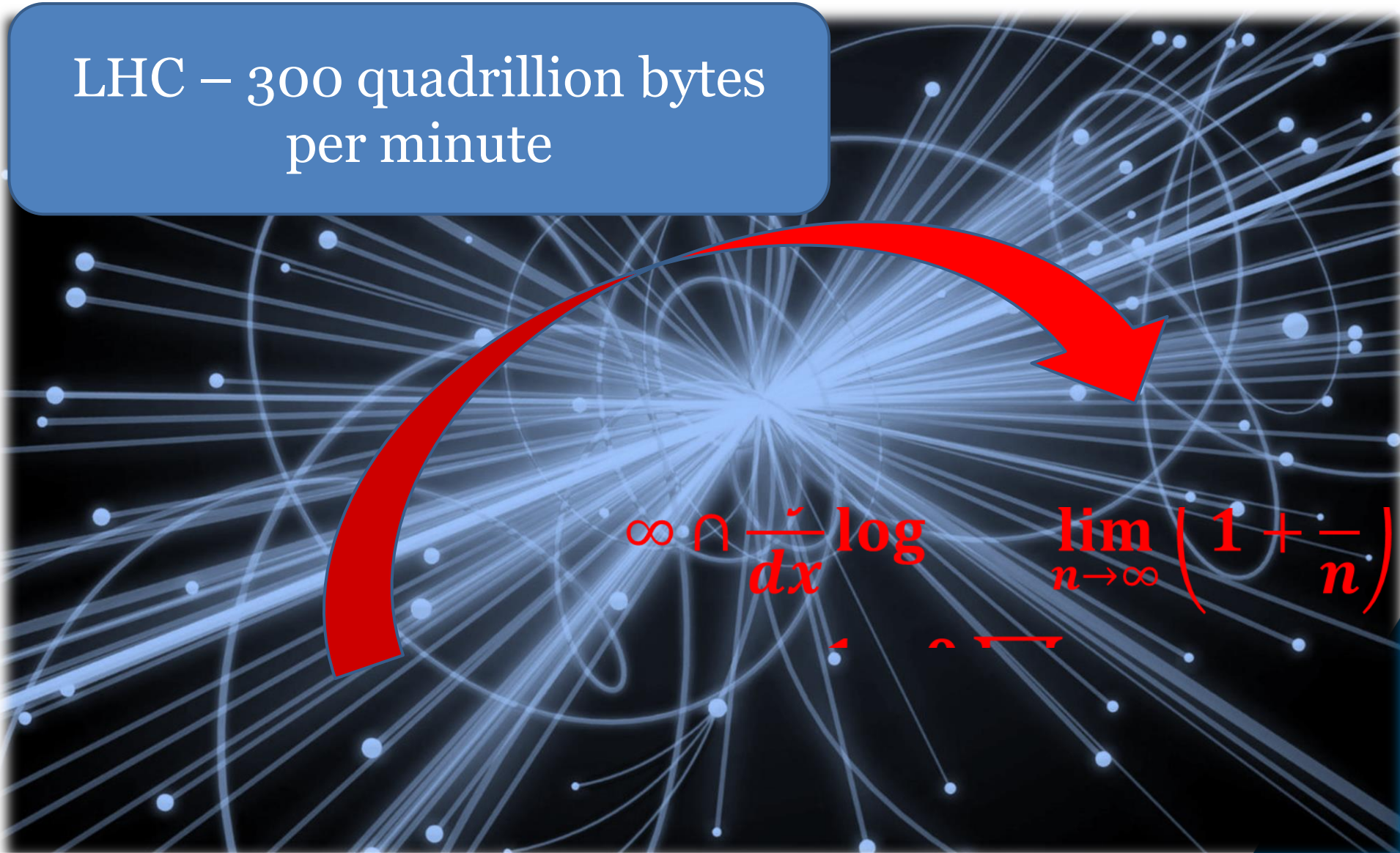
# Prediction

AlphaFold

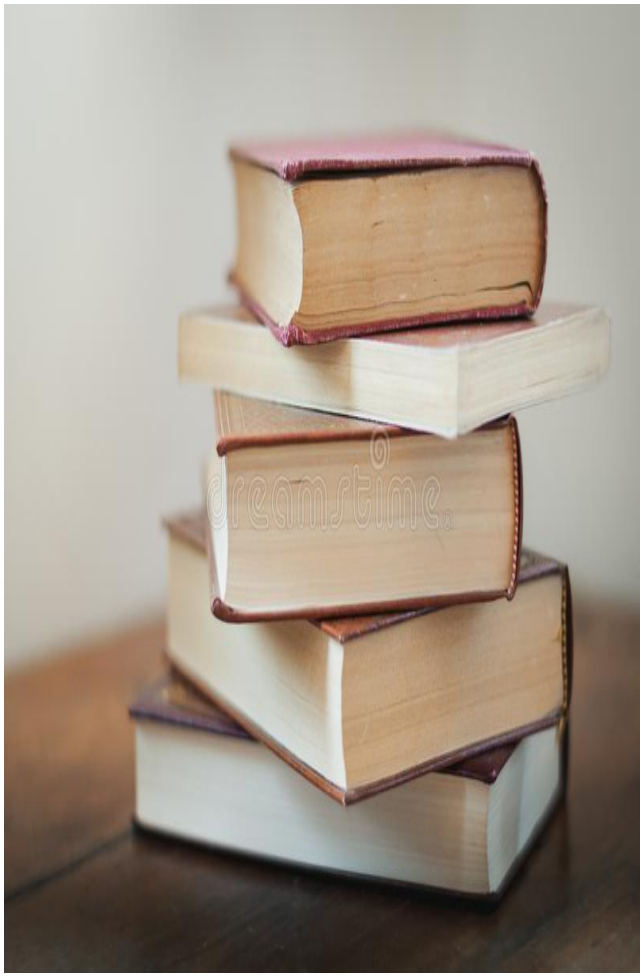


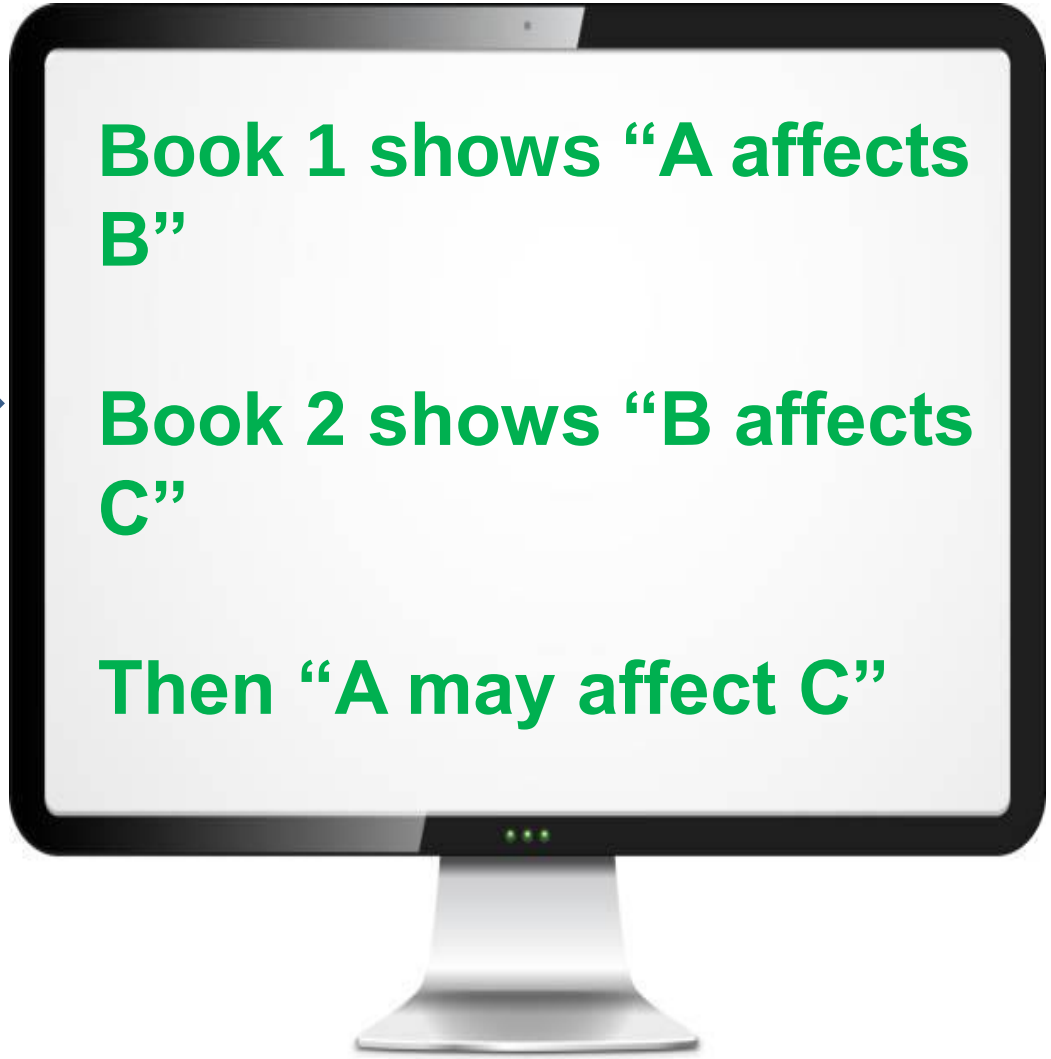
# Generating hypotheses from vast datasets

LHC – 300 quadrillion bytes per minute



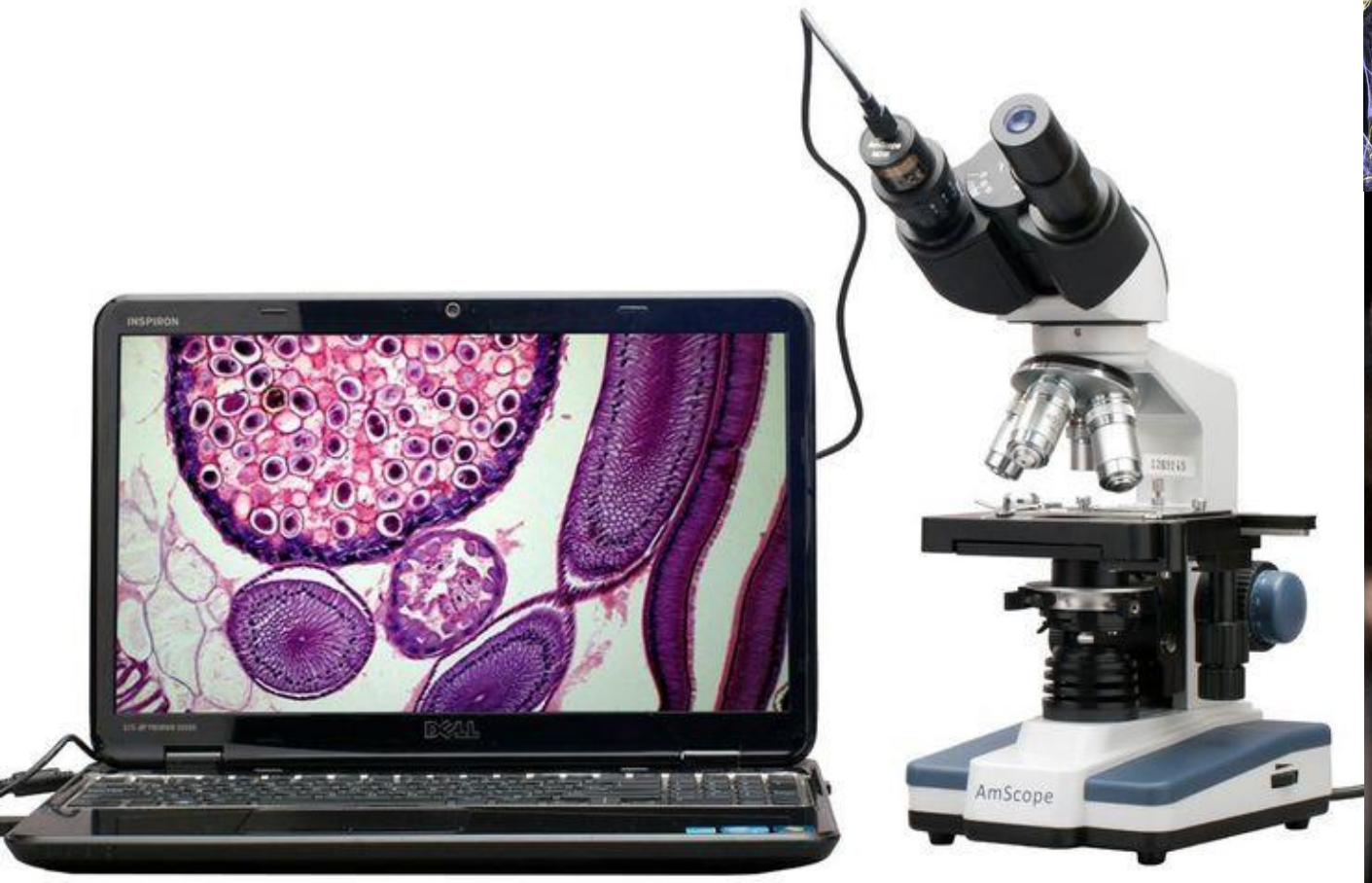
# Finding undiscovered public knowledge (knowledge we don't know we have)

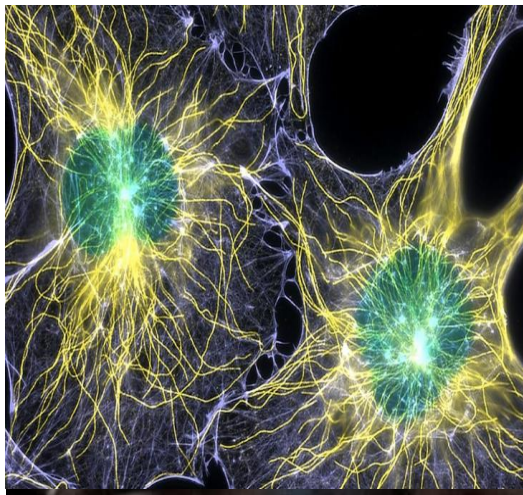


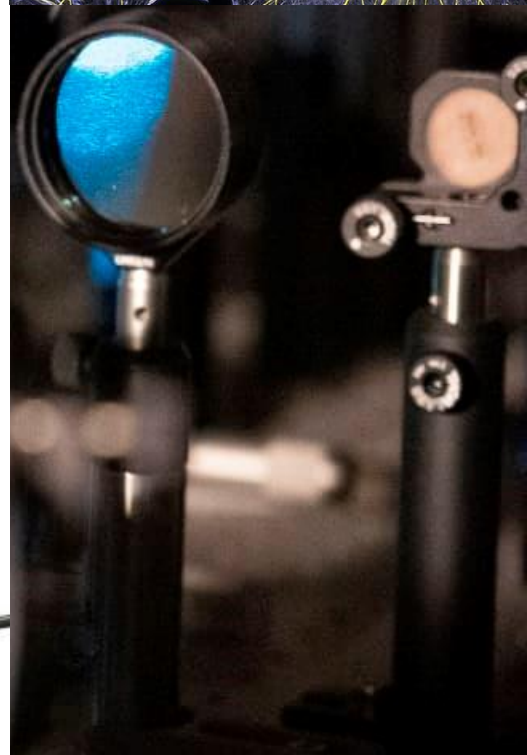


# Novel simulation

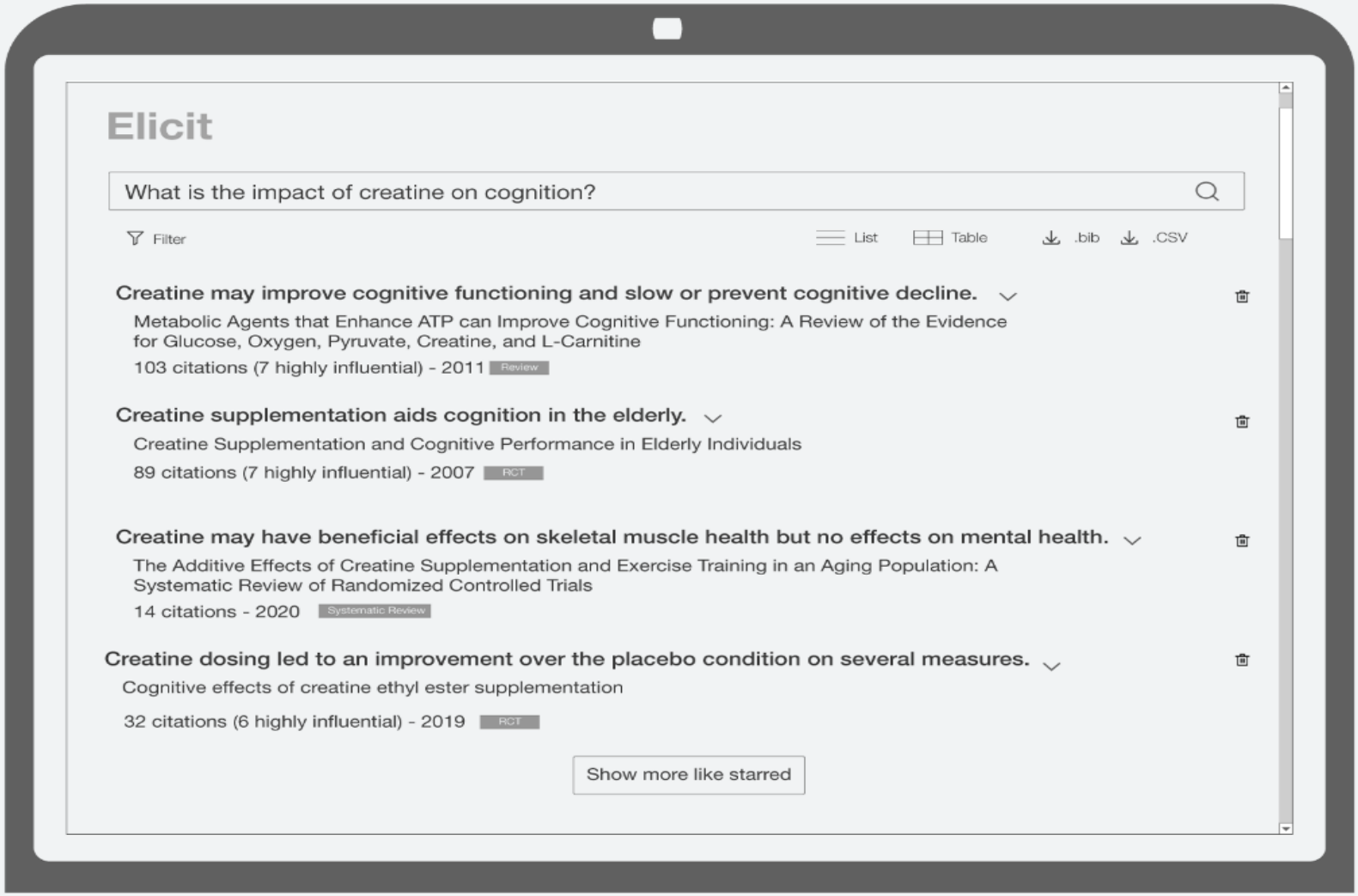
# Revolutionising microscopy





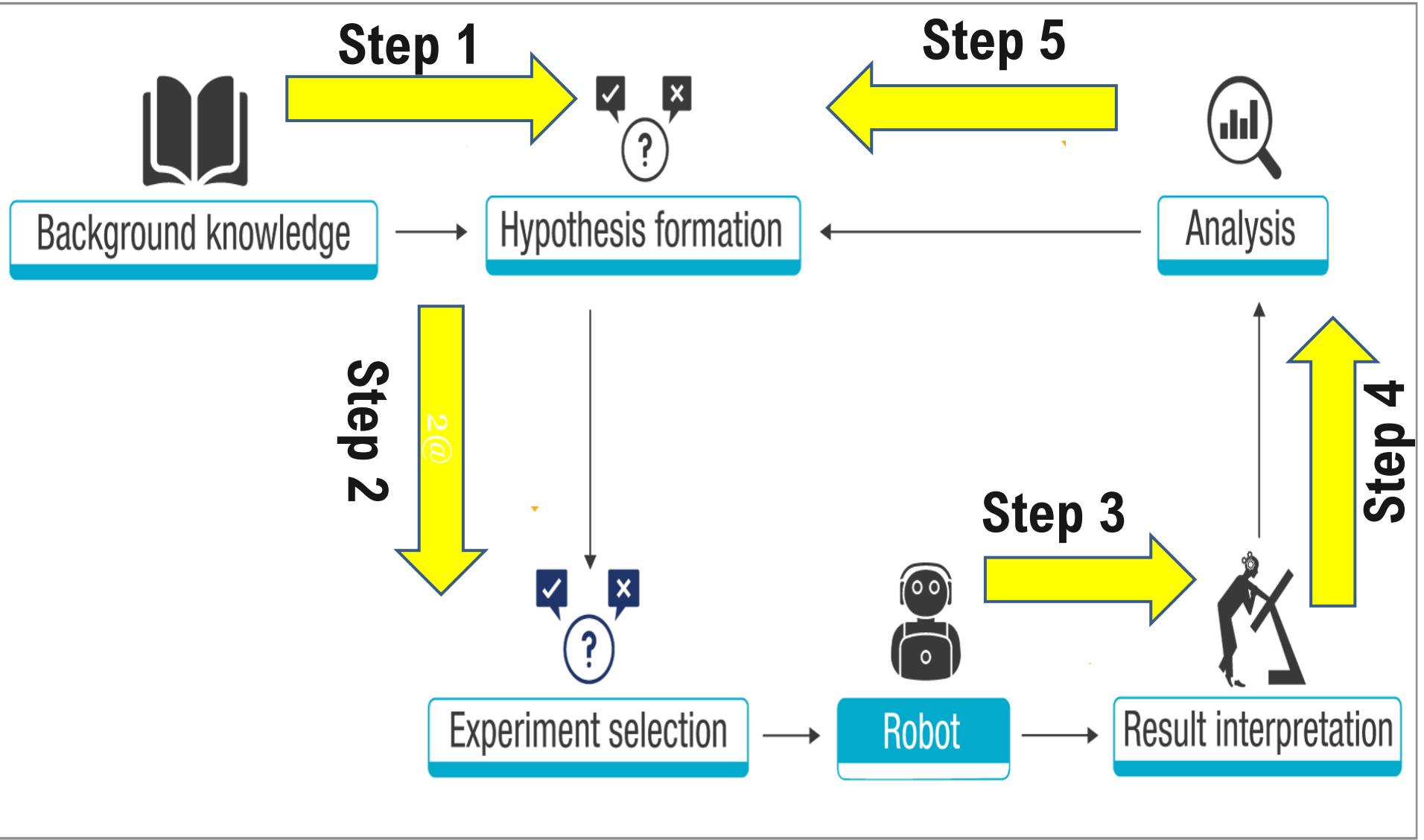
# *Elicit – (Ought.com)* - AI Research assistant – using GPT3



# Many other possible AI applications


**Planning experiments**

# Robot scientists
# Closed-loop cycles of experimentation



# Professor Ross King in front of Adam, the robot scientist

**Triclosan – works against wild-type and drug resistant Plasmodium falciparum, and Plasmodium vivax.**

2008-2015 Eve – Drug Design for Tropical Diseases

**Williams *et al*. (2015) *Royal Society Interface*, DOI 10.1098/rsif.2014.1289**

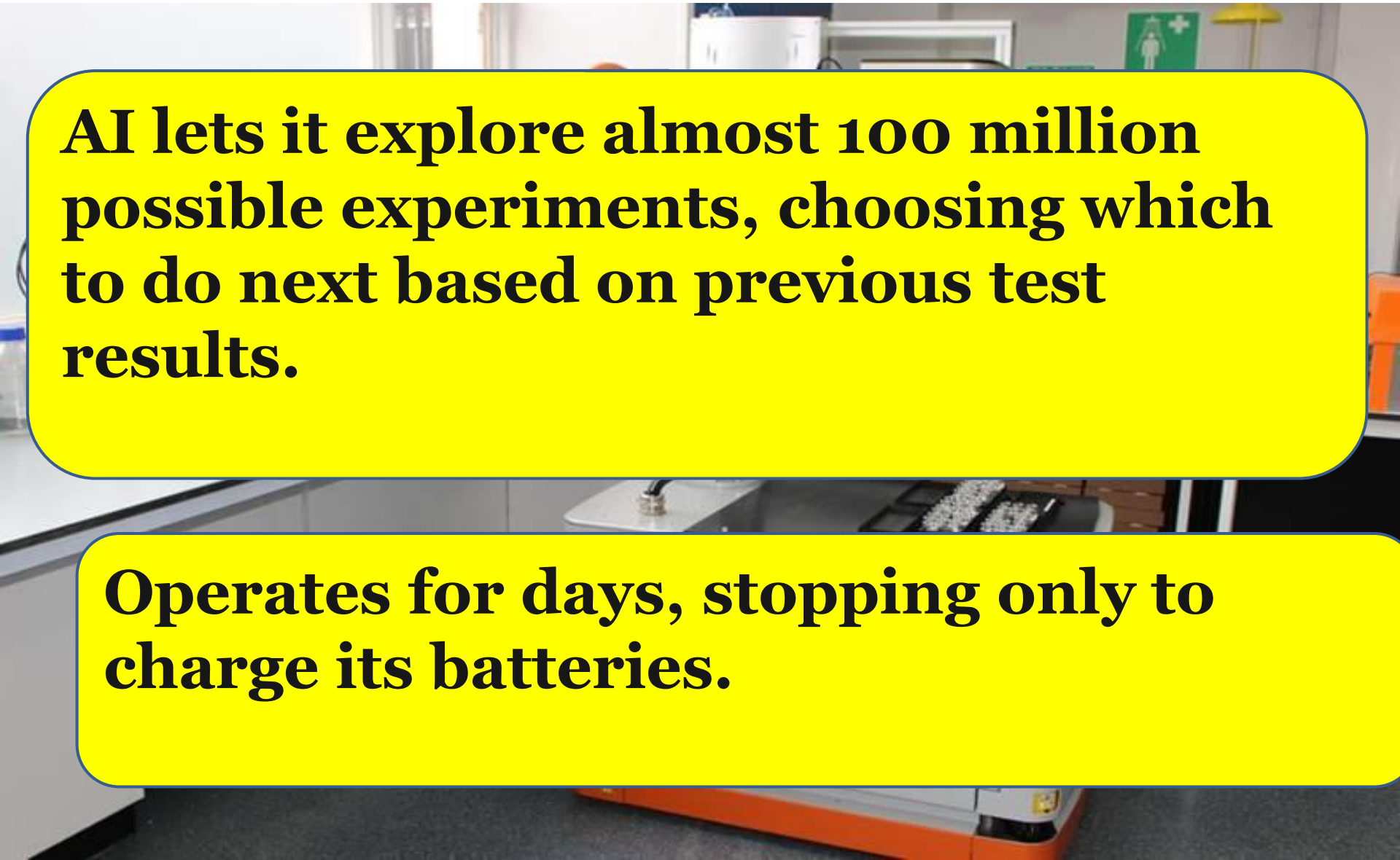
# Effects on research productivity ?

# Robot scientist
# Lowers various types cost

# Robot chemist at the University of Liverpool


**AI lets it explore almost 100 million possible experiments, choosing which to do next based on previous test results.**

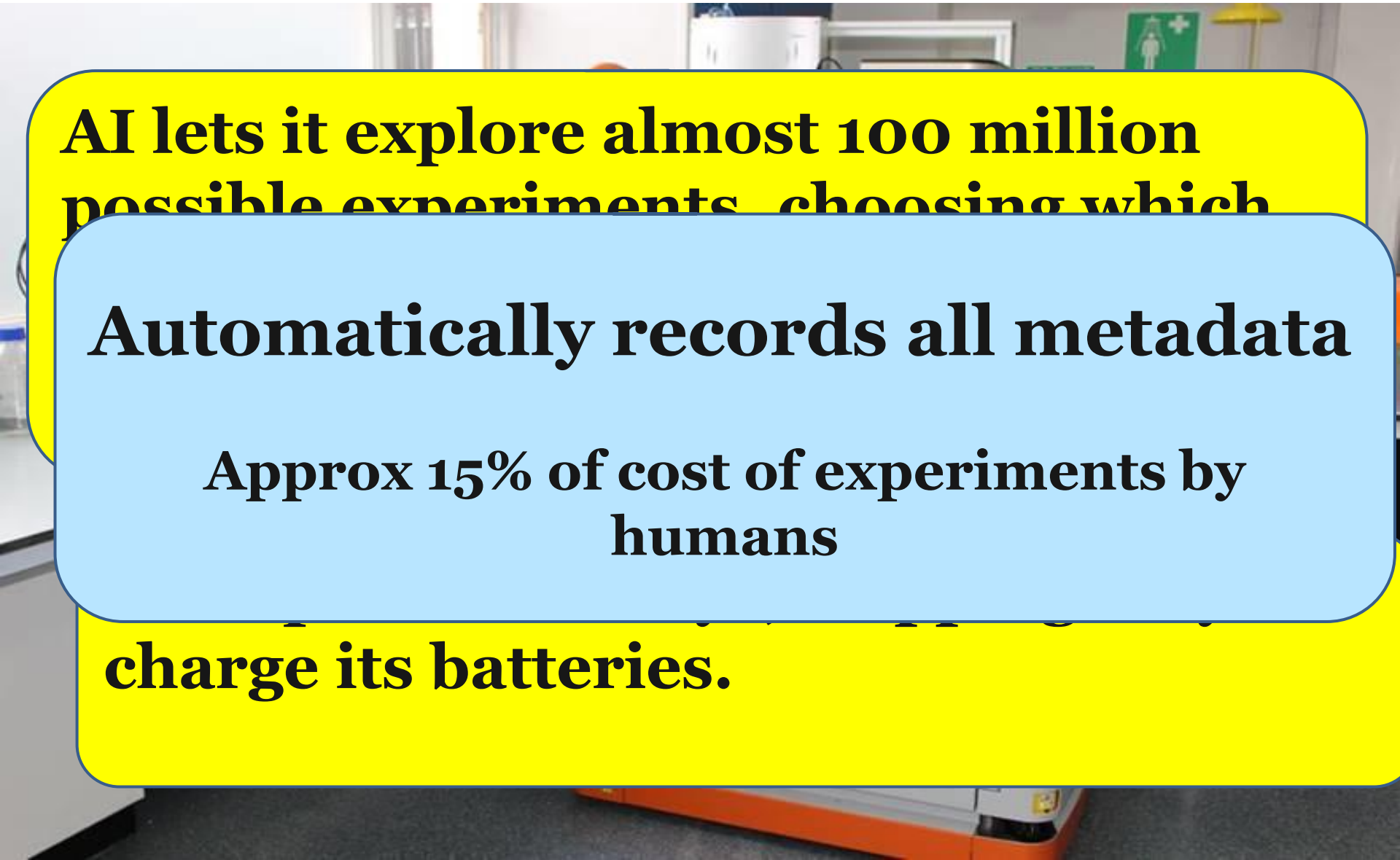
# Robot chemist at the University of Liverpool



**AI lets it explore almost 100 million possible experiments, choosing which to do next based on previous test results.**

**Operates for days, stopping only to charge its batteries.**

# Robot chemist at the University of Liverpool



**AI lets it explore almost 100 million possible experiments, choosing which**

**Automatically records all metadata**

**Approx 15% of cost of experiments by humans**

**charge its batteries.**

# Intelligent data sampling saves compute $$$

# Intelligent research assistants : to save time and money

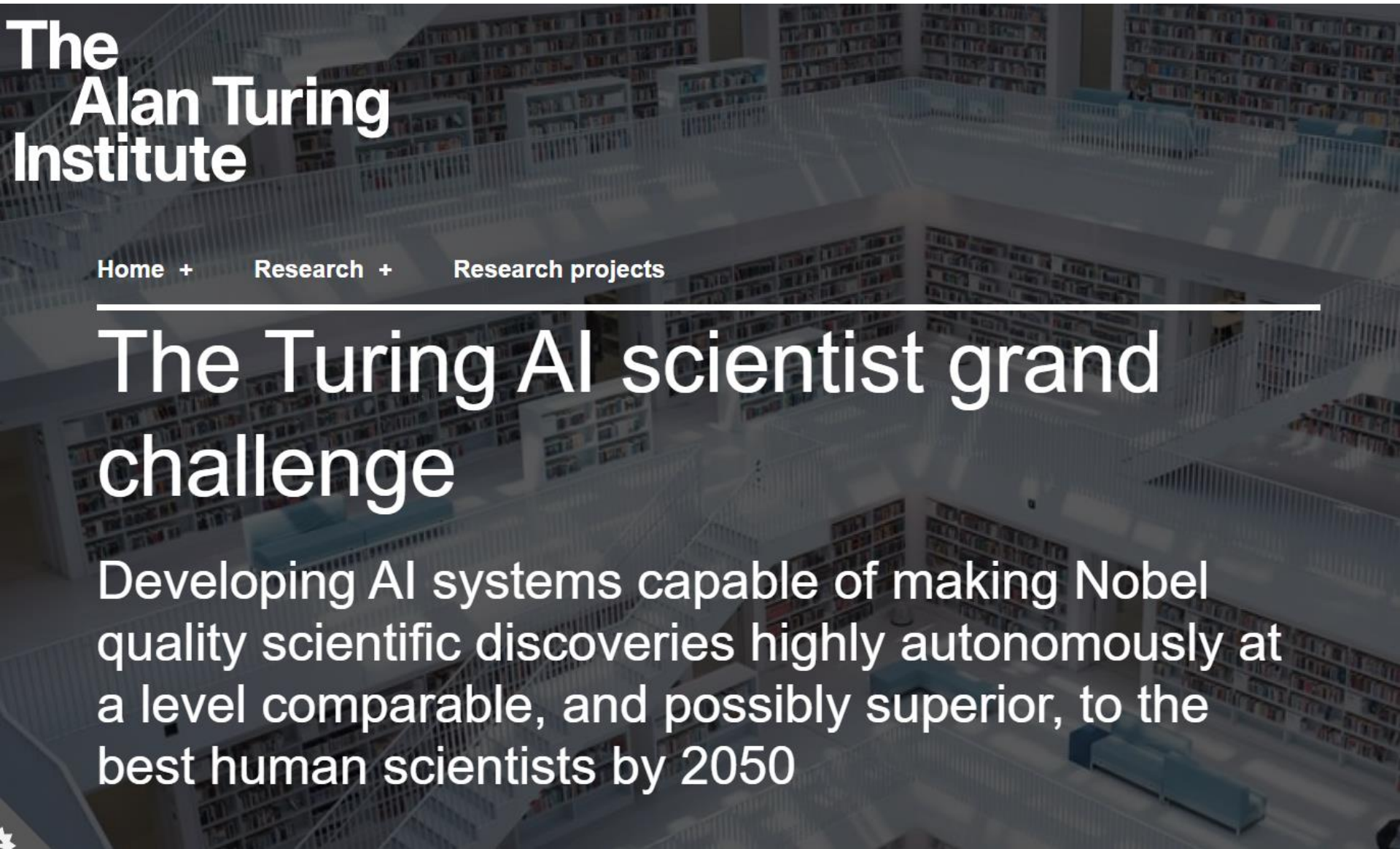
- **8 months to +/- weeks**
- **“Our results show that ChatGPT substantially raises average productivity: time taken decreases by 0.8 SDs and output quality rises by 0.4 SDs.”**
  **https://economics.mit.edu/sites/default/files/inline-files/Noy_Zhang_1.pdf**
- **USD 1.5 billion in 2020 in the US**
  **(Aczel, Szaszi and Holcombe, 2021)**

# Can public policy help ?

# Ambitious multi-disciplinary programmes

# Multi-disciplinarity

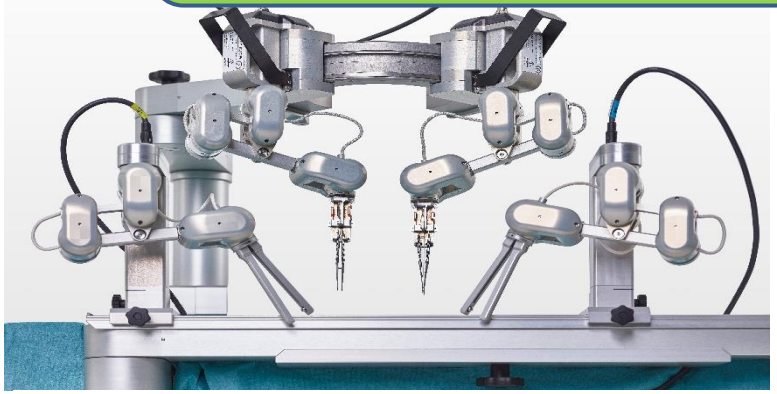
# Ambitious multidisciplinary programmes



# Less than 6% of all LBD publications can be mapped to at least one SDG

# Bring industry, roboticists and domain specialists together

**Strengthen data governance**



# Computational resources

- **National labs, industry and academia could work together to nurture AI ecosystems for tertiary education**

# Computational resources

- Na
  an
  nu
  ter

**Explore pooling resources internationally**

# Curricula

- ***Standard bio-science education doesn't address how to search for new hypotheses.***

- ***New PhD programmes based on knowledge synthesis – aided by AI***
- ***Promote research software engineers and engineering***
- ***Raise awareness of stage of development robot systems***

# Public R&D can advance the field

# Invest in developing new tools for AI in science

AlphaFold

≠

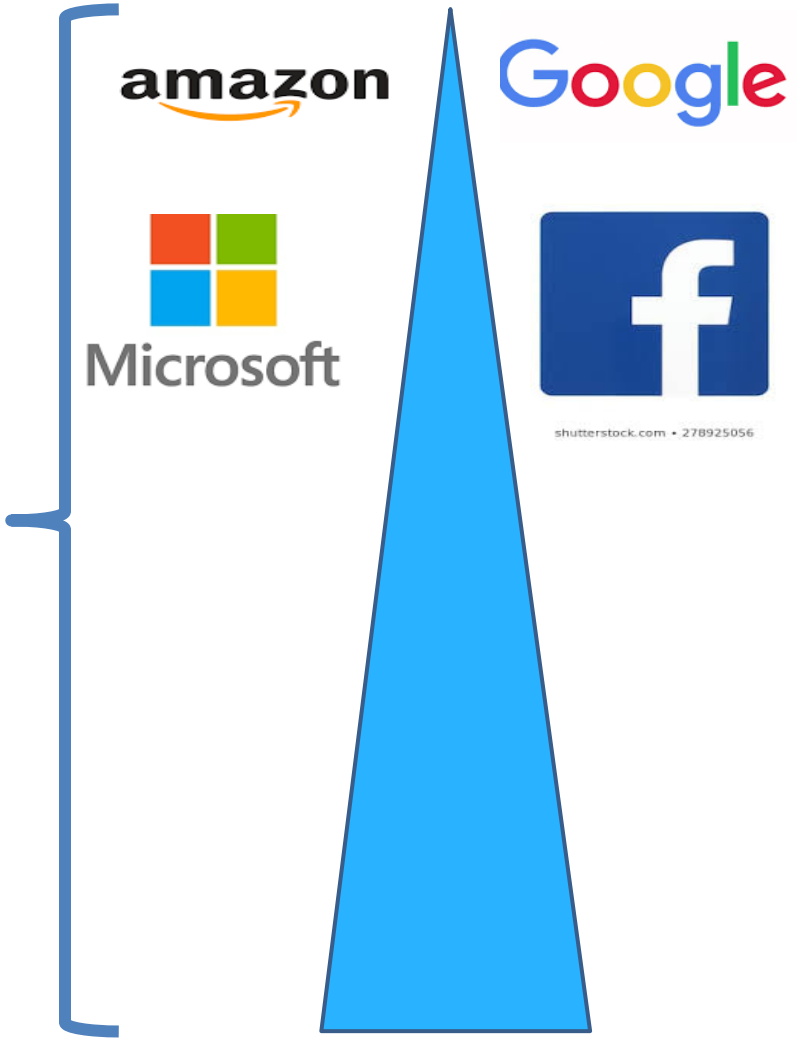
# A narrowing of AI research

AI R&D $$$

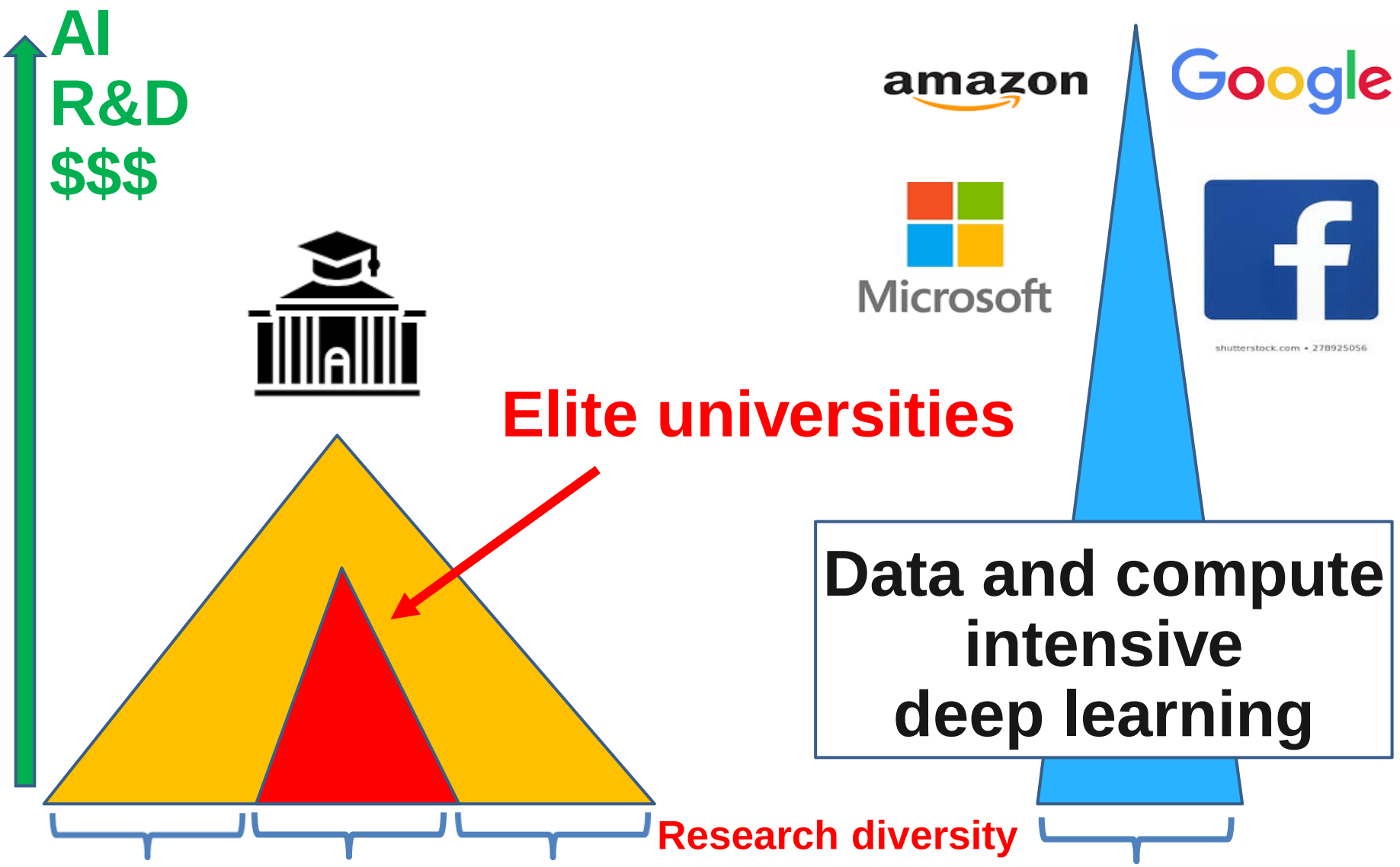
Multiples of public $

Growing faster

21% of AI PhDs went to industry in 2004. 70% in 2020.



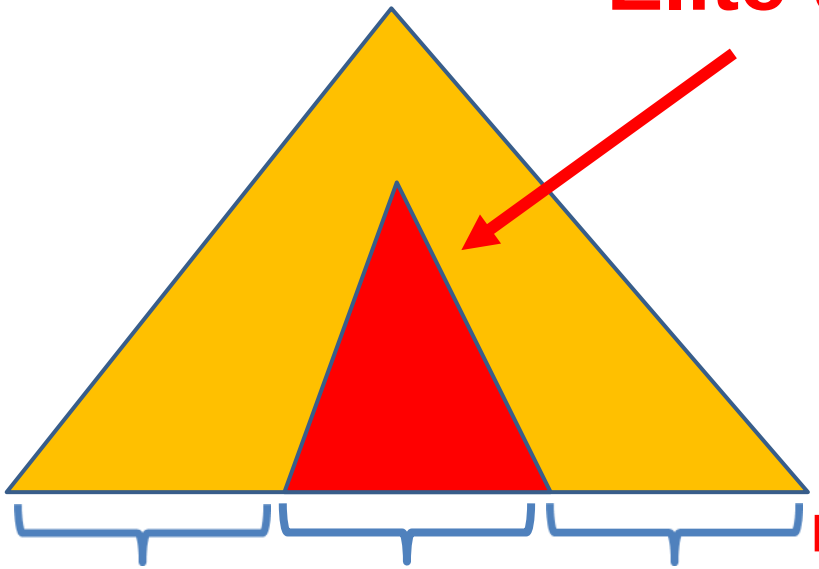
# A narrowing of AI research



# A narrowing of AI research

AI
R&D
$$$

Elite u



Research diversity

amazon

Google

Industry models are on average 29 times larger than academic models (2021)

# Foster more blue sky thinking

**More funding streams and/or publication processes to reward novel methods**

# Funders could help develop specialised tools to enhance collaborative human AI teams

+ AI

# Research governance

nature



NEWS FEATURE | 06 February 2023 | Correction 08 February 2023

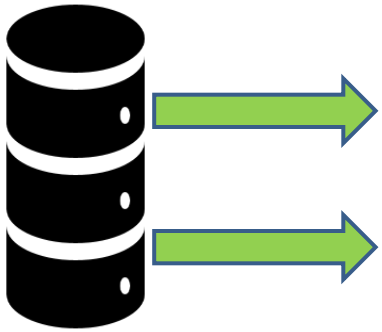
# What ChatGPT and generative AI mean for science

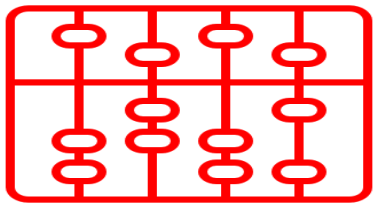
**Researchers are excited but apprehensive about the latest advances in artificial intelligence.**

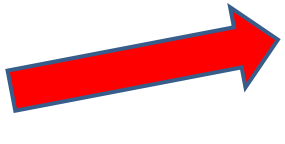
# Dangers of dual use AI in drug design

Publicly available data

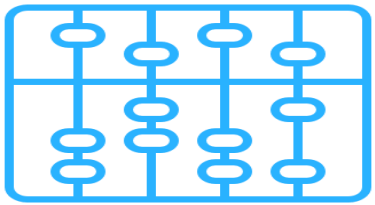
AI molecule design

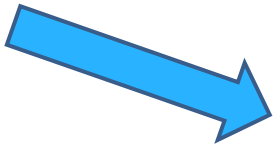






Maximise toxicity





Minimise toxicity

# Dangers of dual use AI in drug design

P
avai

## What to do?

- *High-level recognition of this danger is needed*
- *All parts of the science system have a role in responding*
- *Could draw on existing frameworks for responsible science – but technology-specific measures are needed too*

# Two parting thoughts


***A fast-moving field – much will be new in a year from now.***

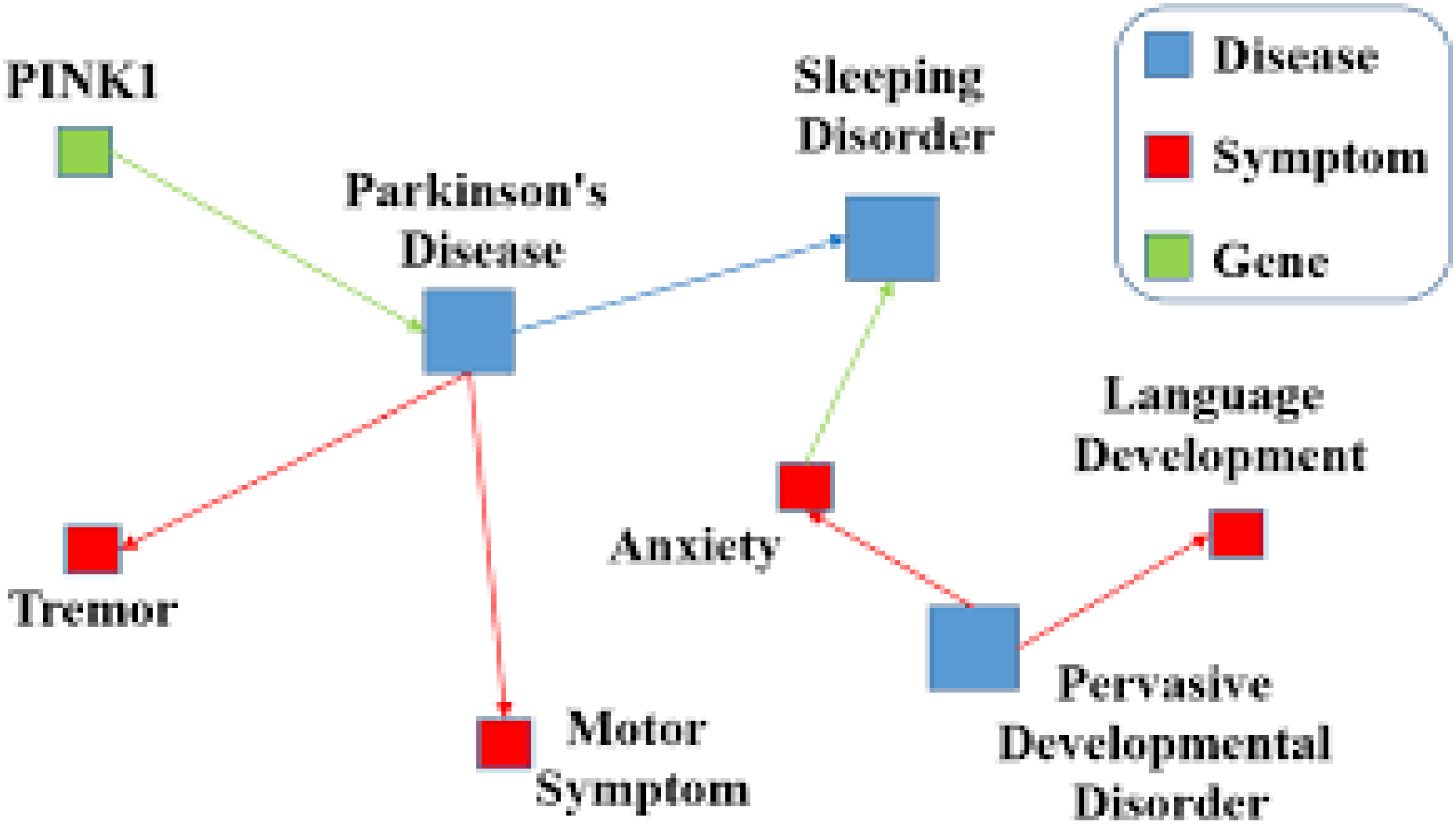
***AI in science may be the most important of all uses of AI.***

# Thank you

alistair.nolan@oecd.org

OECD
BETTER POLICIES FOR BETTER LIVES

# Support an extensive programme to build knowledge essential to AI in science



# What do ChatGPT and LLMs mean for science?


- Change training ?
- Research integrity processes to change?
- Most LLMs are corporate proprietary products.
- Erroneous truth claims.
- Equity?
- Legal implications of LLMs for science?

# AutoML to address skills needs?

Help bu

Will r
more

**Research challenges could be organised around AutoML for science.**

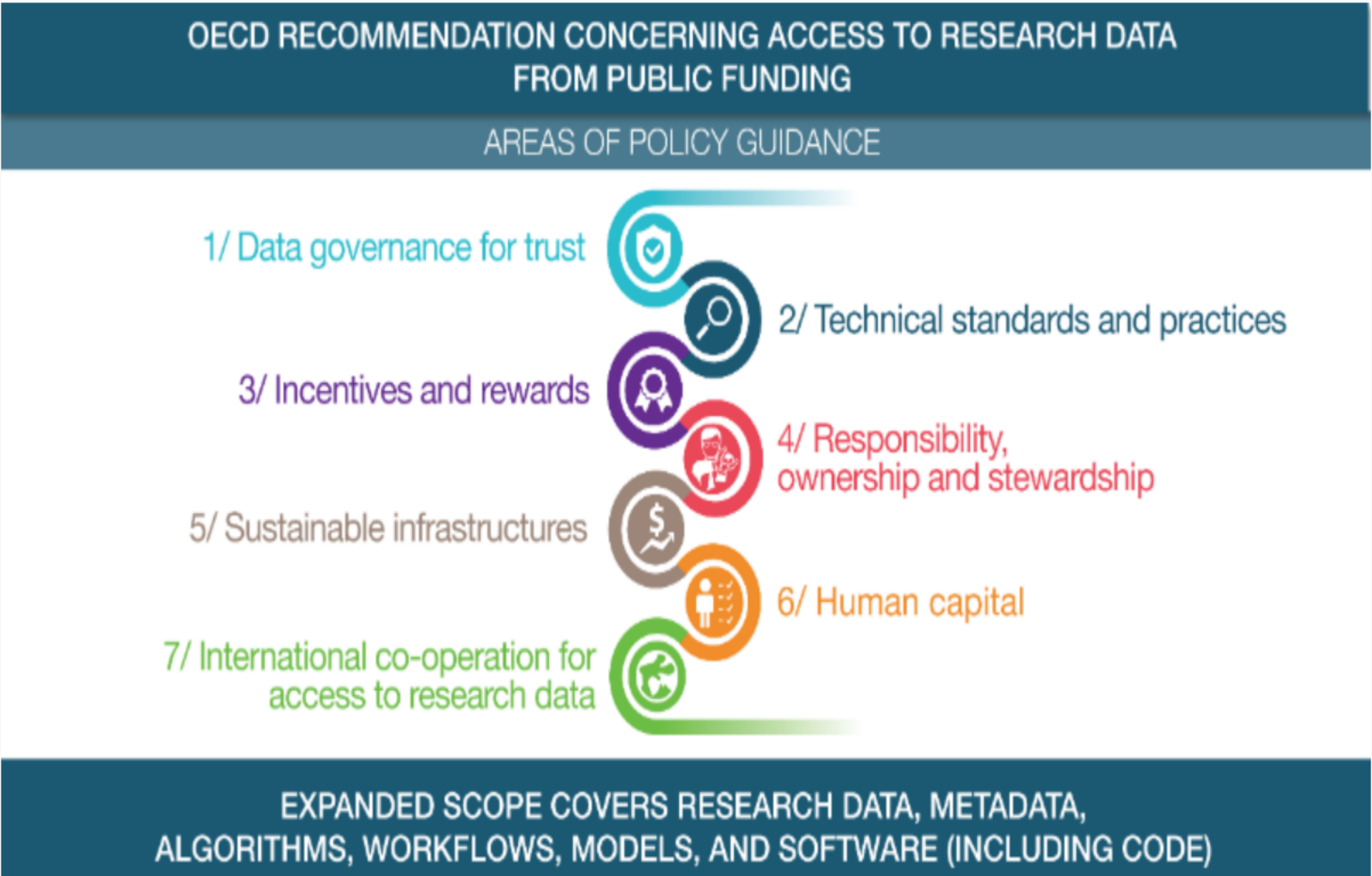
**Research could be funded that involves applying AutoML in AI-driven science.**

# Knowledge without understanding ?

## What if an AI finds something like this:

$$(x-2)^2(y-2x+2)^2(y+2x-10)^2(x-4)^2(y-2x+8)^2(y+2x-16)^2\left(y-3-3\left\lfloor x-\tfrac{11}{2}\right\rfloor^2\right)^2(x-8)^2$$
$$\cdot\left(y-2-3\left\lfloor\tfrac{x-8}{2}\right\rfloor^2\right)^2(x-11)^2\left(y-\tfrac{1}{2}x+\tfrac{5}{2}-3\left\lfloor\tfrac{x-11}{2}\right\rfloor^2\right)^2\left(y+\tfrac{1}{2}x-\tfrac{17}{2}-3\left\lfloor\tfrac{x-11}{2}\right\rfloor^2\right)^2(x-15)^2$$
$$\cdot\left(y-4-3\left\lfloor\tfrac{x-14}{2}\right\rfloor^2\right)^2(y-2x+52)^2(x-17)^2(y+x-21)^2(x-19)^2\left(y-x+17-3\lfloor x-20\rfloor^2\right)^2$$
$$\cdot\left(y+x-23-3\lfloor x-20\rfloor^2\right)^2\left(y-x+19-3\lfloor x-21\rfloor^2\right)^2\left(y-3-3\lfloor x-21\rfloor^2\right)^2(x-25)^2\left(y+\tfrac{1}{4}x-\tfrac{41}{4}-3\left\lfloor\tfrac{x-25}{2}\right\rfloor^2\right)^2$$
$$\cdot\left(y-\tfrac{1}{8}x-\tfrac{1}{8}-3\left\lfloor\tfrac{x-25}{2}\right\rfloor^2\right)^2\left(y+\tfrac{5}{8}x-\tfrac{151}{8}-3\left\lfloor\tfrac{x-25}{2}\right\rfloor^2\right)^2(y-2x+54)^2(y+2x-62)^2\left(y-3-3\left\lfloor x-\tfrac{57}{2}\right\rfloor^2\right)^2$$
$$\cdot(x-31)^2(y+x-35)^2(x-33)^2(x-34)^2\left(y+\tfrac{1}{2}x-21-3\left\lfloor\tfrac{x-34}{2}\right\rfloor^2\right)^2\left(y-\tfrac{1}{2}x+15-3\left\lfloor\tfrac{x-34}{2}\right\rfloor^2\right)^2$$
$$\cdot\left((x-38)^2+(y-3)^2-1\right)^2(x-40)^2(y+2x-84)^2(y-2x+80)^2(x-42)^2(x-43)^2\left(y-2-3\left\lfloor\tfrac{x-43}{2}\right\rfloor^2\right)^2$$
$$\cdot(y-3-|x-47|)^2\left((x-47)^2+\left(y-3+\sqrt{y^2-6y+9}\right)^2\right)^2+\left(y^2-6y+8+\sqrt{y^4-12y^3+52y^2-96y+64}\right)^2=0$$

# Data


OECD RECOMMENDATION CONCERNING ACCESS TO RESEARCH DATA FROM PUBLIC FUNDING
AREAS OF POLICY GUIDANCE
1/ Data governance for trust
2/ Technical standards and practices
3/ Incentives and rewards
4/ Responsibility, ownership and stewardship
5/ Sustainable infrastructures
6/ Human capital
7/ International co-operation for access to research data
EXPANDED SCOPE COVERS RESEARCH DATA, METADATA, ALGORITHMS, WORKFLOWS, MODELS, AND SOFTWARE (INCLUDING CODE)

# Strategic breakthroughs semiconductor manufacture