

February 21, 2025

AI AND BETTER HEALTH: Trust and Governance

CUHK Centre for Bioethics
10th Anniversary Celebration



David Roscoe



THE
HASTINGS
CENTER

A Personal Note: My Interest in AI Started at The Hastings Center in 2016

My Forecast in 2018...

- > AI Will Transform Society by 2030
- > Progress is Unstoppable

Much is at stake

Future of work

Human values

Destiny of human species!

... And My Youngest Grandson Fred Will Be 16



So...
What does
Fred need
to know?

AI and Trust: A Focus of The Hastings Center for Ten Years

Our First AI Project: 2016-18

Project Overview

- Funded by the **Future of Life Institute**
- **AI thought leaders** in tech, academia, government
- Identify **Ethical and Social Issues** raised by AI

My Involvement in the Project

- Board member
- Project Participant
- Report Essayist: **“AI and Trust”**

“AI and Trust”: Identified Concerns

Users

- Safety
- Bias
- Transparency
- Explainability

Public

- Security and Privacy
- Lethal AI Weapons
- Fairness
- Truth

AI-Human Relationships

- False empathy
- Alignment with human values

AI Developers

- Will experts make the right choices for society?

The Modern AI Era: Explosive Growth, More to Come



Predictive AI
2016

- › **Machine Learning:** Find patterns
- › **Perform Tasks:** Identify & predict



Generative AI
2022

- › **Language:** create fresh content
- › **Reasoning:** Solve problems



Agentic AI
2025

- › **Delegate:** autonomous action
- › **Enhance:** use tools, collaborate

Declared Goal:
Artificial General Intelligence by 2030!

Healthcare: A Top Priority for Global AI Companies

“Healthcare is AI's most pressing application.”
—Microsoft

“Healthcare is the largest or second-largest component of the economy.”
—Apple

“Leverage AI to meet the growing demand for equitable access to healthcare”
—Alibaba

AI Targets of Opportunity

- Improve population health
- Improve patient experience
- Enhance caregiver experience
- Reduce rising costs of care

AI and Healthcare: Accelerating Transformation

Prescriptive AI 2016

- › Specific Tools
Narrow Tasks

Generative AI 2022

- › Language Models
Reasoning Models

Agentic AI 2025

- › Autonomous Work
Digital Employees



Stated Goal

AGI 2030

- › Smarter than Humans
Solve “Intractable” Problems

Contributions to Better Health

- › **Diagnostics:** improved imaging
Viral spread: improved forecasting
New drugs: accelerated...Covid, AlphaFold
- › **Patient Records:** automated, less doctor burnout
Emergency Triage: faster, fairer
Complex Diseases: more accurate diagnoses
- › **Surgery:** robotic assistance
Telehealth: 24x7, proactive calls
Caregiving: nursing, therapy, eldercare
- › **Cancers:** earlier detection, personal cures
Pandemics: prevent
Health System: 10x efficiency improvement

AI and Healthcare: Unresolved Trust Issues

Users

- **Transparency:** Am I dealing with AI?
- **Explainability:** Is diagnostic efficacy sufficient?
- **Privacy:** can my data be used to train future AI?

AI-Human Relationship

- **Eldercare:** Empathy, or emotional manipulation?
- **Exploiting “Loneliness”:** risks to mental health?

Public

- **Accountability:** Who is liable for AI errors?
- **Training Bias:** marginalized communities?
- **Benefits:** equitably accessible?

AI Developers

- **Models:** Proprietary or open source?
- **AGI:** Will the experts be ready in 5 years?

AI and Healthcare: Keeping Pace with Furious Progress

Governance Challenges

- > **Balance: Innovation vs. Risk**
- > **Priorities: Immediate vs. Emerging**
- > **Scope: Global vs. Regional**

Guardrails, not Stop Signs

- > **Philosophy: Principles vs. Rules**
- > **Monitor and Adapt: → Health Learning Systems**
- > **Keep an Eye on AGI: Five years** is not that far away!

Closing Thoughts

CUHK Centre for Bioethics: A Thought Leader and Role Model

- > **Develop Regional Governance**
 - Robust, nimble approach
 - Reflecting South Asian culture and values
- > **Global Collaboration**
 - Seeking common ground with other regional cultures and values

The Hastings Center: Proud of our Partnership

- > **Celebrating 10 years of Progress**
- > **Taking on the Challenges Ahead Together**

Fast Forward to 2030: My Birthday Message to Fred



(Fred is Now 11)

I'm now 87, and now you're 16. Your whole life lies ahead of you.
AI has made the world **healthier, safer and fairer.**
Fortunately, all the **needed guardrails are in place.**
(Or not.)

AGI is here now, so you'll never die of cancer.
Artificial Super Intelligence is coming Next.
That's both **exciting and scary...**it could be **either heaven or hell!**

I Only Wish I Could Be There With You!!!



Thank You For Your Attention!!
David Roscoe