



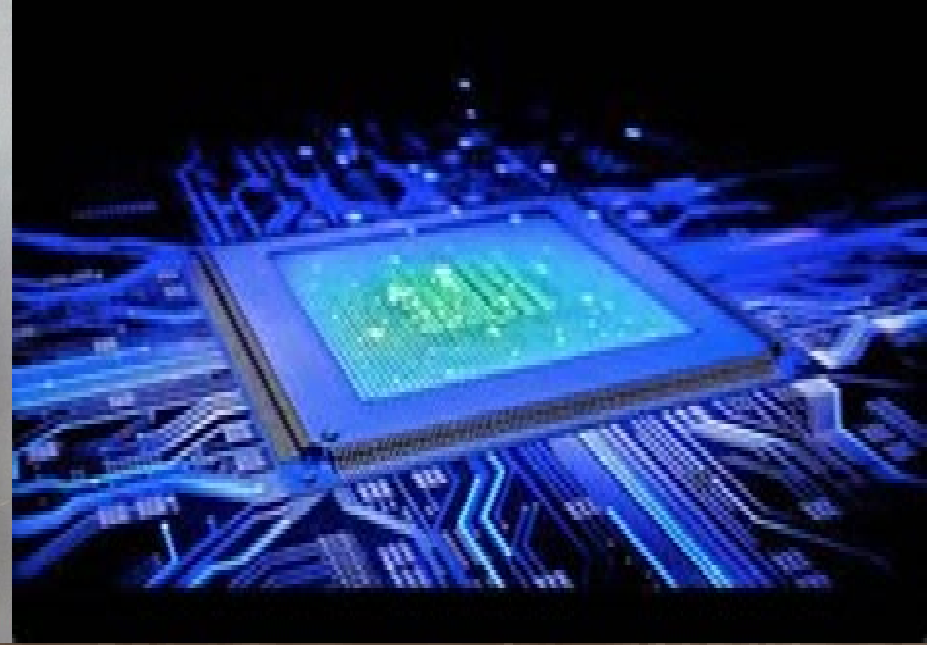
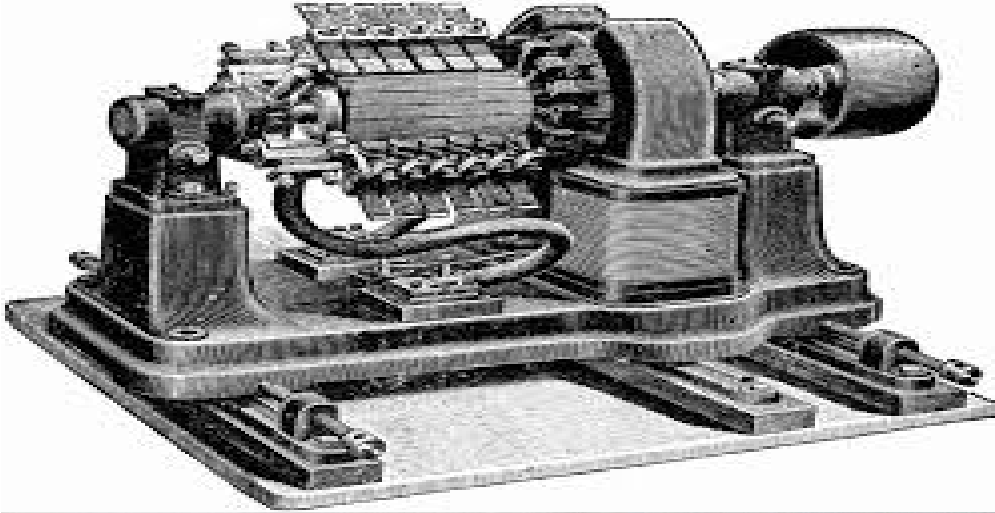
AI in Science, and Science in AI

Alistair Nolan
**Committee for Scientific and
Technological Policy (CSTP)**

GPN, October 11, 2019



Progress in science means growth and better living standards

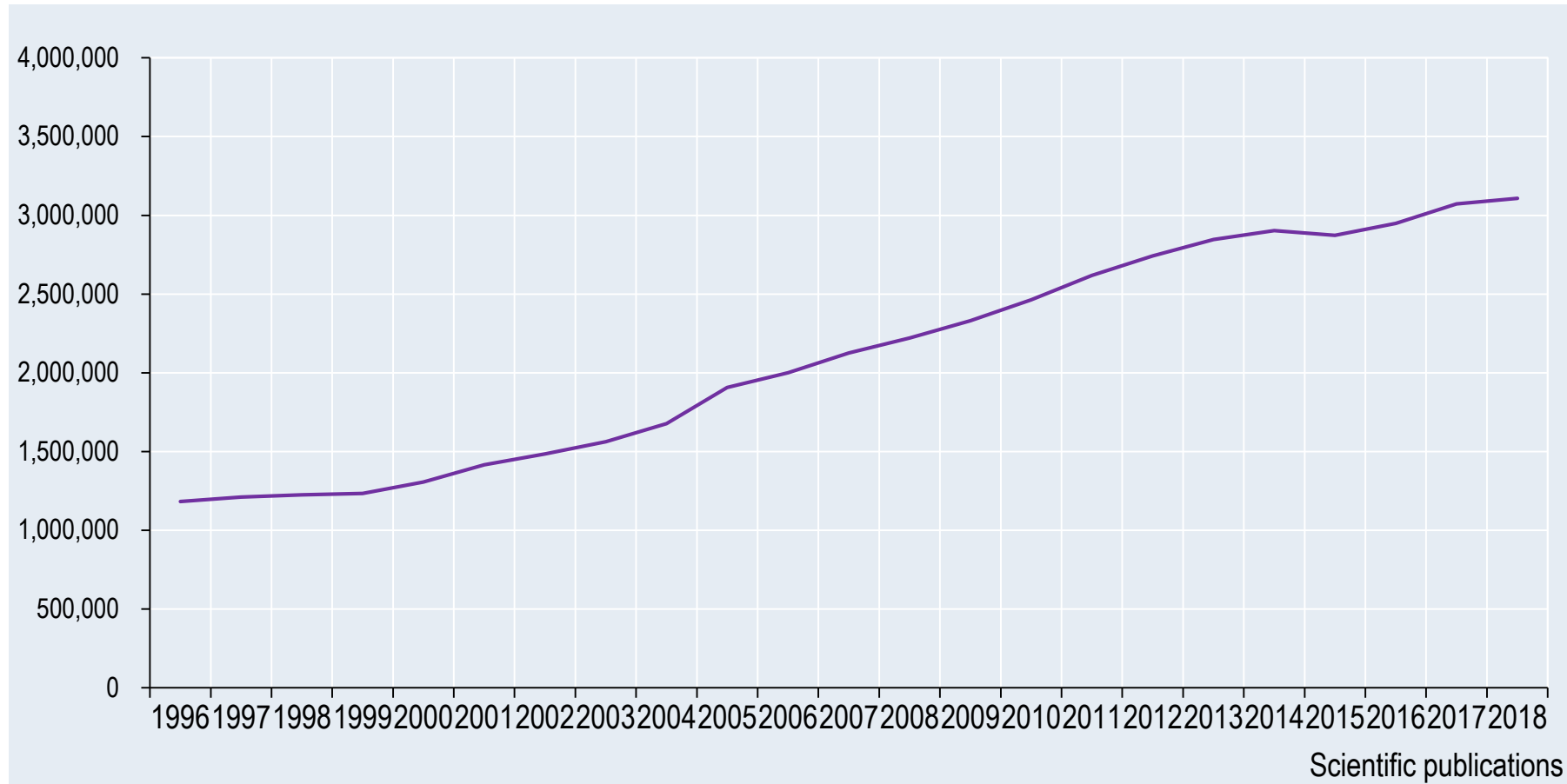




Scientists reaching 'peak reading'

(annual number of scientific publications, 1996-2018)

Scientific publications are defined as the total number of documents published in scholarly journals indexed in Scopus (all document types are included).



Source: OECD calculations based on Scopus Custom Data, Elsevier, Version 5.2019.



Discovery getting harder ?

$$F = m \times a$$

1686

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

1884

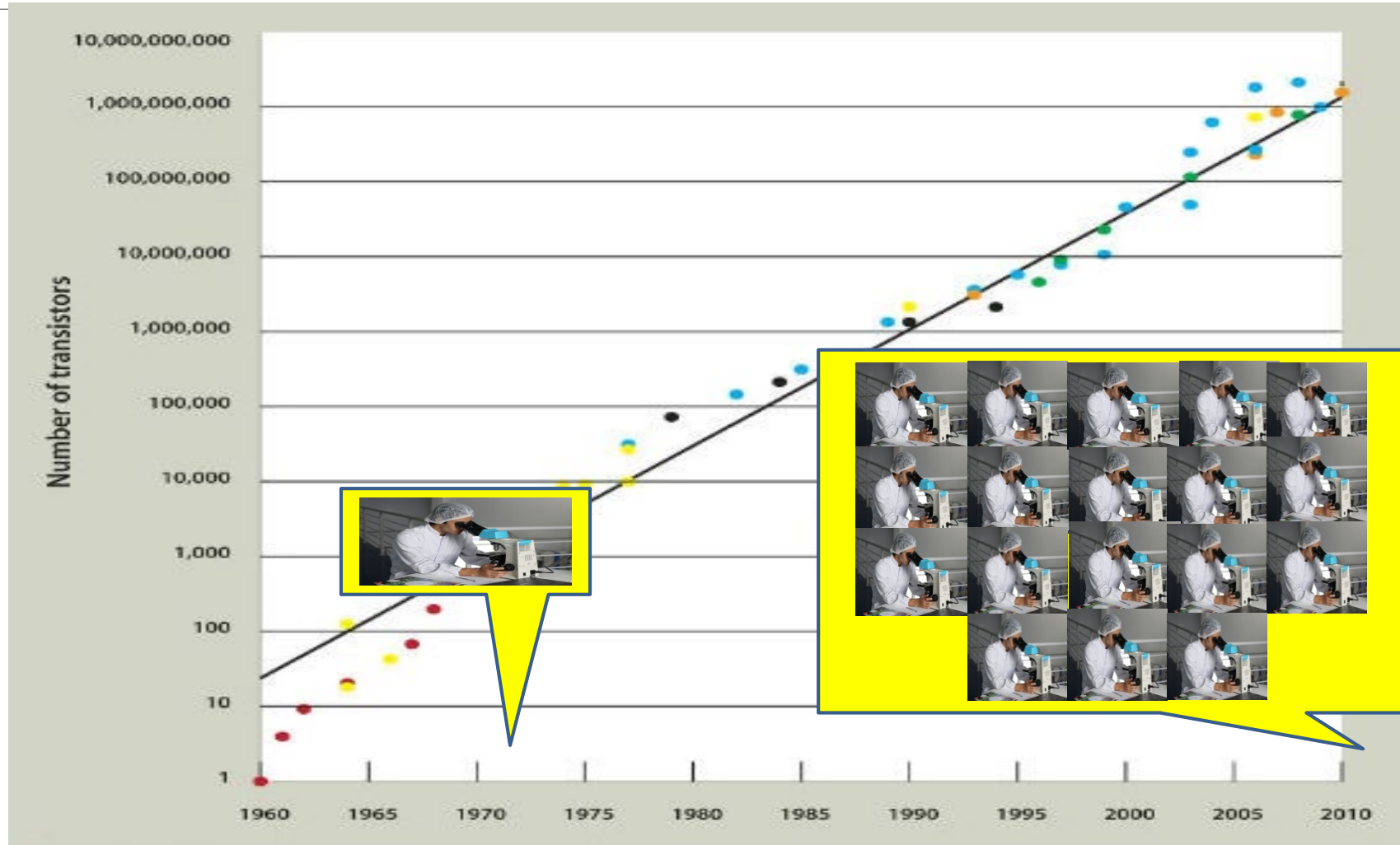
$$(1 - e^{-2\Delta}) r^{D-3} = \frac{2K}{D-2} \int_0^r \rho(r') r'^{D-2} dr' = \frac{2 \cdot 8\pi(D-3)G}{(D-2)} \frac{M}{\Omega_{D-2}} \Rightarrow$$

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

1973



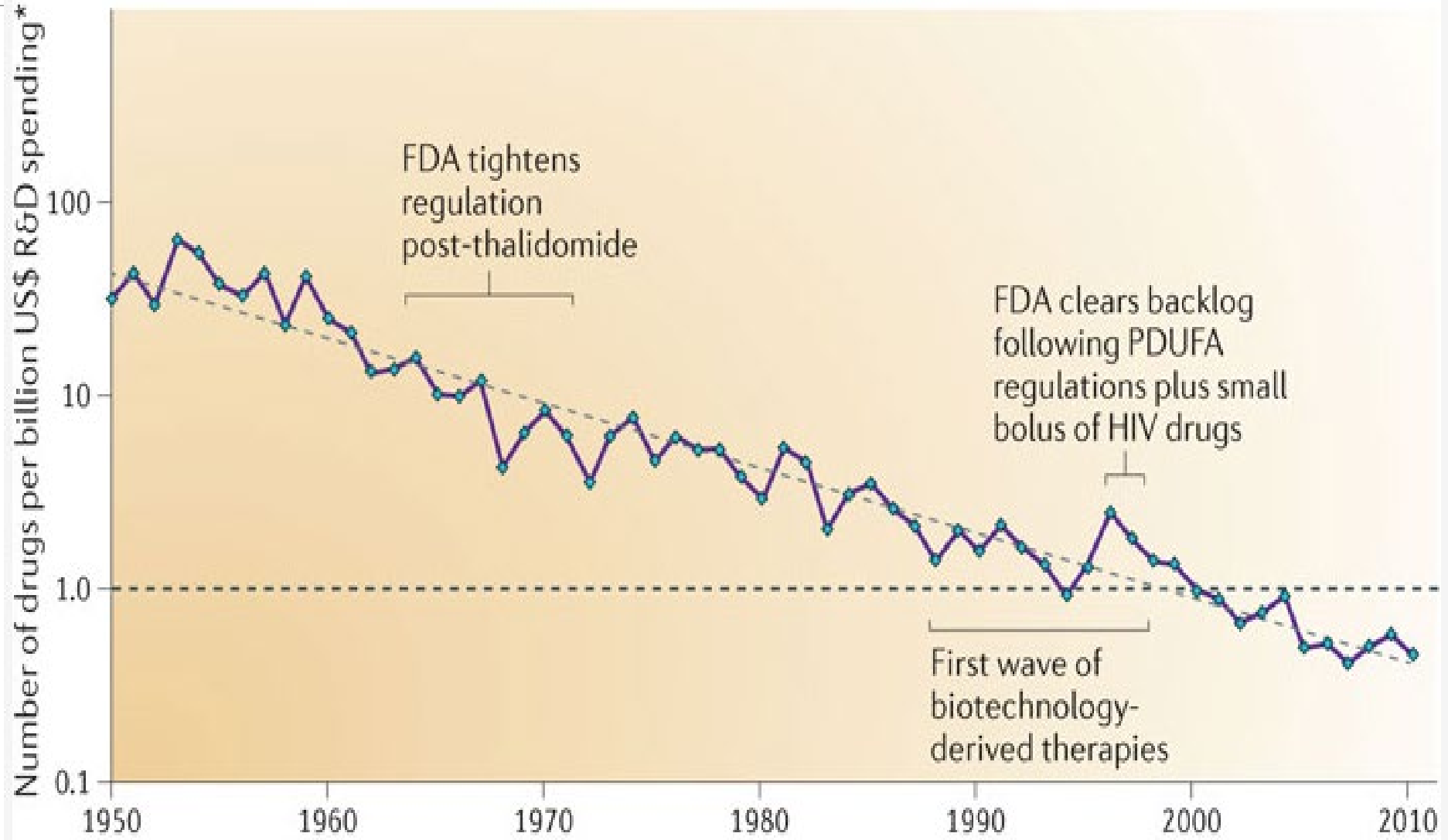
Discovery is getting harder ?



Source: N. Bloom, C.I. Jones, J. Van Reenen and M. Webb (2017), "Are Ideas Getting Harder to Find?", CEP Discussion Paper No 1496, September 2017.



'Eroom's Law'





Science is now 'in the data'

Large Hadron Collider - typical data rate:
300,000,000,000,000,000 bytes *per minute*
(would fill the hard discs of 1,000,000 PCs
in 3.5 minutes)



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



Industrial Research

AI helps discover new metals for jets:

The Periodic Table

1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe

“What would have taken years, it narrowed down to days”





Science may come to know things in new ways



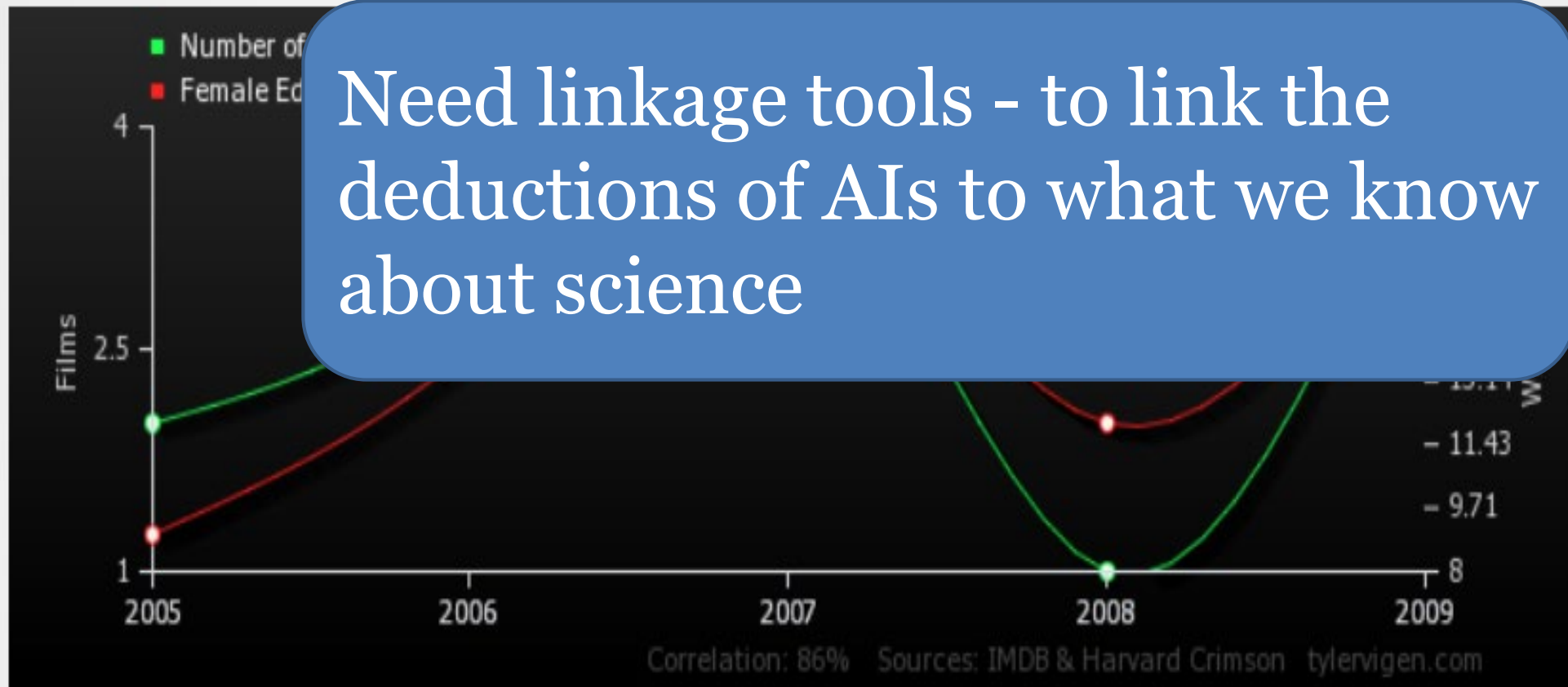
Stephen Roberts,
FREng
RAEng/Man Professor
of Machine Learning,
Oxford University

“We will very soon – perhaps within 3 years – be at a stage when...the human would ask the machine to write an algorithm best suited to treating a given data set. The machines can already write code that is ‘beautiful’ and which human coders would find counter-intuitive.”



But AI is not a panacea

Number of films Nicolas Cage appeared in
correlates with
Female Editors on Harvard Law Review





Again, AI is not a panacea

What if an AI finds this:

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



We also need good policies

- Skills : eg, Curricula ? Enough students into STEM ?
- Data policies: eg. Access, say for medical research, and opening research data, for reproducibility.
- High performance computing: Terms of access?
- Standards: eg. for machine readability of scientific datasets.
- Institutional norms: eg. Should machines be included in academic citations?
- Impacts for developing countries ?



3. A project on AI in science

- **Will AI's impacts be greatest in experimental science, operating in compartmented sequential steps, or will its scope be wider?**
- **Would other institutional and system conditions – such as the drug approval process - effectively nullify some of the productivity gains enabled by AI?**
- **Would the costs of deliberately seeking to deepen the use of AI in science outweigh the benefits, and if so, how would this best be achieved?**



What priorities for public R&D for AI? (project so far unfunded)





What priorities for public R&D for AI ? (project so far unfunded)

Corporate and public AI research goals might not fully align.

Malicious uses of AI: understanding the threat landscape – and developing mitigation measures – is of clear public interest.

Investments in standards, including systems that might test standards.

A new science of machine ecology ?



Thank you
alistair.nolan@oecd.org