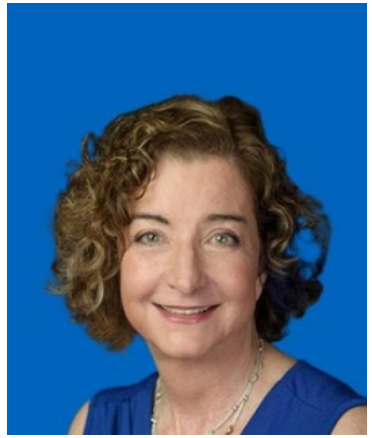
Supercharging Research: Harnessing Artificial Intelligence to Meet Global Challenges

Presidential Council of Advisors on Science and Technology (PCAST)

April 23, 2023

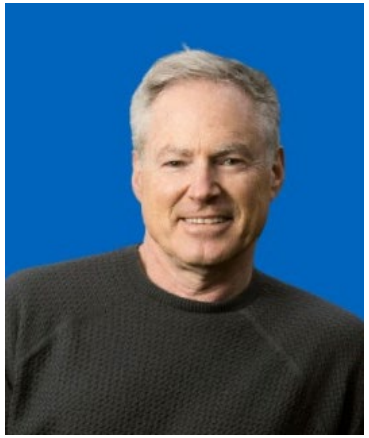
Working group on AI



Laura Greene

John Banovetz

Bill Dally



Eric Horvitz

Jon Levin

Terence Tao

Saul Perlmutter

Bill Press

Lisa Su

Phil Venables

Consults

- Ben Buchanan – NSC
- Alan Mislove – OSTP
- Paula Goldman - NAIAC
- David Danks - NAIAC
- Liz O'Sullivan - NAIAC
- Jatin Aythora – BBC
- Andrew Jenks – Microsoft
- Andy Parsons – Adobe
- Robert Chesney – U. Texas at Austin
- Hany Farid – Berkeley
- Ramayya Krishnan – NAIAC
- Haniyeh Mahmoudian – NAIAC
- Eun-Ah Kim – Cornell University
- Justin Vail – WHO
- Rian Bahran – OSTP
- Nik Marda – OSTP
- Michael Jordan – UC Berkeley
- Wade Shen – OSTP
- Manish Parashar – OSTP
- Tapio Schneider – Caltech
- Anima Anandkumar – Caltech

Executive Order on AI

- On October 30, 2023, President Biden issued an extensive **Executive Order (EO) on AI**, covering a wide range of AI-related issues.

- PCAST was tasked to report on *"the potential role of AI, especially given recent developments in AI, in research aimed at tackling major societal and global challenges."* **This report fulfills this tasking**.

- Broader impacts of AI, outside of scientific research, were tasked to other agencies and groups and are **beyond the scope of this report**.

Synergy between science and AI

- AI is starting to remove many of the barriers that make scientific research slow, expensive, or restricted to a small number of experts.
- The fundamental science of AI can mitigate errors, biases, and other potential harms of AI.
- Our existing scientific values (e.g., validation, reproducibility, openness, expert supervision) will help ensure developing a **culture of responsible use of AI** methods in science.
- There is a **strong use case** for use AI tools in science.

AI transforming science

AI can transform **every** scientific discipline, and **every** aspect of the scientific workflow. We will briefly describe examples in eight representative disciplines to illustrate three representative directions:

I. **Identifying candidate solutions** to scientific problems.
II. **Accelerating and enhancing** scientific simulations and models.
III. **Analyzing new types** of data.

(Other directions include: automation of laboratory work; assistance with coding or writing; and enabling new types of collaboration and engagement.)

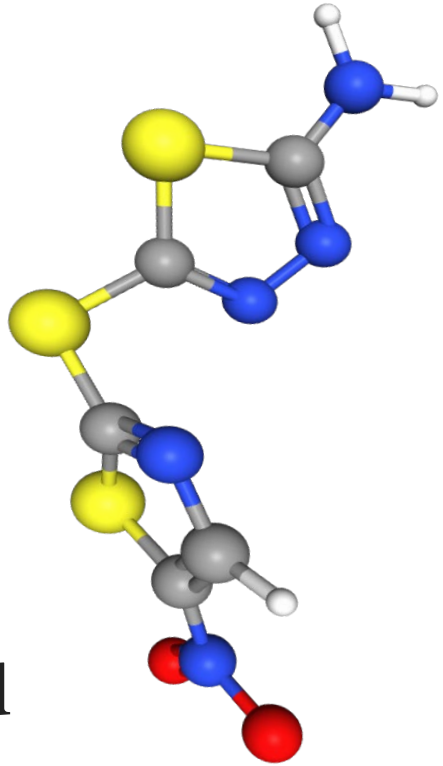
I. Identifying solutions (1/3): materials science

- **Already**, predicted millions of new stable materials that can be automatically synthesized in a laboratory.

- **In the future**, discover room temperature superconductors, new battery materials for electric vehicles…

I. Identifying solutions (2/3): therapeutic drugs

- **Already:**
  - Identified new compounds including the molecule shown here which is a promising new antibiotic;
  - Led to the discovery of a new anti-microbial mechanisms;
  - Used to design vaccines and targeted therapies, such as blocking the binding regions of viruses.

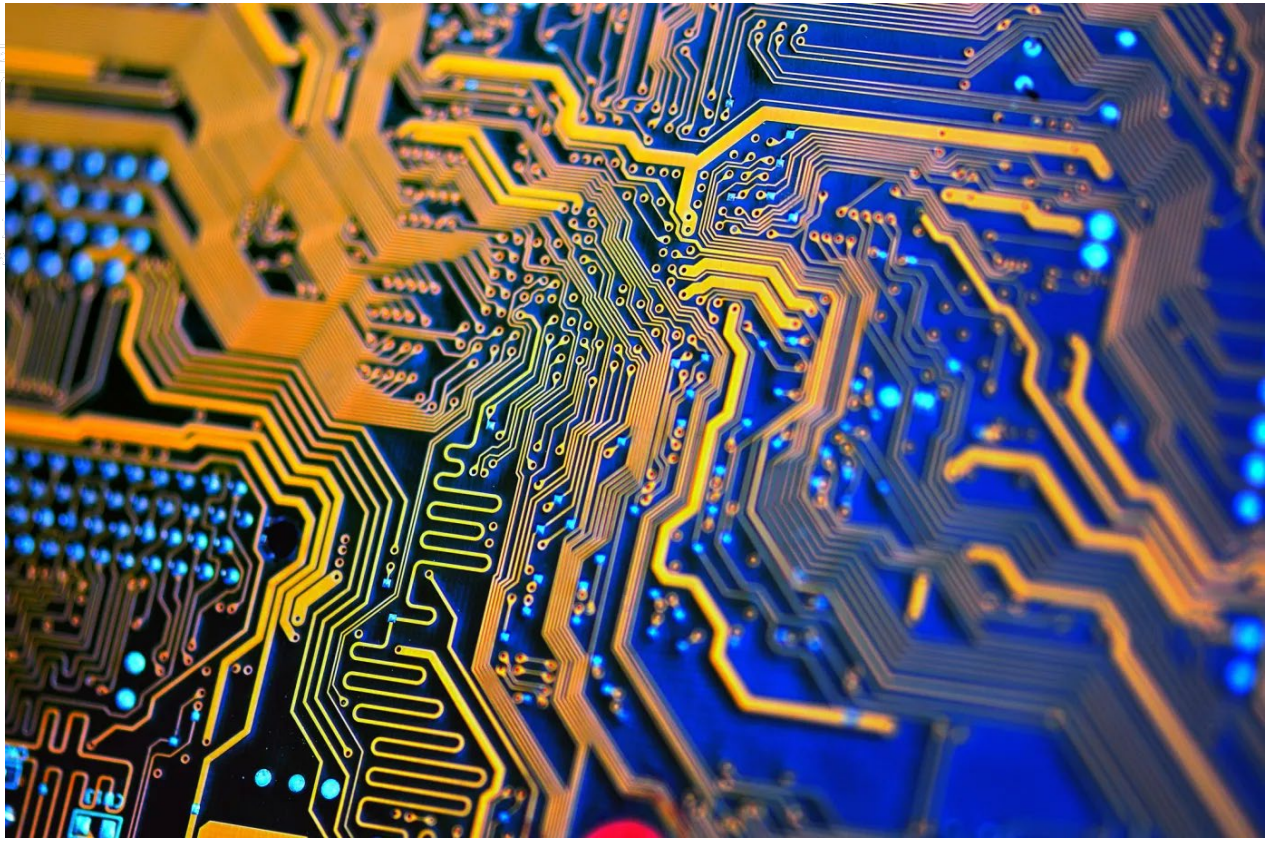


- **In the future**, custom-tailored therapies for cancer and auto-immune disorders, prolonging the vibrancy of cells, and deeper understanding of the mechanisms of antibiotics.

I. Identifying solutions (3/3): semiconductor design

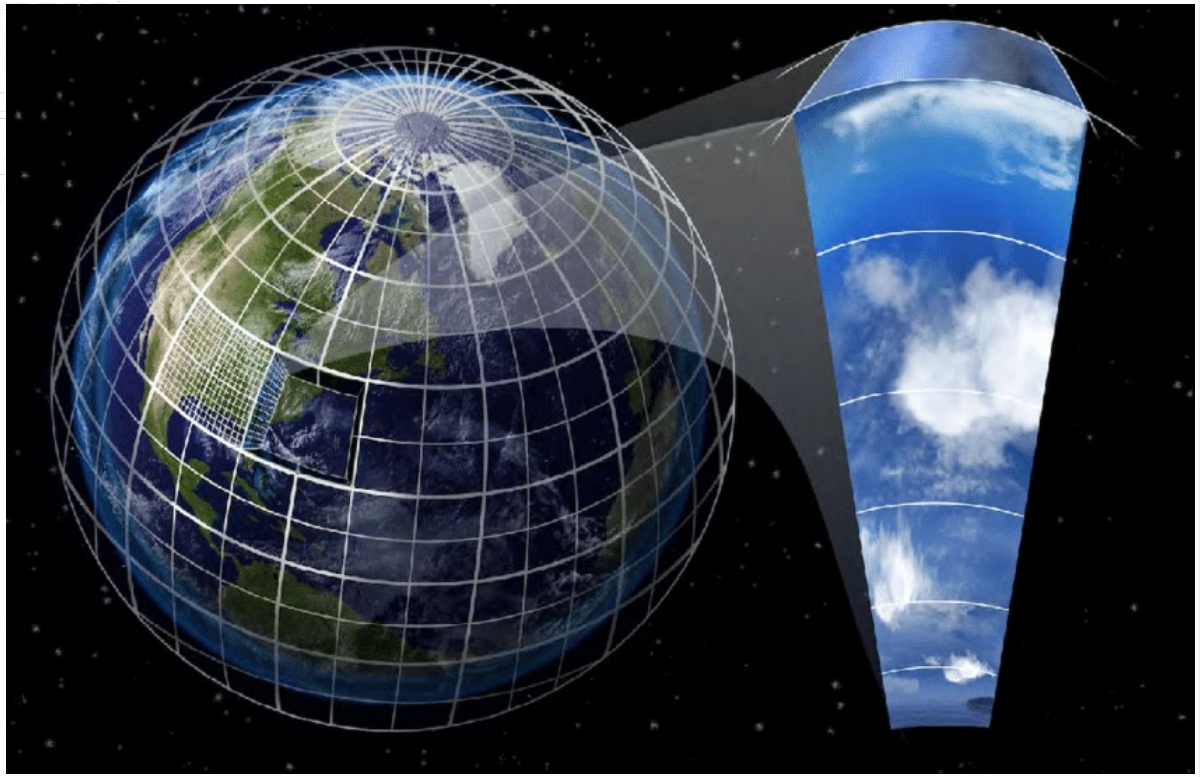
- **Already**, being used to generate initial circuit designs for chips.

- **In the future**, will perform lower-level chip design tasks, improving designer productivity by more than an order of magnitude.



II. Accelerating modeling (1/3): climate science

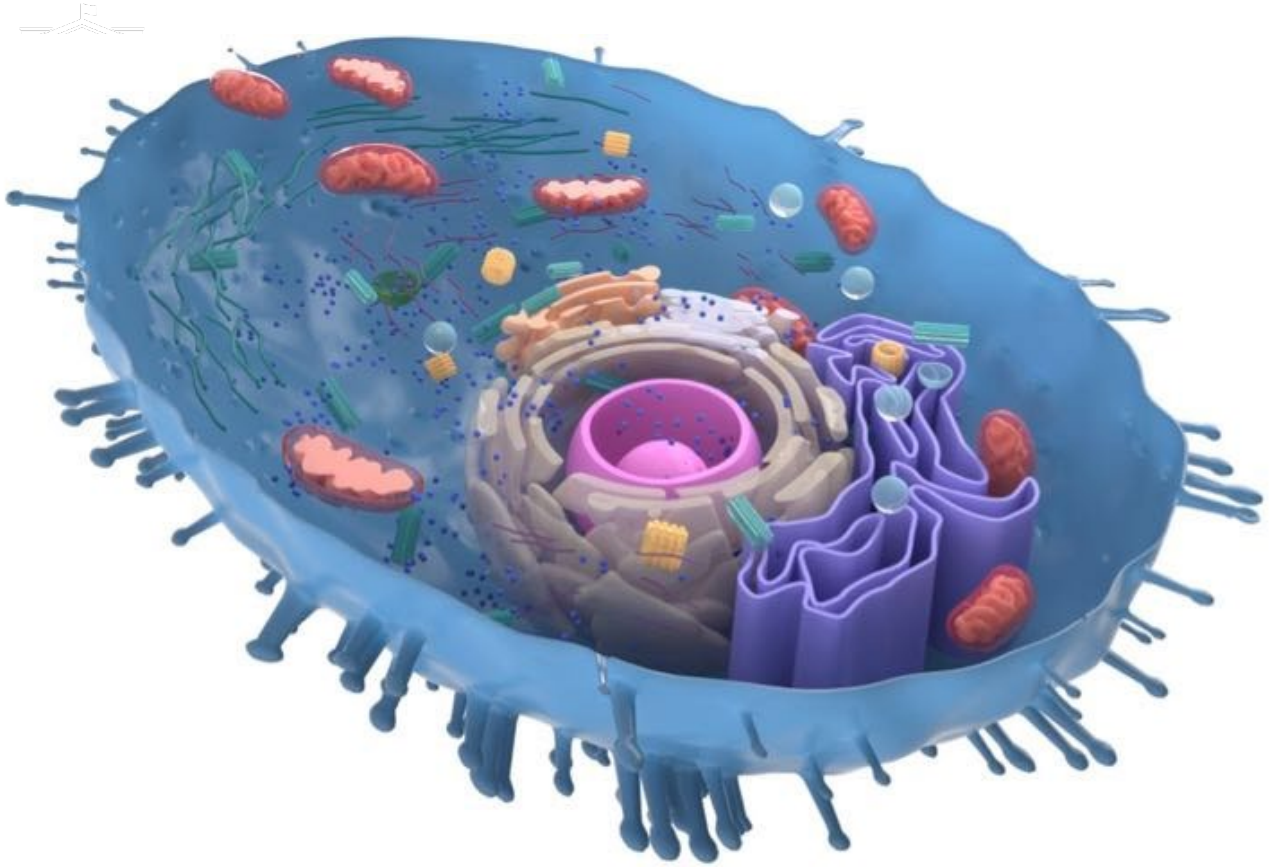
- **Already**, can greatly accelerate some aspects of weather forecasting or climate modeling, as much as a thousandfold.
- **In the future**, whole Earth models will advance many areas of science, including catastrophe modeling, water management, greenhouse gas monitoring…



II. Accelerating modeling (2/3): cellular biology

- **Already**, has successfully predicted the structure and function of proteins.

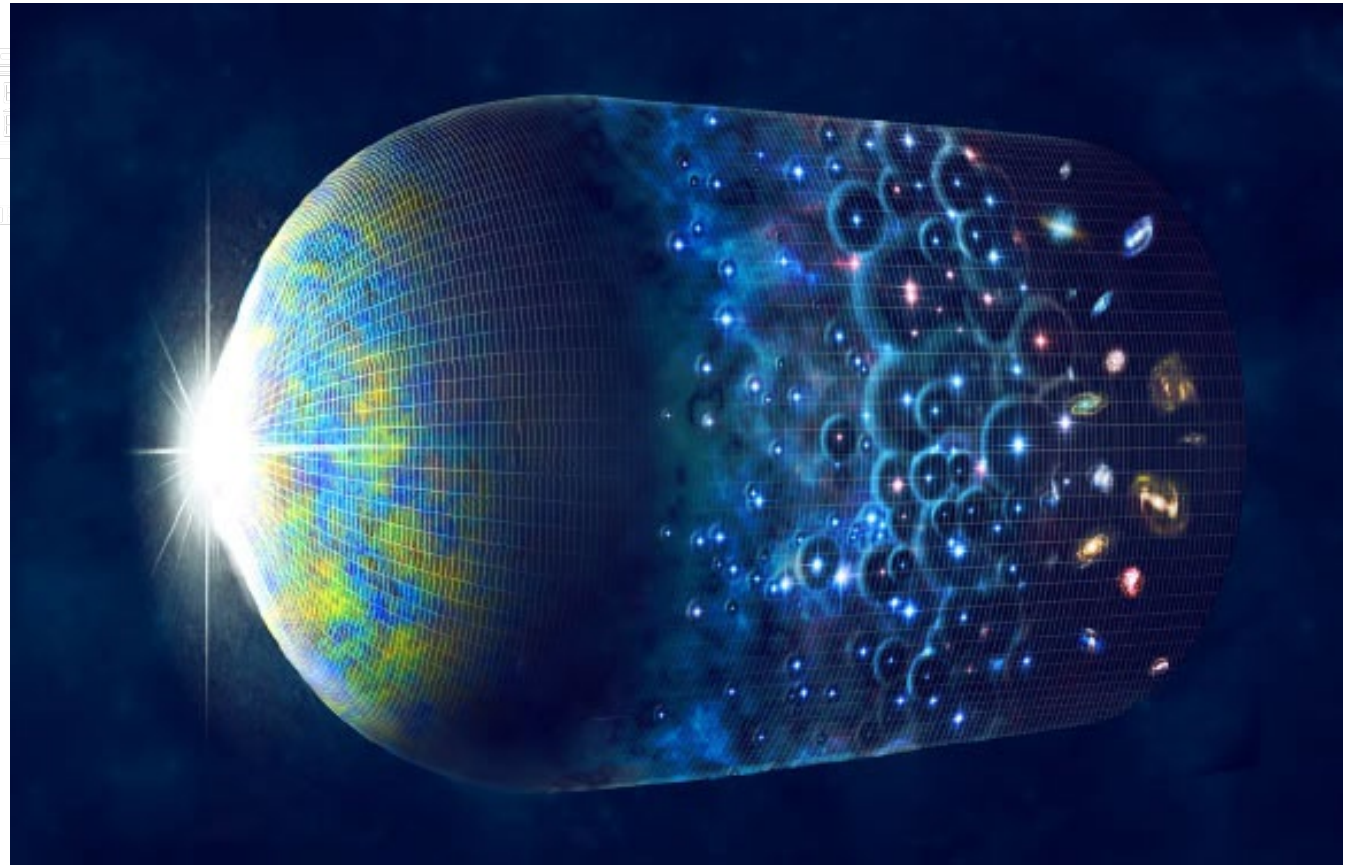
- **In the future**, models of the cell will unlock the secrets of cellular biology and allow new therapies to be tested *in silico*.



II. Accelerating modeling (3/3): cosmology

- **Already**, permits rapid simulation of cosmological hypotheses on the nature of the universe.

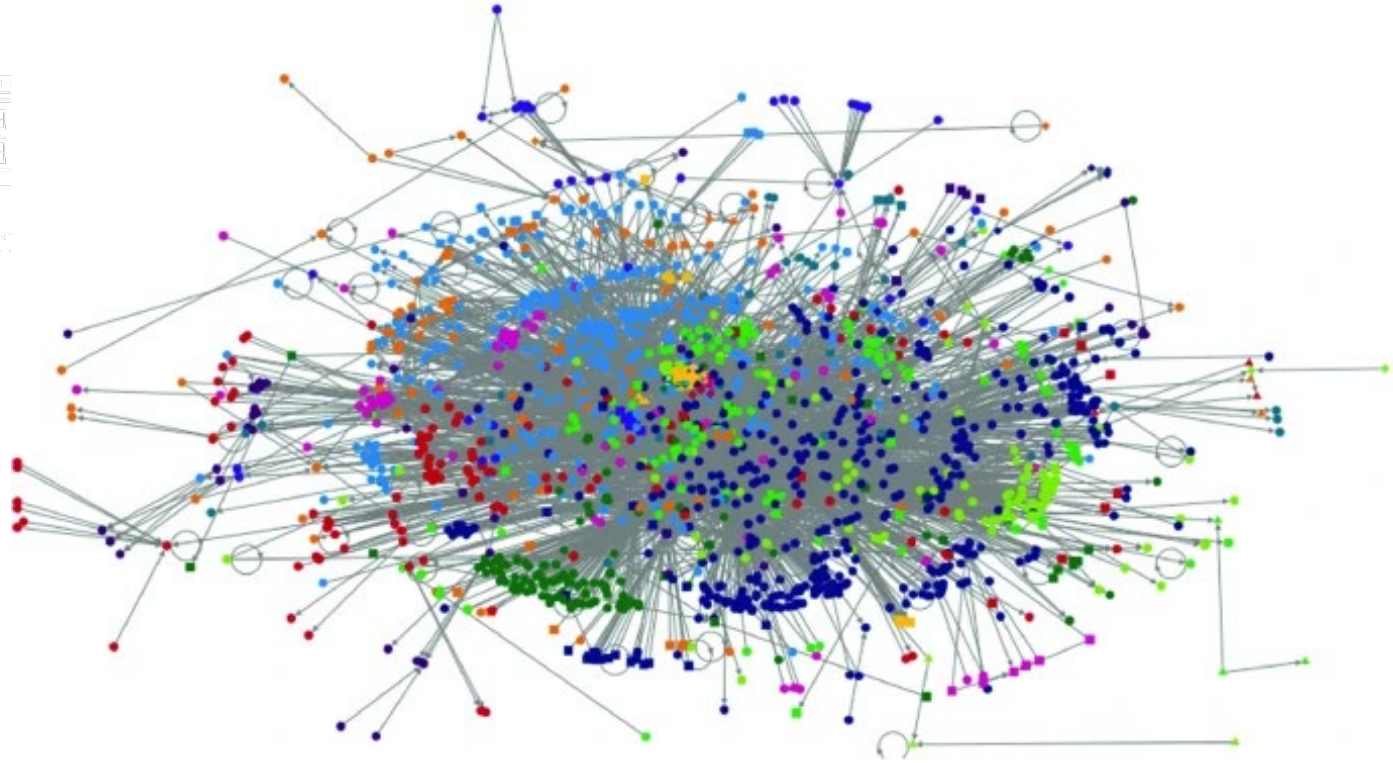
- **In the future**, could generate new fundamental theories in physics.



III. New types of data (1/2): social science

- **Already**, large language models analyze non-quantitative data, such as large-scale behavioral data in social media posts.

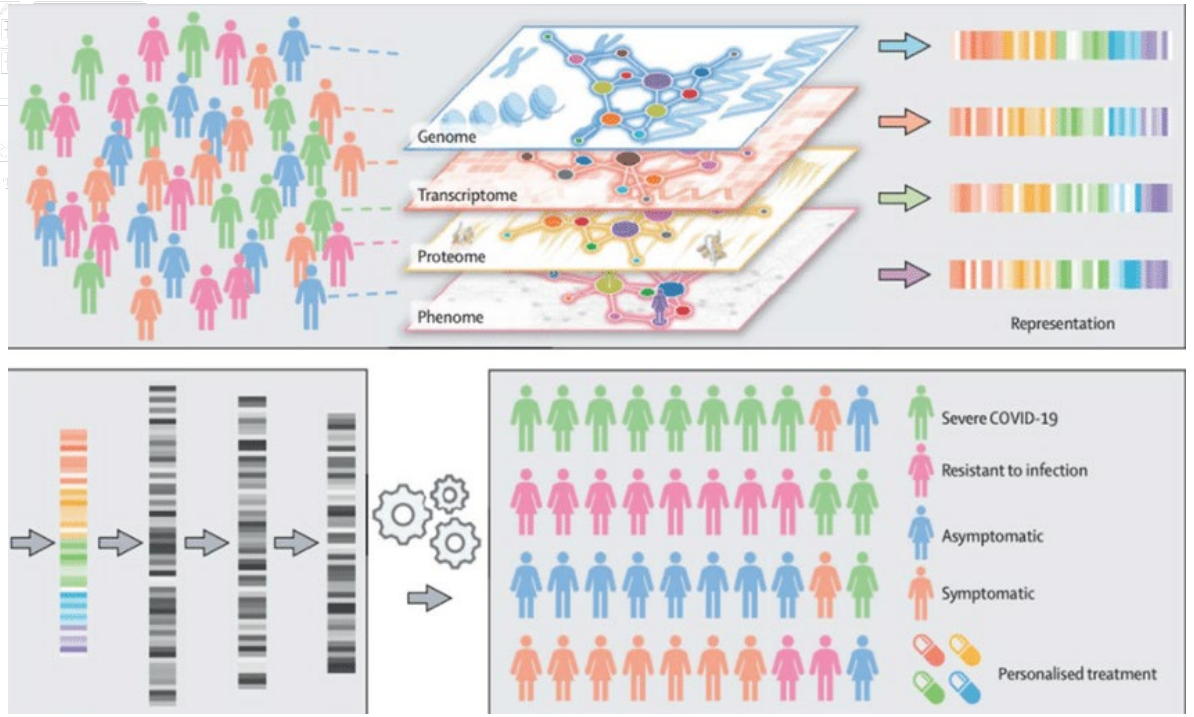
- **In the future**, data-driven social science will enable more effective, responsive, and fair delivery of public services.



III. New types of data (2/2): health and wellness

- **Already**, aids physicians to successfully make early diagnosis of cancer and detect potential errors to enhance patient safety.

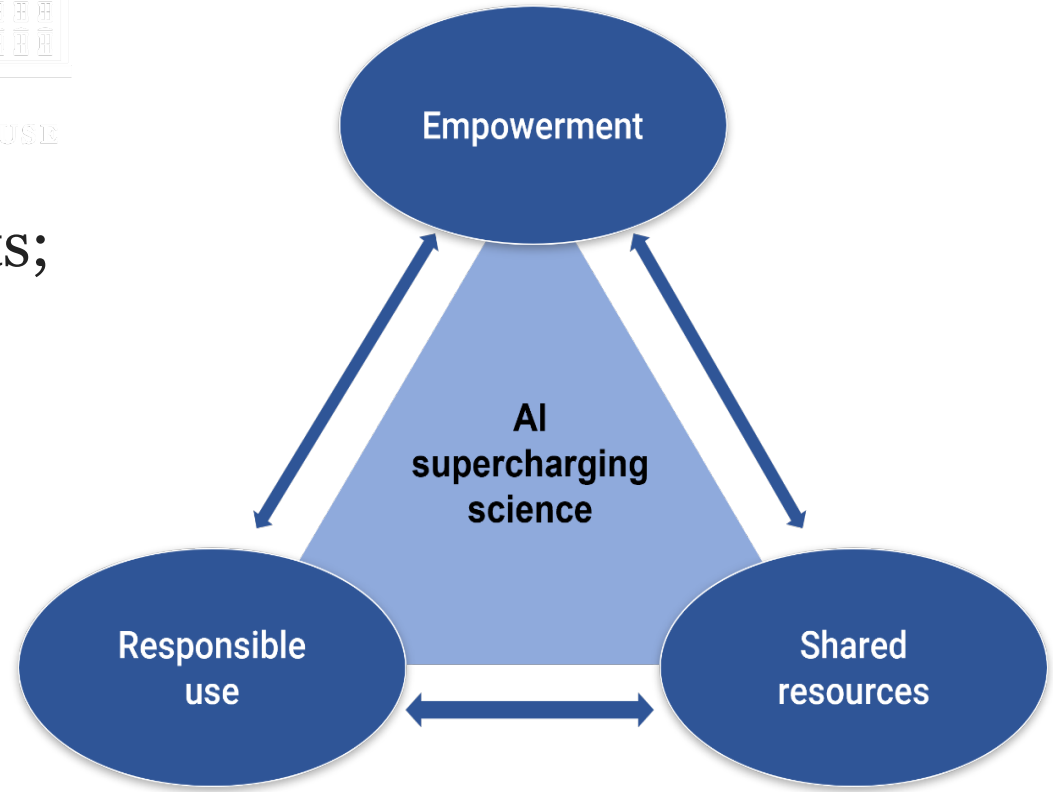
- **In the future**, ultra-personalized medicine will tailor health care to an individual's specific genetics and medical history.



Our vision of AI-assisted research and development

To gain the greatest net benefit from AI adoption in science, we envisage pursuing three broad goals:

I. **Empowerment** of human scientists;
II. **Responsible use** of AI tools; and
III. **Sharing** of AI resources.



Five recommendations to advance these three goals follow:

I. Empowerment of human scientists (1/3)

- **AI assistants** can be developed to **complement** and **empower** human scientists, rather than **replace** them.
- **AI tools** can be harnessed to process huge streams of data, increasingly manage such tasks as laboratory work, coding, and writing so identify promising solutions to scientific problems.
- This can enable **human scientists to focus on high-level directions**.
- **New ways of collaborating** can be developed, such as human scientists directing a network of interlinked AI assistants and make possible extremely large, interdisciplinary, and/or decentralized projects.

I. Empowerment of human scientists (2/3)

Recommendation: Support both basic and applied research in AI that involves collaborations across academia, industry, national laboratories, and federal agencies.

- An existing example is NSF's **Materials Innovation Platforms**, which is developing a data sharing infrastructure while employing AI tools as part of the community building with **other agencies and industrial partners**.

- Future projects could include collaborations to develop next-generation quantum computing qubits, whole cell modeling, whole Earth foundation models, or high-quality scientific databases in a broad range of disciplines.

I. Empowerment of human scientists (3/3)

Recommendation: Encourage innovative approaches to integrating AI assistance into scientific workflows.

- Funding agencies should recognize the emergence of new workflows and design flexible procedures, metrics, funding models, and problems that encourage strategic experimentation with new AI-assisted ways to organize a scientific project.

- Although AI-assisted science can be superior to either unassisted human science or fully automated science, these tools must be complemented by traditional human scientific research.

II. Responsible use of AI (1/2)

- A **culture of responsible AI use** in science should be be developed, in which AI outputs are externally verified, protections for private data are in place, algorithmic bias is measured and compensated for, and models and data are as transparent, replicable, open, and explainable as possible.

- This culture should be shaped by continuous dialogue between the **physical and social sciences, the humanities, and policy makers**.

- With the development of more **lightweight models and shared resources**, environmental impacts and other costs of AI use will likely be reduced.

II. Responsible use of AI (2/2)

Recommendation: Adopt principles of responsible, transparent, and trustworthy AI use throughout all stages of the scientific research process.

- AI tools could **strengthen scientific integrity**, e.g., in detecting algorithmic biases and enabling more efficient replication and validation of scientific findings, protecting personal privacy, and intellectual property rights.

- Funding agencies should require **responsible AI use plans** from researchers that would assess potential AI-related risks, including aligning with the principles called out in the blueprint for the AI Bill of Rights and the NIST AI Risk Management Framework as models.

III. Shared and open resources (1/3)

- **Multiple open-source models** that are powerful enough for research applications will emerge avoiding dependence on a few private technology companies.
- Through **shared AI infrastructure** (e.g., models, data, standards, best practices), redundant efforts can be avoided, access to AI for research purposes can be democratized, and the time and resources needed to use these tools can be significantly reduced.
- Significant scientific **federal data sets can be shared**, as noted in recent PCAST Reports: FEMA data to improve catastrophe modeling; defense satellite data to improve wildfire prediction; healthcare data to make patient safety measures more effective; and much more.

III. Shared and open resources (2/3)

Recommendation: Expand existing efforts to broadly and equitably share basic AI resources.

- We strongly recommend that the **National Artificial Intelligence Research Resource (NAIRR)** be fully funded, as an important **first step** towards establishing a broader national AI infrastructure of compute, data, practices, and human expertise, that can compete with private and international efforts.
- In the longer term, **NAIRR could be used as a stepping-stone** for even more ambitious "moonshot" programs, such as multimodal foundation models of complex systems such as the Earth or the human cell.

III. Shared and open resources (3/3)

Recommendation: Expand secure access to federal data sets for approved critical research needs, with appropriate protections and safeguards.

- We encourage **expansion** of the **National Secure Data Service Demonstration (NSDS-D)** and the **Federal Statistical Research Data Centers** with strengthened enforcement of data sharing mandates for federally funded research.
- **Shared AI infrastructure** may make feasible the upgrading of raw federal data sets to more valuable highly curated databases to democratize resources and benefit society.

Conclusion / Executive summary

- Scientific research will be radically accelerated by a broad spectrum of AI technologies.
- **With the right AI infrastructure**, this will enable scientists to address urgent challenges.
- While powerful, AI has weaknesses (e.g., algorithmic bias) and can require enormous amounts of compute, energy, and data.
- There are ways to mitigate the weaknesses and reduce the net resources required, particularly in scientific applications. **When used responsibly, AI can unlock great advances in science and technology**.