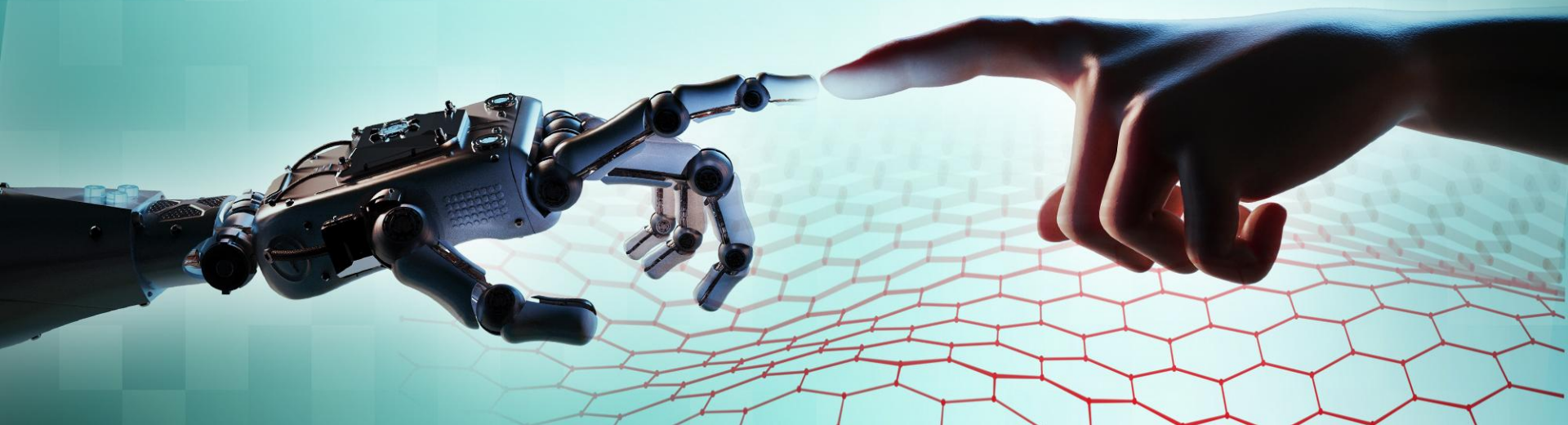


# From AI Ethics to Governance: An Overview of the Landscape



**Hoda Heidari**

K&L Gates Career Development  
Assistant Professor in Ethics and  
Computational Technologies



K&L Gates Initiative in  
**ETHICS & COMPUTATIONAL  
TECHNOLOGIES**

# AI in Medicine and Healthcare

## A.I. Chatbots Defeated Doctors at Diagnosing Illness

A small study found ChatGPT outdid human physicians when assessing medical case histories, even when those doctors were using a chatbot.

### ARTIFICIAL INTELLIGENCE

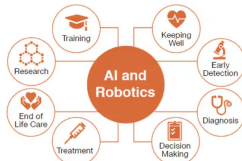
**AI is dreaming up drugs that no one has ever seen. Now we've got to see if they work.**

AI automation throughout the drug development pipeline is opening up the possibility of faster, cheaper pharmaceuticals.

No longer science fiction, AI and robotics are transforming healthcare



AI is getting increasingly sophisticated at doing what humans do, but more efficiently, more quickly and at a lower cost. The potential for both AI and robotics in healthcare is vast. Just like in our every-day lives, AI and robotics are increasingly a part of our healthcare ecosystem.



We have highlighted eight ways that showcase how this transformation is currently underway.

Blog > Healthcare & AI

## The Power of AI in Streamlining Healthcare Workflows

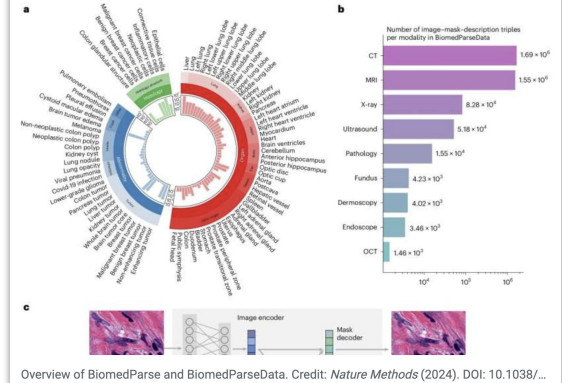
November 19, 2024 | 1 min watch

**VIDEO: AI may improve maternal health outcomes by identifying high-risk patients**

**AI program enhances echocardiogram interpretation and reduces wait times**

## Q&A: A new medical AI model can help spot systemic disease by looking at a range of image types

by Stefan Milne, University of Washington



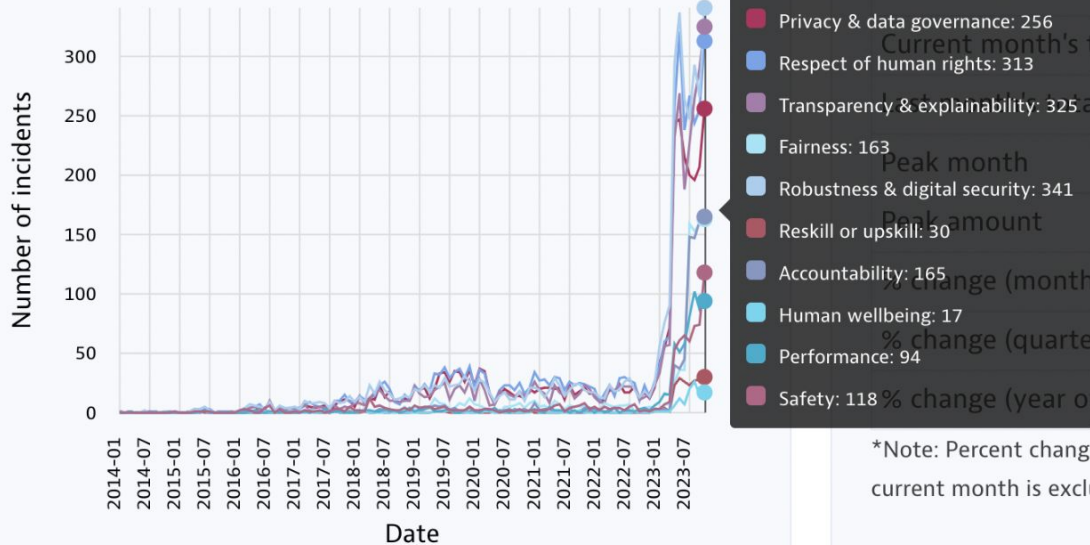
FORBES > LEADERSHIP

**Elevating Elderly Care With AI-Powered Digital Identity Systems**

# AI-related Incidents/Harms on the Rise

## Summary visualisations

Evolution of incidents by AI principle



## Summary statistics

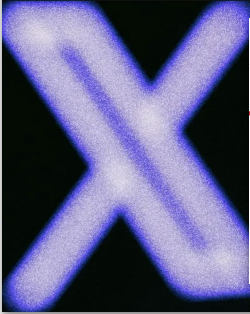
	Incidents	Articles
All time total	6264	36345
Current month's total	317	1768
Peak month	616	3227
Peak amount	<u>2023-10</u>	<u>2023-10</u>
% change (month-over-month)	23.2	51.22
% change (quarter-over-quarter)	13.01	13.87
% change (year over year)	961.58	690.9

\*Note: Percent change is calculated based on preceding full months (i.e. the current month is excluded).

NOV 20, 3:14 PM EDT / BY FRANK LANDYMORE

## EXPERTS ALARMED BY PEOPLE UPLOADING THEIR MEDICAL SCANS TO ELON MUSK'S GROK AI

"THIS IS VERY PERSONAL INFORMATION, AND YOU DON'T EXACTLY KNOW WHAT GROK IS GOING TO DO WITH IT."



MOTHERBOARD  
TECH BY VICE

## 'He Would Still Be Here': Man Dies by Suicide After Talking with AI Chatbot, Widow Says

naturemedicine

View all journals

Search

Log in

Explore content ▾ About the journal ▾ Publish with us ▾

nature > nature medicine > comment > article

Comment | Published: 26 August 2024

### Not all AI health tools with regulatory authorization are clinically validated

## ChatGPT is little help for doctors in diagnosing diseases, study finds

The research, conducted with 50 physicians last year, found that using ChatGPT did not significantly improve doctors' diagnostic reasoning.

## FDA strengthens AI regulation to ensure patient safety and innovation in healthcare

'Bias in, bias out': Tackling bias in medical artificial intelligence

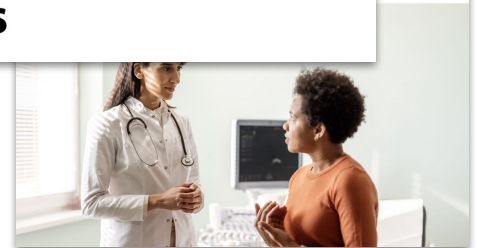
HEALTH AND HEALTHCARE SYSTEMS

## Billions of dollars have been invested in healthcare AI. But are we spending in the right places?

DIGITAL

## What doctors wish patients knew about using AI for health tips

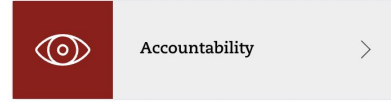
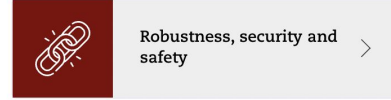
Health Equity and Ethical Considerations in Using Artificial Intelligence in Public Health and Medicine



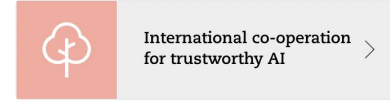
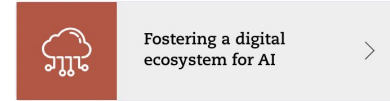
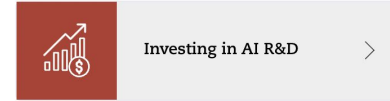
# Principles

- Fairness
- Accountability
- Transparency
- Safety and reliability
- Security
- Privacy
- ...

## Values-based principles



## Recommendations for policy makers



Safe and Effective  
Systems



Algorithmic  
Discrimination  
Protections



Data Privacy



Notice and  
Explanation



Human Alternatives,  
Consideration, and  
Fallback

# AI Governance Challenges

## 1- What to comply with?

- Developing **standards and best practices**
  - uncertain, fast-moving, dynamic environment
  - general-purpose technology
- Process to reach normative **consensus**
  - tradeoffs among various stakeholders
  - accounting for different contexts
  - While avoiding conflicting laws and policies

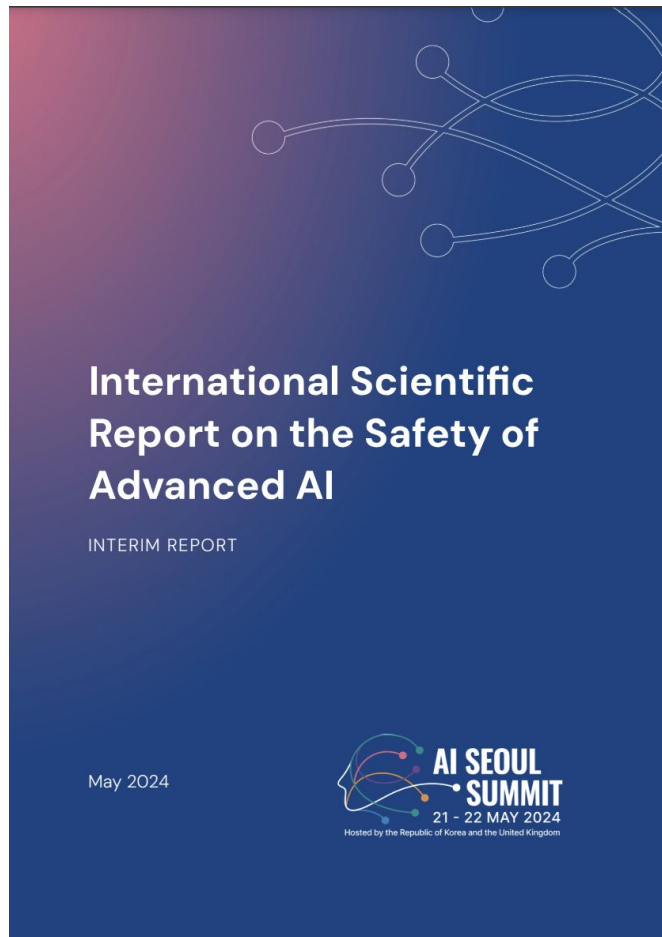
## 2- How to incentivize compliance?

- **Assessing compliance**
  - Information asymmetries and access
  - Model opaqueness and inscrutability
- **Attributing liability** for harms
  - Harms are diffuse and hard to foresee
  - Intentionality, Recklessness, Negligence
  - creator/ owner/ operator/ user?
- **Enforcing consequences**
  - Balancing retributive justice and deterrence with innovation and growth




# Challenges Unique to GenAI

- Unclear (hyped-up) **capabilities**
- Challenges of evaluating **risks**
  - Broad risk surface
  - Vacuum of robust methodologies



# Risk Assessment

1. Model testing
2. Red-teaming
3. Field testing



**NIST** NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY  
U.S. DEPARTMENT OF COMMERCE

**ARIA**

## Assessing Risks and Impacts of AI

A compelling set of scenarios will aim to explore risks and related impacts across three levels of testing: model testing, red-teaming, and field testing.

[Evaluation Plan](#) [Sign ups for APPLICATIONS TEAMS are open](#)

The banner features a dark blue background with glowing green and cyan digital lines and particles, creating a futuristic, data-driven aesthetic. The NIST logo is in the top left, and the ARIA logo is in a dark blue box on the left. The main title and text are centered, and two call-to-action buttons are at the bottom.



# Upcoming Event at CMU

Generative AI, Ethics, and Governance

- Early 2025 (stay tuned!)
- Expert panels and roundtables



**K&L Gates** Initiative in

**ETHICS & COMPUTATIONAL  
TECHNOLOGIES**

---

**Carnegie  
Mellon  
University**

The logo for K&L GATES, featuring the text in white on a teal square background. The background of the slide is a colorful, abstract pattern of diagonal lines in shades of red, orange, yellow, green, blue, and purple.

K&L GATES

## Navigating Ethical Challenges in Artificial Intelligence in MedTech

## Introduction: Use Cases and need for governance

- Blood loss AI tool
- Iron deficiency AI tool
- Is there a need for governance frameworks?

## Focus: Existing frameworks and guidelines in healthcare

- World Health Organization
- Singapore Health Sciences Authority
- US Food and Drug Administration
- Japan Pharmaceutical and Medical Devices Agency
- Korea Ministry of Food and Drug Safety

## Main AI Ethical Principles identified

- **Human centricity and oversight**
- **Explainability and transparency**
- **Fairness**
- **Safety and security**
- **Equity and quality of access**
- **Social and environmental well-being**

## Risk-based approach considerations

- **EU AI Act - Regulation (EU) 2024/1689**
  - Unacceptable risk AI systems (Article 5)
  - High-risk AI systems (Article 6)
  - Limited risk AI systems
  - Minimal risk AI systems



Questions?

K&L GATES