Peter Novitzky
Senior Research Fellow (UCL / PETRAS)
Assoc Prof — Ethics of Emergent Technologies (Avans)

22 May 2024

Empowerment and Ethics at the Edge

The Benefits and Risks of Edge Technologies

The Foundation for Science and Technology

The Royal Society, London

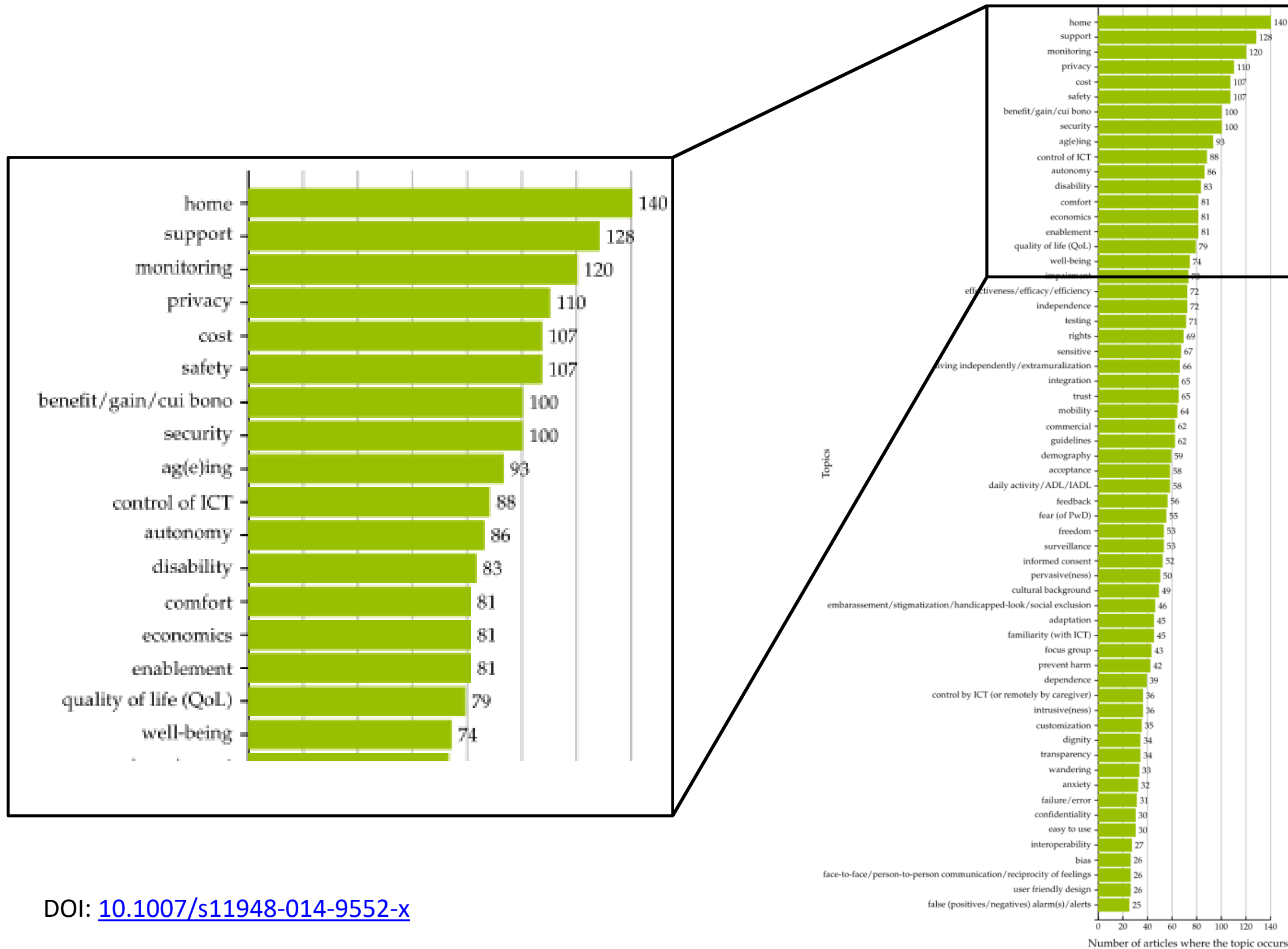# ETHICAL ASPECTS OF EMERGENT TECHNOLOGIES (AAL)

Fig. 4 Frequency of ethical terms

Fig. 4 continued

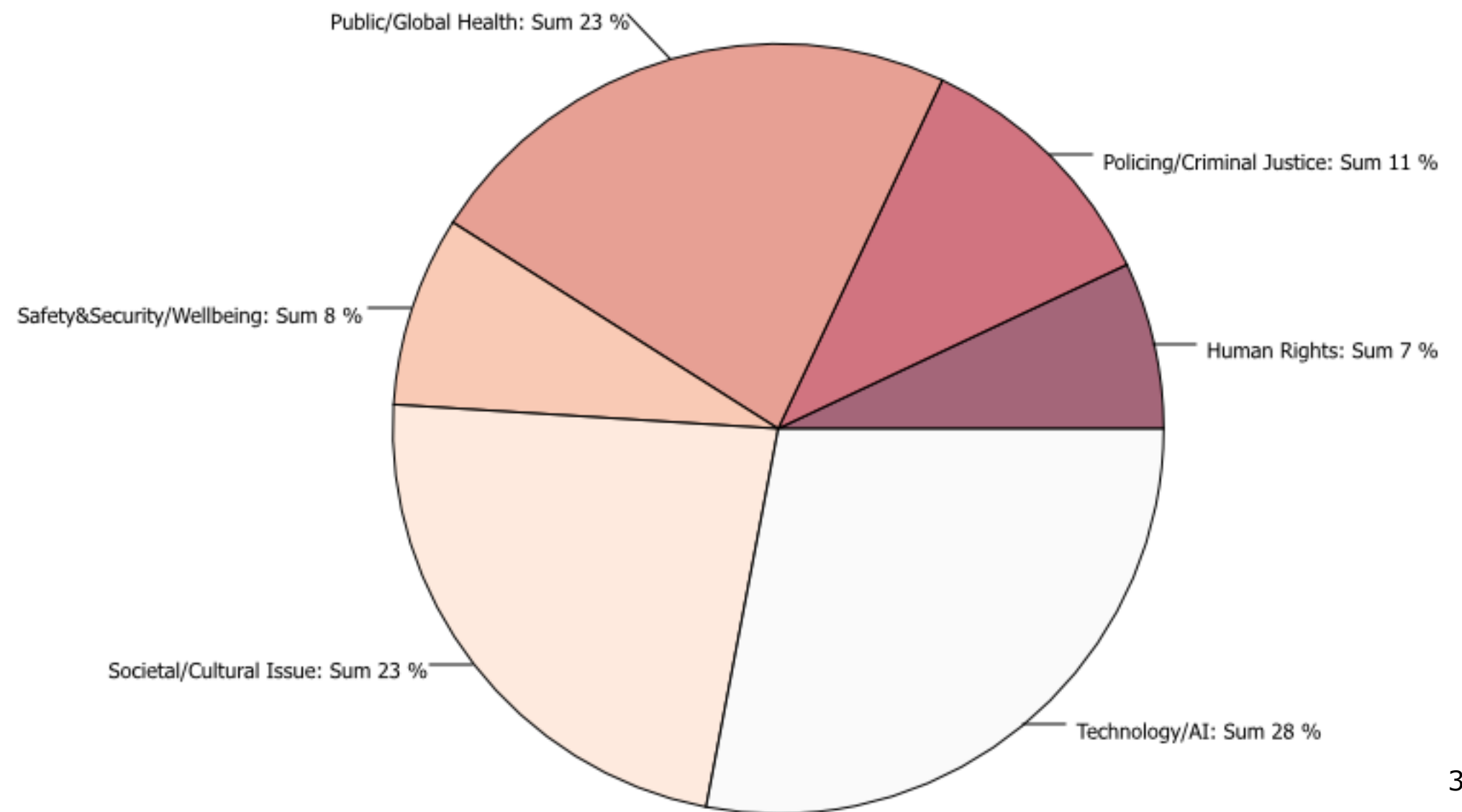# ETHICAL ASPECTS OF EMERGENT TECHNOLOGIES (IPV)

- Modern technologies have a tremendous amount of **ethically relevant challenges**

- The numbers might even increase

- Copying the **complex entanglements** of our societies, e.g.:

  - Private/public;

  - Communal/individual

- Copying the **complex links** between stakeholders, actors, interests

  - PwD/(in)formal carer/HCP

  - Victim/perpetrator

- **Wicked problems**

  - By definition: impossible to solve

  - Even with technology

  - But we still try…

The 'art' of the ethicist-researcher is to **pick those most relevant** to ensure respect, rights, universal and societal values

Table 4

Categorization of relevant publications based on source of domestic violence/IPV. The categories may overlap across a single publication.

Category	Publications
Personal Safety & security, wellbeing	[14–20]
Technology, AI	[14–17,19–35]
Policing, Criminal Justice, Law enforcement	[14,23,26,36–41]
Societal, Cultural issue	[14,17,22,26,28,35,38–48]
Global or Public health issue	[2,14,15,17,22,25,27,29–32,42,45,49–52]
Human rights issue	[2,33,34,41,46]

3

DIGITAL, PHYSICAL, AND POLITICAL CONNOTATIONS OF AI-CYBERSECURITY
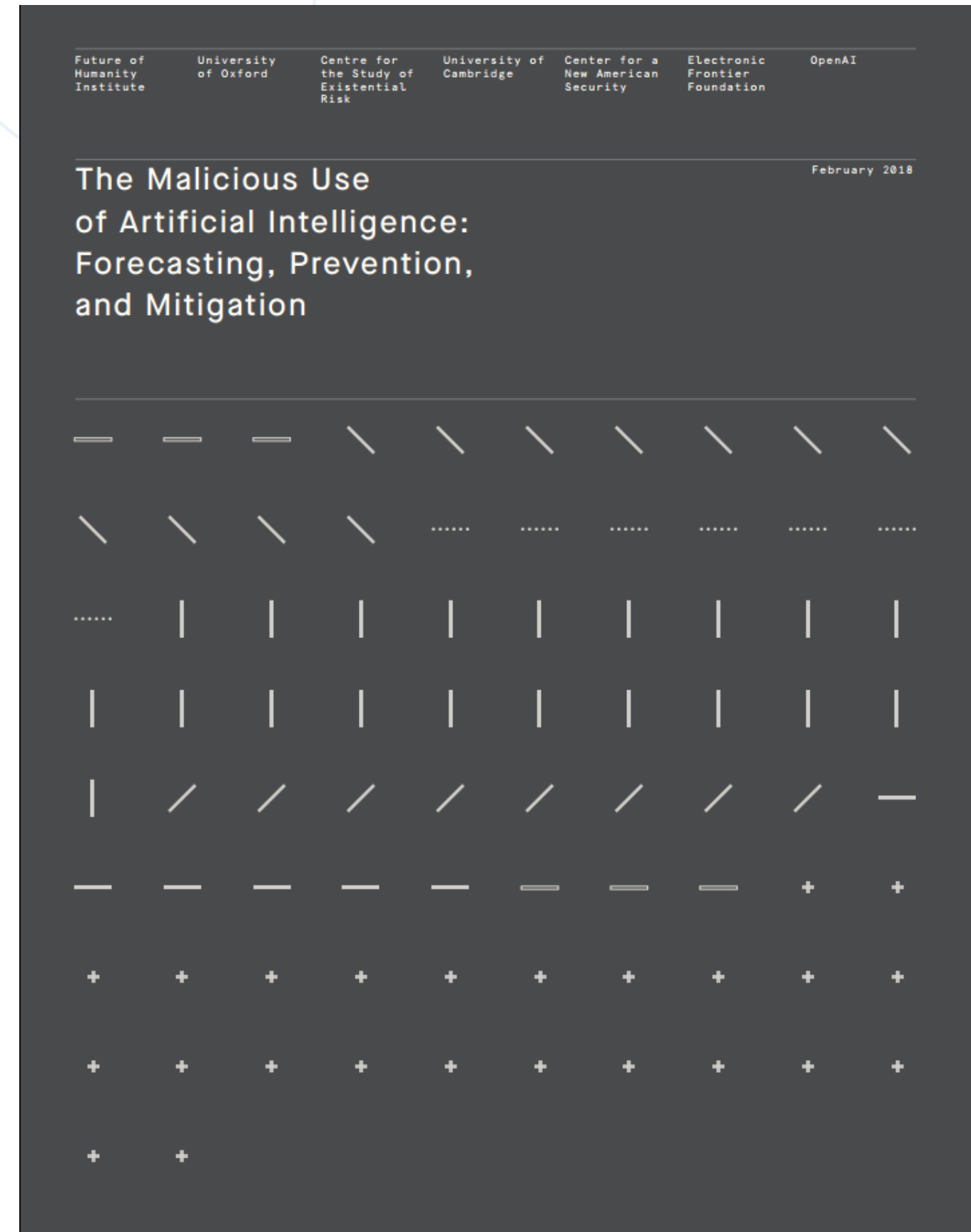
DIGITAL SECURITY
- Automation of social engineering attacks
- Automation of vulnerability discovery
- More sophisticated automation of hacking
- Human-like DoS
- Automation of service tasks in criminal cyber-offense
- Prioritising targets for cyber attacks using ML
- Exploiting AI used in applications, esp. information security
- Black-box model extraction of proprietary AI system capabilities

PHYSICAL SECURITY
- Terrorist repurposing of commercial AI systems
- Endowing low-skill individuals with previously high-skill attack capabilities
- Increased scale of attacks
- Swarming attacks
- Attacks further removed in time and space (autonomy)

POLITICAL SECURITY
- State use of automated surveillance platforms to suppress dissent
- Fake news reports with realistic fabricated video and audio
- Automated, hyper-personalised disinformation campaigns
- Automating influence campaigns
- Denial-of-information attacks (information overload)
- Manipulation of information availability

Future of Humanity Institute | University of Oxford | Centre for the Study of Existential Risk | University of Cambridge | Center for a New American Security | Electronic Frontier Foundation | OpenAI

February 2018

The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation

4

DIGITAL, PHYSICAL, AND POLITICAL CONNOTATIONS OF AI-CYBERSECURITY
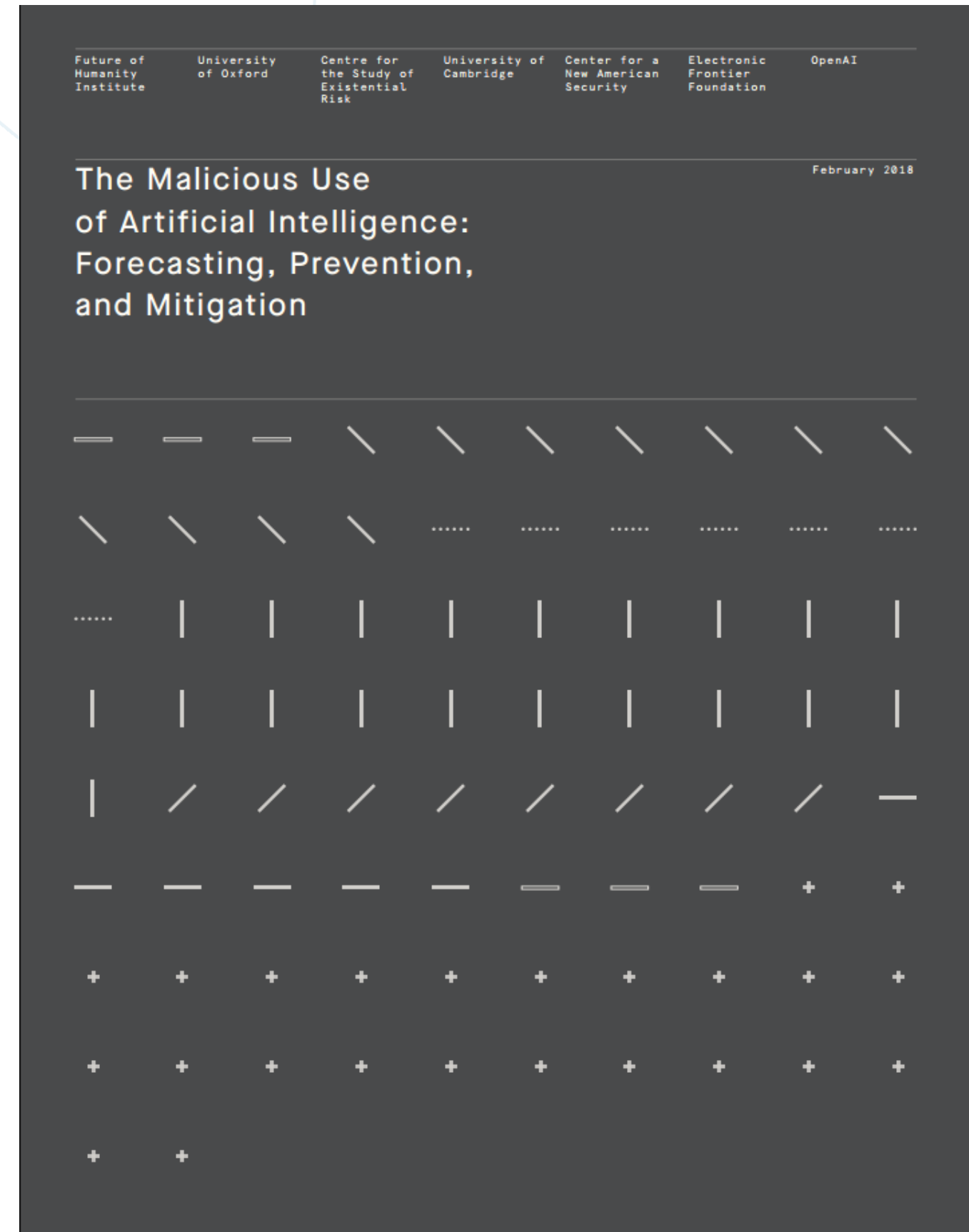
DIGITAL SECURITY

- Automation of social engineering attacks [2024a; 2024b]
- Automation of vulnerability discovery [2024c]
- More sophisticated automation of hacking [2024d]
- Human-like DoS [2023a]
- Automation of service tasks in criminal cyber-offense [2023b]
- Prioritising targets for cyber attacks using ML [2023c]
- Exploiting AI used in applications, esp. information security [2023d]
- Black-box model extraction of proprietary AI system capabilities [2021]
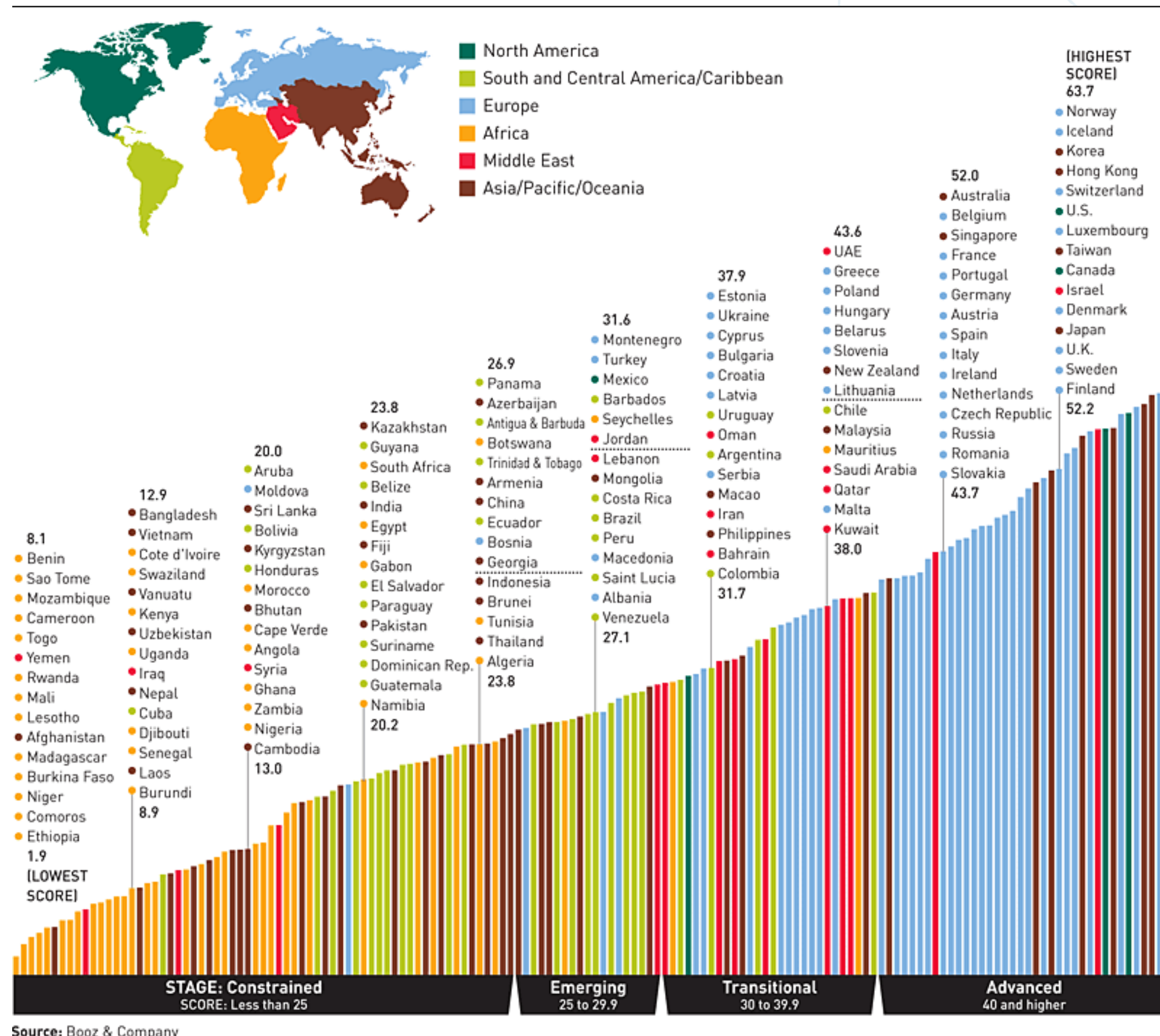
PHYSICAL SECURITY

- Terrorist repurposing of commercial AI systems
- Endowing low-skill individuals with previously high-skill attack capabilities [2024e]
- Increased scale of attacks [2024f]
- Swarming attacks [2024g]
- Attacks further removed in time and space (autonomy) [2024h]

POLITICAL SECURITY

- State use of automated surveillance platforms to suppress dissent [2022]
- Fake news reports with realistic fabricated video and audio [2024i]
- Automated, hyper-personalised disinformation campaigns [2024i]
- Automating influence campaigns [2018]
- Denial-of-information attacks (information overload) [2024j]
- Manipulation of information availability [2022]

Future of Humanity Institute · University of Oxford · Centre for the Study of Existential Risk · University of Cambridge · Center for a New American Security · Electronic Frontier Foundation · OpenAI

The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation

February 2018

5

DIGITAL, PHYSICAL, AND POLITICAL CONNOTATIONS OF AI-CYBERSECURITY
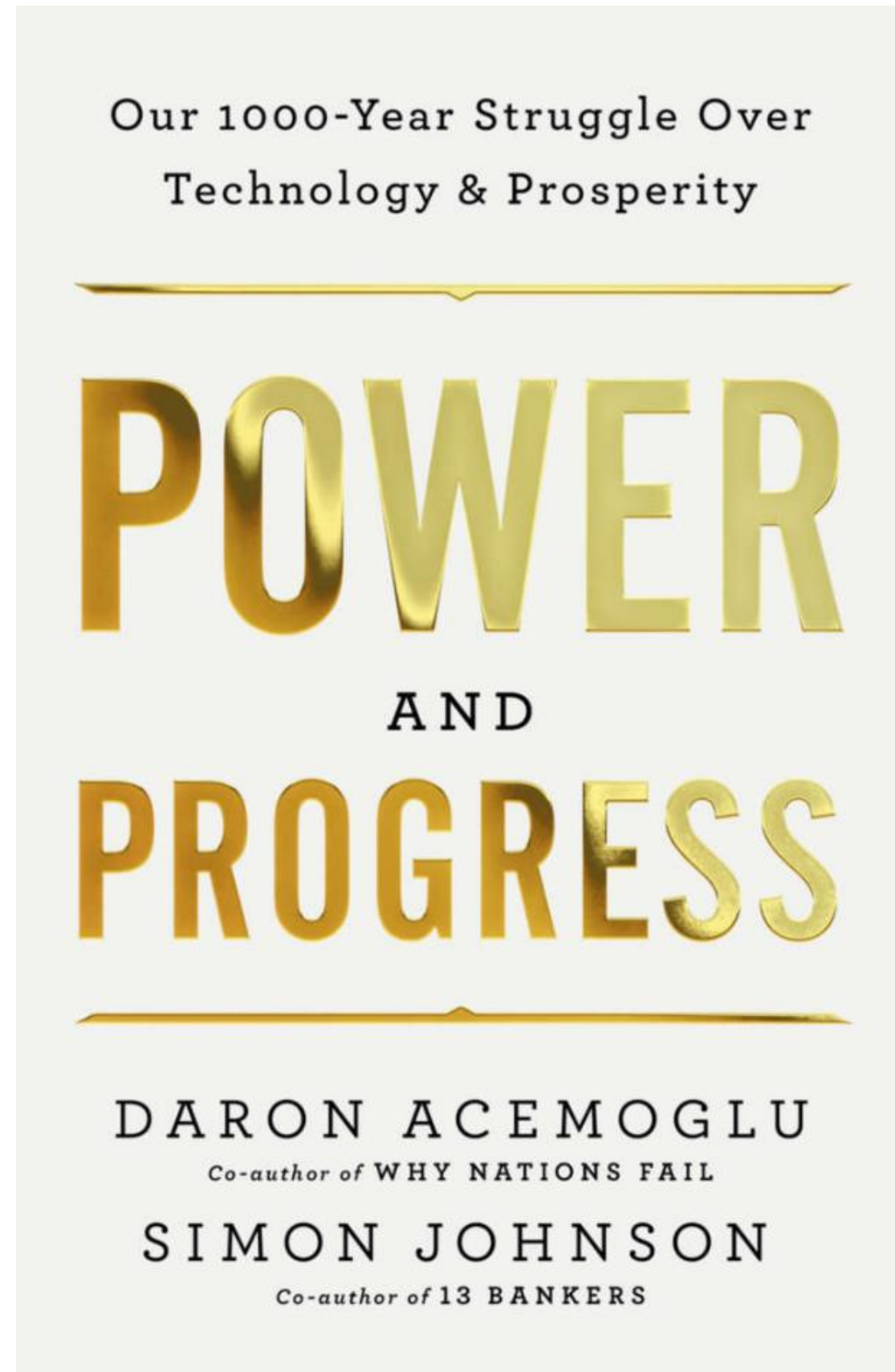

Source: Booz & Company

Booz & Company's Digitization Index is a composite score that calculates the level of a country's digitization using **23 indicators** to measure the following **six key attributes**:

- **Ubiquity**: consumer / enterprise universal access to digital services and applications.
- **Affordability**: digital services pricing making them available to as many people as possible.
- **Reliability**: The quality of available digital services.
- **Speed**: digital service accessibility in real time.
- **Usability**: The ease of use of digital services and the ability of local ecosystems to boost the adoption of these services.
- **Skill**: The ability of users to incorporate digital services into their lives and businesses.

The Digitization Index measures a country's level of digitization on a scale of 0 to 100, with 100 signifying the most advanced, to identify its distinct stage of digital development: **constrained**, **emerging**, **transitional**, or **advanced**.

~~ETHICAL~~ POLITICAL ASPECTS OF EMERGENT TECHNOLOGIES

Our 1000-Year Struggle Over
Technology & Prosperity

POWER
AND
PROGRESS

DARON ACEMOGLU
Co-author of WHY NATIONS FAIL

SIMON JOHNSON
Co-author of 13 BANKERS

(2023)

VISION
- Most of our technology- and AI-related innovations originate from a **small number of visionaries**
- **Small number = higher probability of going something wrong**
- A vision might be **skewed** (focus, interest, etc)

DEMOCRATIC DECISION-MAKING
- We need to reinstate democratic institutions and decision making
- Involvement of citizens, democratic public
- **Delegating** decision-making **to expert groups**/professionals is **insufficient**

POLITICAL SECURITY
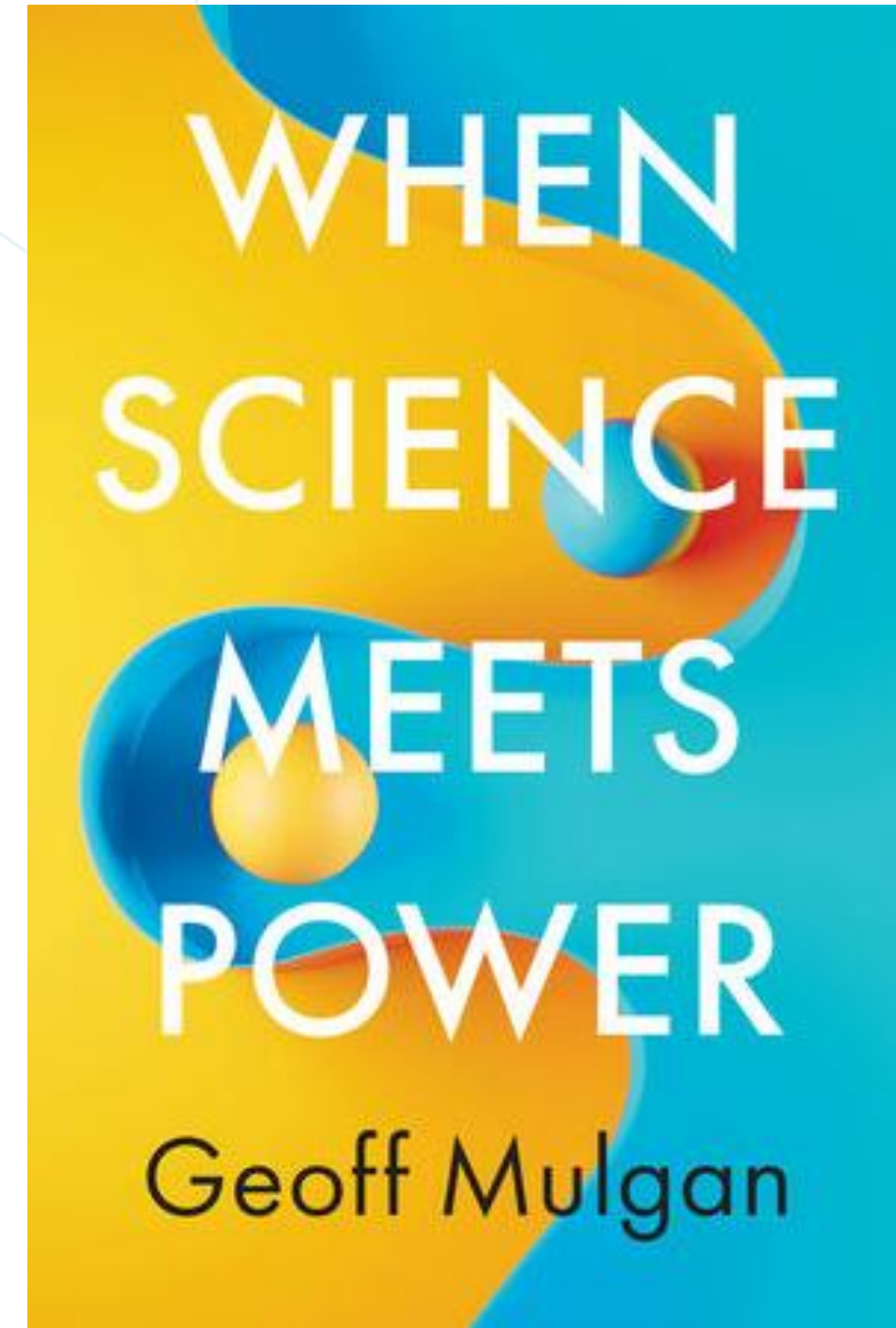- Throughout history **trickle-down prosperity** never worked

~~ETHICAL~~ POLITICAL ASPECTS OF EMERGENT TECHNOLOGIES

"... In the years that followed many came to view much of the work on **AI ethics as a distraction rather than a solution**. Dozens of books and conferences worried about issues such as the moral dilemmas of '**trolly problems**'... Others explored the possible dynamics of '**the singularity**' a generation or more into the future, when AI might suddenly take off and dramatically supersede human intelligence.

Yet, at the time, **AI was being applied in almost every area of daily life with very little serious analysis of the social and ethical issues**... The work on the distant horizon was **intellectually appealing** but arguably a **diversion from the harder problems of the present**"

(p. 78, 2023)

WHEN SCIENCE MEETS POWER

Geoff Mulgan

Peter Novitzky
Senior Research Fellow (UCL / PETRAS)
Assoc Prof — Ethics of Emergent Technologies (Avans)

22 May 2024

Thank you!

p.novitzky@avans.nl

p.novitzky@ucl.ac.uk