stfm.org

teach&
transform

Ethical Use of AI in the Family Medicine Clinic

- Ioannis Kakadiaris, PhD
- Vasiliki N. Rahimzadeh, PhD
- Samira A. Rahimi, B.Eng, PhD
- Winston Liaw, MD, MPH

Learning Objectives

- Identify key ethical considerations relevant to the use of artificial intelligence / machine learning (AI/ML) in primary care
- Discuss strategies for ethically using AI/ML tools to enhance patient care

Overview

1) Overview of the application of ethical principles to AI (Kakadiaris)
2) Examples of how AI challenges ethical principles (Rahimzadeh)
3) What can practices do to address ethical challenges (Rahimi and Liaw)

Poll Questions

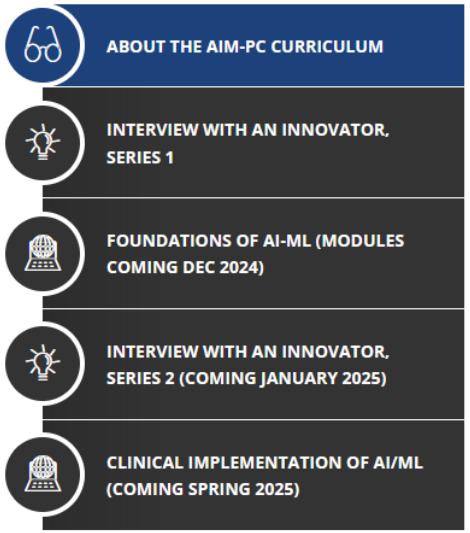
1. What best describes your role?
- Family physician
- Medical resident or student
- AI researcher or developer
- Health policy maker or administrator
- Patient of public representative
- Other

2. How familiar are you with the concept of ethical AI?
- Very familiar
- Somewhat familiar
- Heard of it, but not clear
- Not familiar at all

Overview

AiM-PC: Artificial intelligence & Machine learning for Primary Care

- **Vision**
  - Education in pursuit of AI-augmented care for all
- **Mission**
  - To elevate the practice of primary care by teaching learners to use artificial intelligence / machine learning (AI)
- **Audience**
  - Medical students, primary care residents, practicing primary care physicians
- **Launch**
  - Free
  - Email wliaw@central.uh.edu if you want to pilot the curriculum



teach & transform

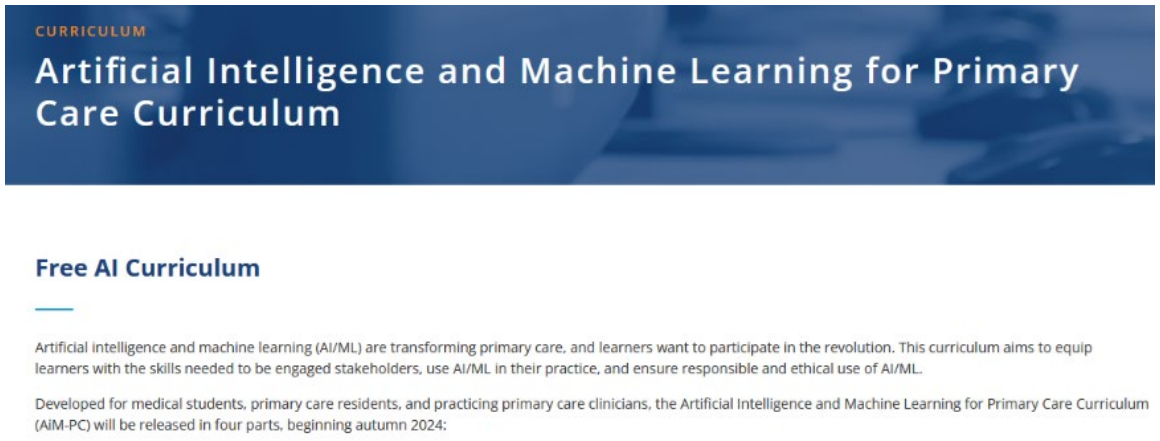
AiM-PC: Artificial intelligence and Machine learning for Primary Care

Modules

- AI Essentials
- Social and Ethical Implications of AI
- Evidence-Based Evaluation of AI-Based Tools
- AI-Enhanced Clinical Encounters
- Integrating AI into the Clinic

Features

- Learning activities
- Chatbot
- Video series: Interview with an Innovator



teach & transform

Video Series

Steven Lin, MD

- Clinical Professor, Stanford

Bob Wachter, MD

- Chair, Department of Medicine, University of California, San Francisco

Interview With an Innovator, Episode 2: Bob Wachter, MD

Katie Link

- Product Manager, NVIDIA
- Medical student, Icahn School of Medicine at Mount Sinai

Jenna Wiens, PhD

- Associate Professor, Computer Science, University of Michigan

Nipa Shah, MD

- Chair, Department of Community Health and Family Medicine, University of Florida, Jacksonville

Vasiliki Nataly Rahimzadeh, PhD

- Assistant Professor, Center for Medical Ethics and Health Policy, Baylor College of Medicine

Judy Gichoya, MD, MS

- Associate Professor, Department of Radiology and Imaging Sciences, Emory University School of Medicine

A Story

- Dr. Thomas has recently started using an AI-powered tool to transcribe her clinical notes. After using the tool for several weeks, she notices discrepancies in how the AI-based scribe has documented certain conversations and interpreted patient responses. Dr. Thomas is worried that the AI-based scribe's notes may be inaccurate and, worse still, be entering this inaccurate exchanges in the patient's electronic health record (EHR).
- Patient privacy, consent, and explainability take center stage at this point, along with considerations of how AI/ML tools can either reduce or exacerbate health disparities.
- Moreover, Dr. Thomas becomes concerned about patient privacy as she learns more about the data collection practices of the tool's developer. Where will patient data ultimately be stored? Who will have access to the recordings? Will the data be sold? If so, how will the data be used?
- These experiences force Dr. Thomas to address the ethical complexities of relying on AI/ML in health care and prompt her to advocate for greater transparency and oversight.

“Our future is a race between the growing power of our technology and the wisdom with which we use it. Let’s make sure that wisdom wins.”

- Stephen Hawking

Ethical Failures

- Research is important for the advancement of technology
- However, progress has too often been achieved at the expense of vulnerable populations, who have been exploited and subjected to experimentation without their consent, without disclosure of risks, and without any benefit to themselves

Ethical Failures

- Nazi experimentation
- Eugenics
- Nuremberg Trials

Courtesy of the National Archives

Ethical Failures

- Willowbrook Hepatitis Studies
- Tuskegee Syphilis Study
- Belmont Report

The Willowbrook State School. Credit: U.S. National Library of Medicine (NLM): Images from the History of Medicine

Ethical Failures

- Henrietta Lacks

Statue of Henrietta Lacks by sculptor Helen Wilson-Roe at Royal Fort House, Bristol. This file is licensed under the Creative Commons Attribution-Share Alike 4.0 International license.

teach & transform

Responsible AI Framework for Decision Support Systems

teach & transform

Need, Knowledge Gap and Objective

- **Need**
- A universal accountability framework for diverse industries and algorithms is missing

- **Knowledge Gap**
- Current high-level regulations are difficult for industries to implement practically

- **Objective**
- Develop an algorithm accountability benchmark

Value Proposition

- **AI Developers and Deployers**
- A benchmark for assessing AI systems' compliance with accountability standards

- **Legislators and Policymakers**
- A reference guide for shaping future AI legislation, policies, and monitoring compliance

- **Researchers**
- State-of-the-art benchmark for AI systems

System Cards for AI-Based Decision-Making for Public Policy

- **Need**
- Lack of an accountability framework that can be applied across a comprehensive range of algorithms being used in public policy

- **Knowledge Gap**
- Standards of accountability reflecting current legal obligations and societal concerns have lagged algorithms' extensive use and influence

- **Objective**
- Develop an algorithm accountability benchmark

- **Value Proposition**
  - **AI Developers and Deployers**: A benchmark to evaluate AI systems' compliance to accountability standards
  - **Legislators and Policymakers**: A reference guide for shaping future legislation and policies on AI
  - **Researchers**: State-of-the-art comprehensive benchmark for AI systems

F. Gursoy and I. A. Kakadiaris, "System Cards for AI-Based Decision-Making for Public Policy," 2022, arXiv:2203.04754

System Cards for AI-Based Decision-Making for Public Policy

- **System Accountability Benchmark (SAB)**
  - Fifty-six criteria organized within a framework of four-by-four matrix

Rows:

- Data
- Model
- Code
- System

Columns:

- Development
- Assessment
- Mitigation
- Assurance

	Development	#	Assessment	#	Mitigation	#	Assurance	#
Data	Data Dictionary	C111	Privacy, Data	C211	Anonymization	C311	Data Protection	C411
	Datasheet, Collection Process	C112	Fairness, Data	C212	Security	C312	Datasheet, Maintenance	C412
	Datasheet, Composition	C113	Quality, Labels	C213			Datasheet, Uses	C413
	Datasheet, Motivation	C114	Inspectability	C214				
	Datasheet, Preprocessing	C115						
Model	Reproducibility, Model	C121	Interpretability	C221	Adversarial, Training	C321	Privacy, Model	C421
	Design Transparency, Model	C122	Fairness, Model	C222	Explanations, Mitigation	C322	Uses, Model	C422
	Documentation, Model	C123	Testing, Adversarial	C223	Fairness, Mitigation	C323	Documentation, Capabilities	C423
	Selection, Model	C124			Privacy, Training	C324	Explainability	C424
Code	Reproducibility, Code	C131	Privacy, Code	C231	Review, Code	C331	Certification, Developer	C431
	Design Transparency, Code	C132	Security, Code	C232	Diversity, Team	C332	Due Diligence	C432
	Documentation, Code	C133	Testing Cards	C233				
System	Documentation, Development	C141	Awareness, Public	C241	Monitoring, Fairness	C341	Record Keeping, Operational	C441
	Plans, Maintenance	C142	Risk, Humans	C242	Monitoring, Performance	C342	Uses, System	C442
			Training, Operator	C243	Oversight, Human	C343	Documentation, Acceptability	C443
			Accuracy, System	C244	Harms, Remedies	C344	Insurance	C444
					Mechanism, Feedback	C345	Rating, Risk	C445
					Security	C346		

System Cards for AI-Based Decision-Making for Public Policy

- **System Cards**
  - Overall outcome of the evaluation for a specific decision-aiding system

- Each circle corresponds to a column in the SAB
- Each quarter corresponds to a row in the SAB
- Each arc corresponds to a criterion, a cell in the SAB
- Colors reflect evaluation outcomes

f (Machine Learning-Based Decision-Aiding Systems, System Accountability Benchmark) = System Cards

System Card+ Accountability Benchmark

1. **Accountability Performance Layer:** Measures and evaluates the performance and accountability of AI systems (baseline requirements)
2. **Fairness Layer:** Evaluates AI systems are fair and unbiased in their operations and outcomes
3. **Inclusivity Layer:** Assess the inclusion of diverse groups and perspectives in AI system design and implementation
4. **Ethical Compliance Layer:** Monitors AI systems are aligned with ethical standards and guidelines
5. **Legal Compliance Layer:** Ensures AI systems adhere to all relevant laws and regulations

System Card+ for AI-Based Decision-Support Systems

Each layer contains a four-by-four matrix of evaluation criteria.

Rows

1. **Data**
   Refers to the information collected, processed, and used by the AI system.
2. **Model**
   The algorithms and methods used to process data and make decisions.
3. **Code**
   The programming and software implementations that support the model and system.
4. **System**
   The overall infrastructure, including hardware and software, that hosts and executes the AI model.

Columns

1. **Development**
   The process of creating and building the components of the AI system.
2. **Assessment**
   Evaluating the performance, accuracy, and effectiveness of the AI components.
3. **Mitigation**
   Identifying and addressing potential risks and issues within the AI system.
4. **Assurance**
   Ensuring that the AI system meets required standards and performs reliably.

System Card+: Layer - Performance Accountability

Category	Development	Assessment	Mitigation	Assurance
Data	Data Dictionary (A111)	Inspectability (A211)	Anonymization (A311)	Data Protection (A411)
	Datasheet, Collection Process (A112)			Datasheet, Maintenance (A412)
	Datasheet, Composition (A113)			Datasheet, Uses (A413)
	Datasheet, Motivation (A114)			
	Datasheet, Preprocessing (A115)			
Model	Reproducibility, Model (A121)	Interpretability (A221)	Adversarial Training (A321)	Uses, Model (A421)
	Design Transparency, Model (A122)	Testing, Adversarial (A223)	Explanations, Mitigation (A322)	Documentation, Capabilities (A423)
	Documentation, Model (A123)			Explainability (A424)
	Selection, Model (A124)			
Code	Reproducibility, Code (A131)	Testing Cards (A231)	Review, Code (A331)	Certification, Developer (A431)
	Design Transparency, Code (A132)	Compliance Review (A232)	Diversity, Team (A332)	Due Diligence (A432)
	Documentation, Code (A133)			
System	Documentation, Development (A141)	Training, Operator (A241)	Oversight, Human (A341)	Record Keeping, Operational (A441)
	Plans, Maintenance (A142)		Mechanism, Feedback (A345)	Uses, System (A442)
		Security (A346)		Documentation, Acceptability (A443)
				Insurance (A444)
				Rating, Risk (A445)

- Ensures transparency and reliability
- Implements robust evaluation methods
- Enhances trust and accountability
- Promotes continuous improvement

- Tibebu, H., & Kakadiaris, I. A. "System Card+: A Theoretical Framework for AI-Based Responsible Decision Support Systems." Proc of the IEEE International Conference on Big Data (IEEE BigData 2024), Washington, DC, USA, 2024
- F. Gursoy and I. A. Kakadiaris, “System Cards for AI-Based Decision-Making for Public Policy,” 2022, arXiv:2203.04754

System Card+: Layer - Fairness

Category	Development	Assessment	Mitigation	Assurance
Data	Equity, Data (B111)	Fairness, Data (B211)	Bias Mitigation, Data (B311)	Third Parties assessment, Data (B411)
	Bias Detection, Data (B112)	Impact Assessment, Data (B212)	Rebalancing Techniques (B312)	Transparency Reports, Data (B412)
	Sampling Integrity, Data (B113)			
	Fairness Metrics, Data (B114)			
Model	Fairness-by-Design, Model (B121)	Fairness Metrics, Model (B221)	Bias Mitigation, Model (B321)	Third Parties assessment, Model (B421)
	Bias Detection, Model (B122)	Sensitivity Analysis, Model (B222)	Fair Optimization, Model (B322)	Transparency Reports, Model (B422)
	Fairness Benchmarking, Model(B123)			
Code	Automated Testing, Code (B131)	Compliance Review, Code (B231)	Fairness Refactoring, Code (B331)	Compliance Certificates, Code (B431)
	Audit Trails, Code (B132)	Fairness Test, Code (B232)	Equity Enhancements, Code (B332)	Transparency Reports, Code (B432)
System	Alert, System (B141)	User Feedback, System (B241)	Operation Protocols, System (B341)	Assurance Policies, System (B441)
	Stress Tests, System (B142)	Demographic Performance, Syste (B242)	Calibration, System (B342)	Fairness Logs, System (B442)
	Impact Assessments, System (B143)			

- Enhances fairness through refactoring
- Promotes demographic performance reviews
- Ensures equity and bias detection
- Implements fairness metrics and assessments

System Card+: Layer - Inclusivity

Category	Development	Assessment	Mitigation	Assurance
Data	Inclusion Criteria, Data (C111)	Coverage Evaluation, Data (C211)	Translation Mitigation (C311)	Third Parties assessment, Data (C411)
	Language Variety, Data (C112)	Translation assessment, Data (C212)	Language Expansion, Data (C312)	Transparency Report, Data (C412)
	Diverse Demography, Data (C113)			
Model	Selection Criteria, Model (C121)	Inclusion Impact, Model (C221)	Bias Mitigation, Model (C321)	Third Parties assessment, Model (C421)
	Interpretability, Model (C122)	Cultural Appropriateness, Model (C222)	Representation Repair, Model (C322)	Accessibility Reports, Model (C422)
			Accessibility Enhancements, Model (C323)	Inclusion Monitoring, Model (C423)
			Cultural Sensitivity, Model (C324)	
Code	Inclusive Design, Code (C131)	Accessibility Audit, Code (C231)	Accessibility Enhancements, Code (C331)	Third Parties assessment, Code (C431)
	Language Support, Code (C132)			Compliance Audits, Code (C432)
System	Accessibility Standards, System (C141)	Cultural Testing, System (C241)	Cultural Feedback, System (C341)	Inclusive Logs, System (C441)
			Inclusive Updates, System (C342)	Third-party assessment, System (C442)

- Ensures diverse representation
- Implements accessibility measures
- Enhances cultural appropriateness
- Promotes inclusive practices

System Card+: Layer - Ethical

Category	Development	Assessment	Mitigation	Assurance
Data	Consent Protocols, Data (D111)	Consent Assessment, Data (D211)	Consent Verification, Data (D311)	Third-party Review, Data (D411)
	Ethical Source, Data (D112)	Impact Assessment, Data (D212)	Diversity Mitigation, Data (D312)	Ethics Report, Data (D412)
Model	Ethical Guidelines, Model (D121)	Ethical Compliance, Model (D221)	Value Adjustments, Model (D321)	Third-party Review, Model (D421)
	Value Alignment, Model (D122)	Environmental Impact, Model (D222)	Impact Mitigation, Model (D322)	Value Audits, Model (D422)
Code	Ethical Guidelines, Code (D131)	Ethical Compliance, Code (D231)	Mitigation, Code (D331)	Third-party Review, Code (D431)
		Transparency Assessment, Code (D232)	Transparency Mitigation, Code (D332)	Developer Ethical Training, Code (D432)
				Ethics Report, Code (D433)
System	Stakeholder Engagement, System (D141)	Transparency Assessment, System (D241)	Stakeholder Feedback, System (D341)	Third-Party Review, System (D441)
	Transparency, System (D142)		Ethical Mitigation, System (D342)	Transparency Log, System (D442)
				Stakeholder Reviews, System (D443)

- Ensures ethical data sourcing and consent
- Implements value alignment and compliance checks
- Mitigates diversity and environmental impacts
- Promotes transparency and stakeholder engagement

System Card+: Layer - Legal

Category	Development	Assessment	Mitigation	Assurance
Data	License Verification (E111)	Legal Review (E211)	Breach Protocols (E311)	Compliance Certification (E411)
	Regulation Tracking (E112)	Privacy Assessment (E212)	Consent Updates (E312)	Privacy Guarantees (E412)
	Consent Logs (E113)	Consent Audit (E213)	Rectification Processes (E313)	Legal Reporting (E413)
Model	Contract Compliance (E121)	Legal Testing (E221)	Compliance Adjustments (E321)	Legal Clearance (E421)
	Patent Verification (E122)	Liability Review (E222)	Liability Mitigation (E322)	Regulatory Compliance (E422)
	Compliance Documentation (E123)	Regulatory Review (E223)	Legal Remediation (E323)	Audit Trails (E423)
Code	License Audit (E131)	Security Check (E231)	Code Corrections (E331)	License Audits (E431)
	Secrets Management (E132)	Compliance Monitoring (E232)	Compliance Patching (E332)	Ethical Coding (E432)
	Export Compliance (E133)	Documentation Check (E233)	Security Updates (E333)	Document Compliance (E433)
System	Regulatory Submissions (E141)	Compliance Verification (E241)	Compliance Upgrades (E341)	Legal Conformance (E441)
	Privacy Compliance (E142)	Safety Check (E242)	Regulatory Adjustments (E342)	Safety Compliance (E442)
	Accessibility Audit (E143)	Legal Monitoring (E243)	Legal Reassessments (E343)	Accessibility Verification (E443)

- Ensures compliance with legal standards
- Implements privacy and security assessments
- Mitigates legal risks and liabilities
- Promotes ongoing legal monitoring and certification

System Card+ Conclusions

- Proposed a unified, accessible, and straightforward mechanism for compliance and certification.
- The System Card+ methodology offers a considerable step forward in the ethical assessment of AI systems.
- This framework addresses this critical gap by introducing a five-layer holistic benchmark to verify that the principles of performance accountability, fairness, inclusivity, and ethical and legal compliance are embedded within AI systems.
- We recommend a dynamic approach to AI regulation that promotes continuous learning, adaptation, and ethical vigilance.

Ethical Failures in AI/ML

Denied by AI: How Medicare Advantage plans use algorithms to cut off care for seniors in need

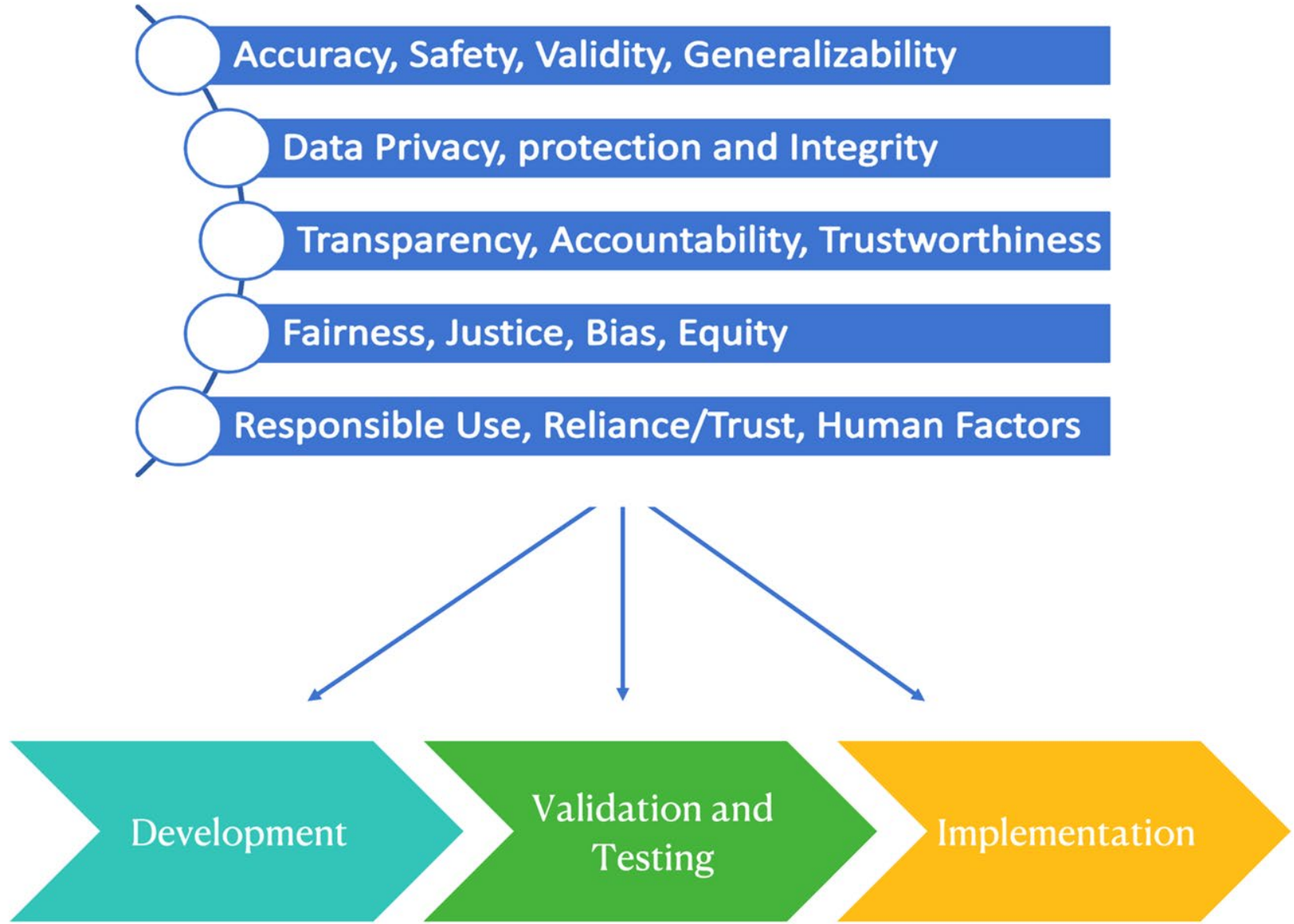
AI Ethics Essentials: Lawsuit Over AI Denial of Healthcare

Examples

Core Principles of Medical Ethics

- Autonomy
- Nonmaleficence
- Beneficence
- Justice

Varkey B. Principles of Clinical Ethics and Their Application to Practice. *Med Princ Pract*. 2021;30(1):17-28.

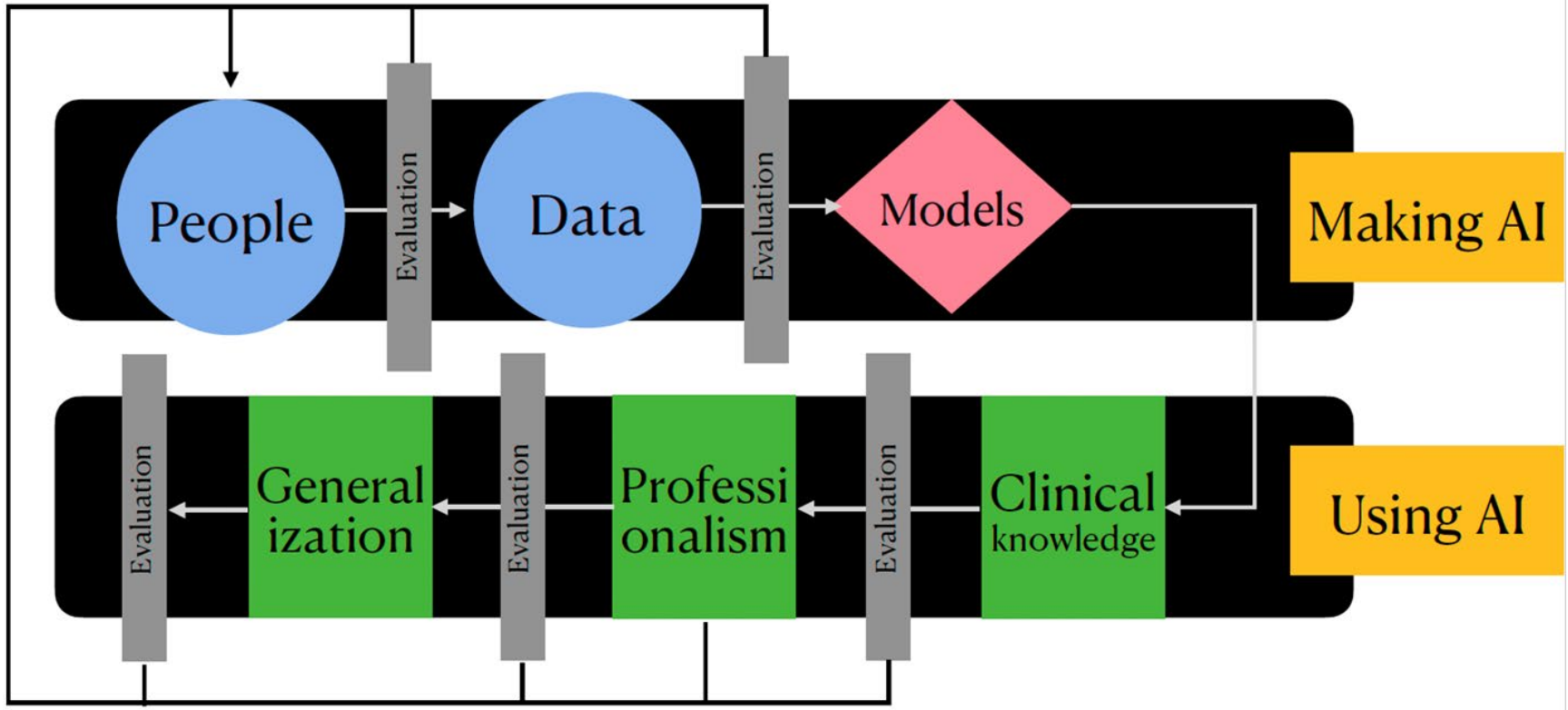


Ethics across the health AI lifecycle

DEVELOPMENT, VALIDATION & EVALUATION

Baylor College of Medicine

CENTER FOR MEDICAL ETHICS & HEALTH POLICY



Slide credit - Brad Malin, Vanderbilt University

teach & transform

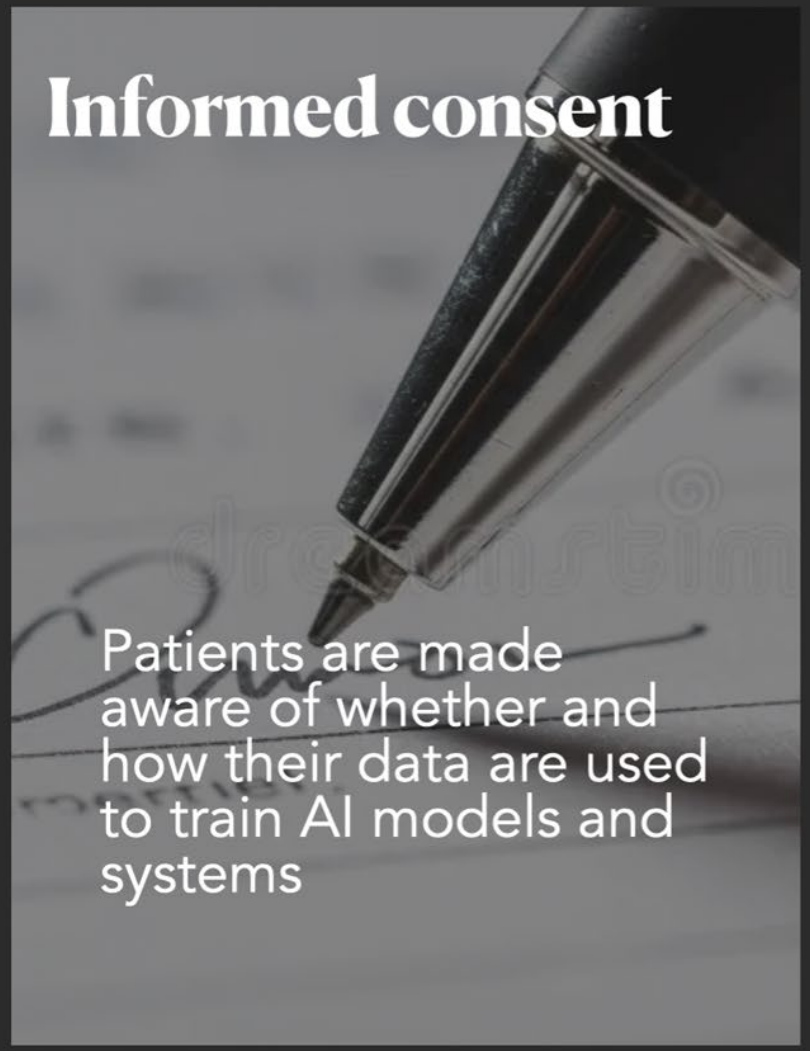
Autonomy & Health AI

Alexander Semenov photography

Autonomy & Health AI

Alexander Semenov photography

Dreamstime



STFM

Frank Gehry - AI art

Limited opportunities to consent

Individuals may not be given the opportunity to consent to the use of AI/ML tools in their care; EHR data used to train AI/ML models may be lawfully used without patient knowledge.

Frank Gehry - AI art

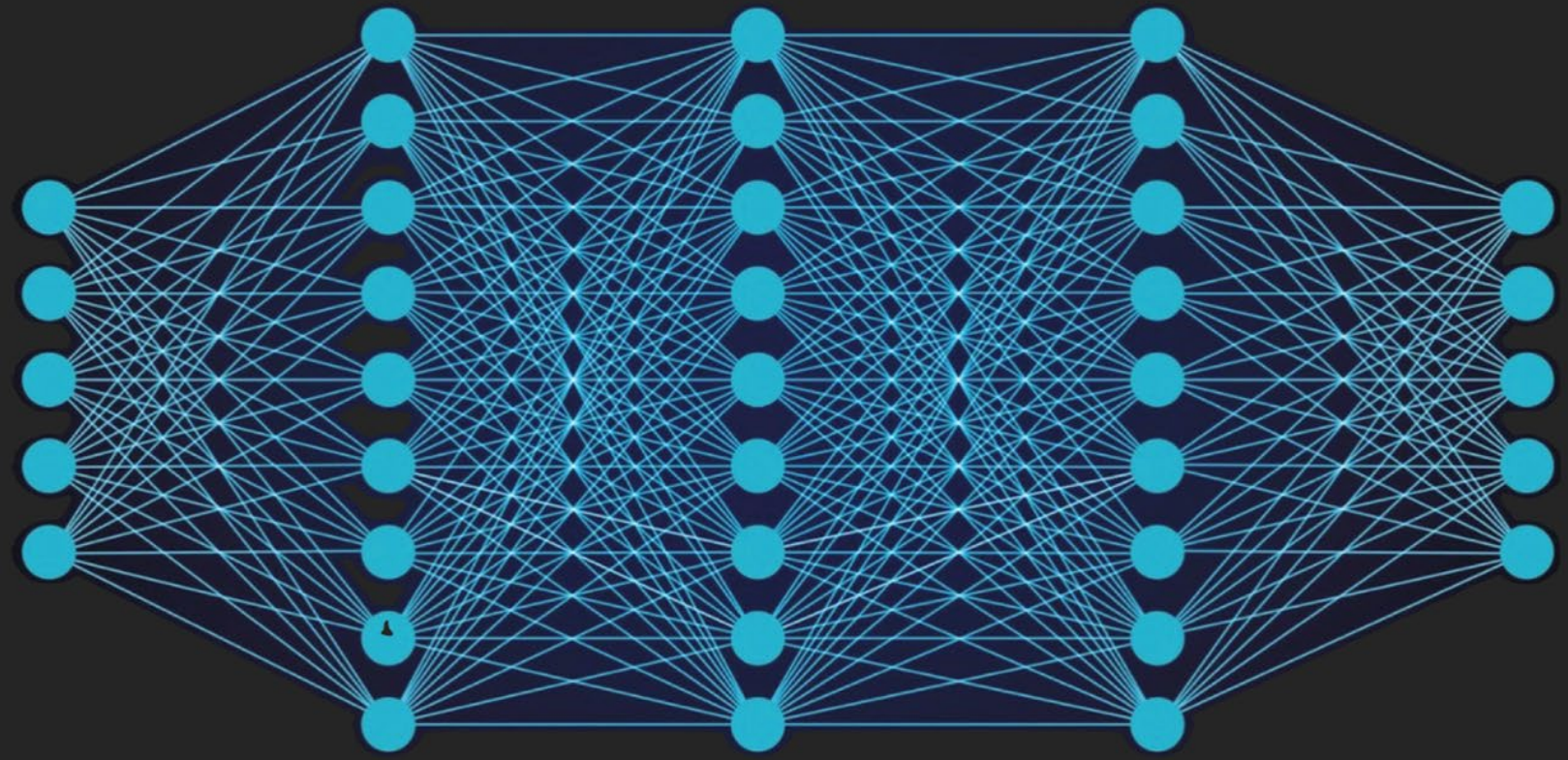
STFM

Threats to confidentiality

AI/ML may lead to the disclosure of patient information to outside parties without the individuals' consent

Ethical Dilemmas in Action

Health AI development & Use

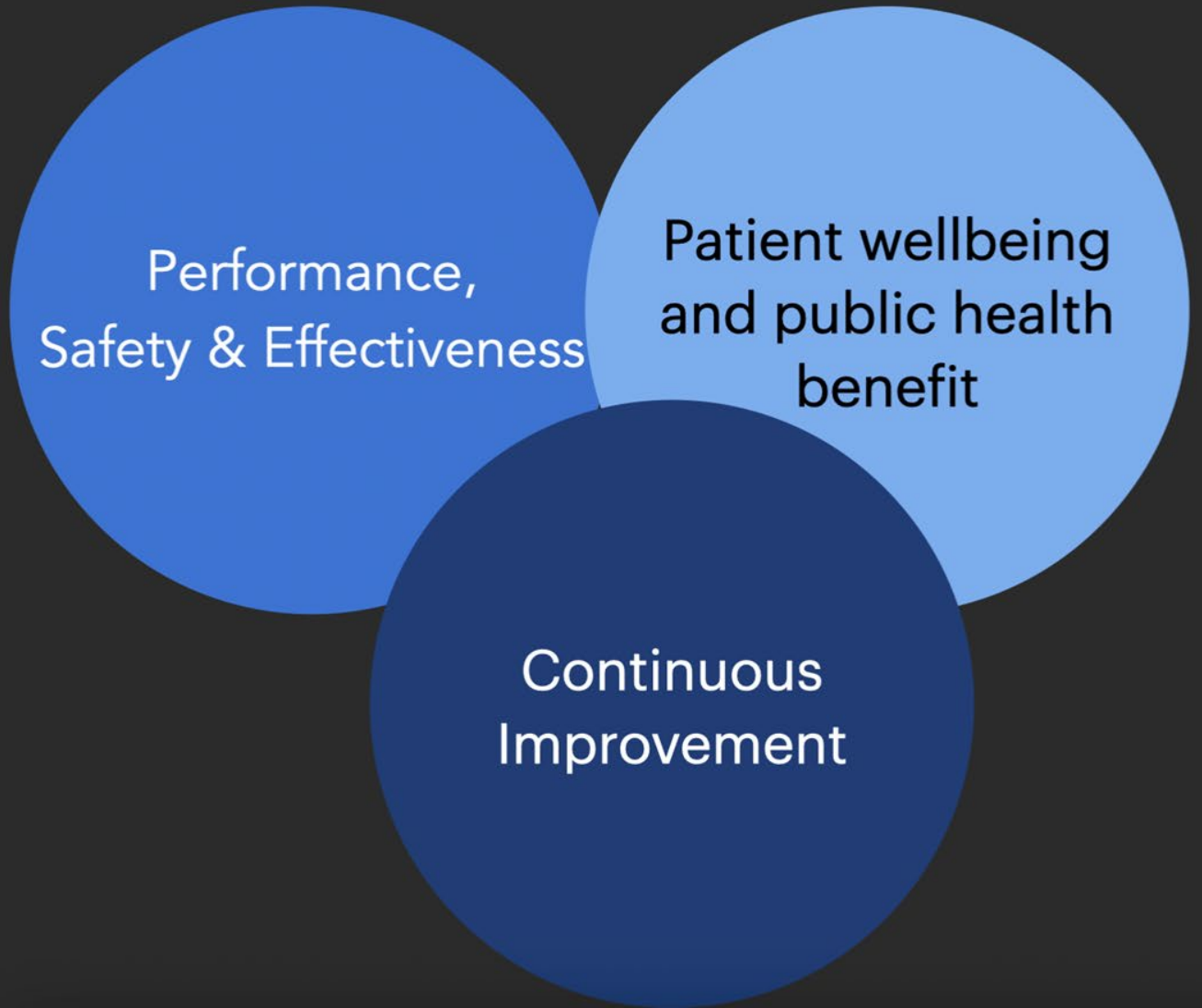


AI-recommended treatments may be too complex to explain clearly, undermining informed consent.

AI Business

Beneficence & Health AI

AI should be used to improve health outcomes and support clinical decision-making.



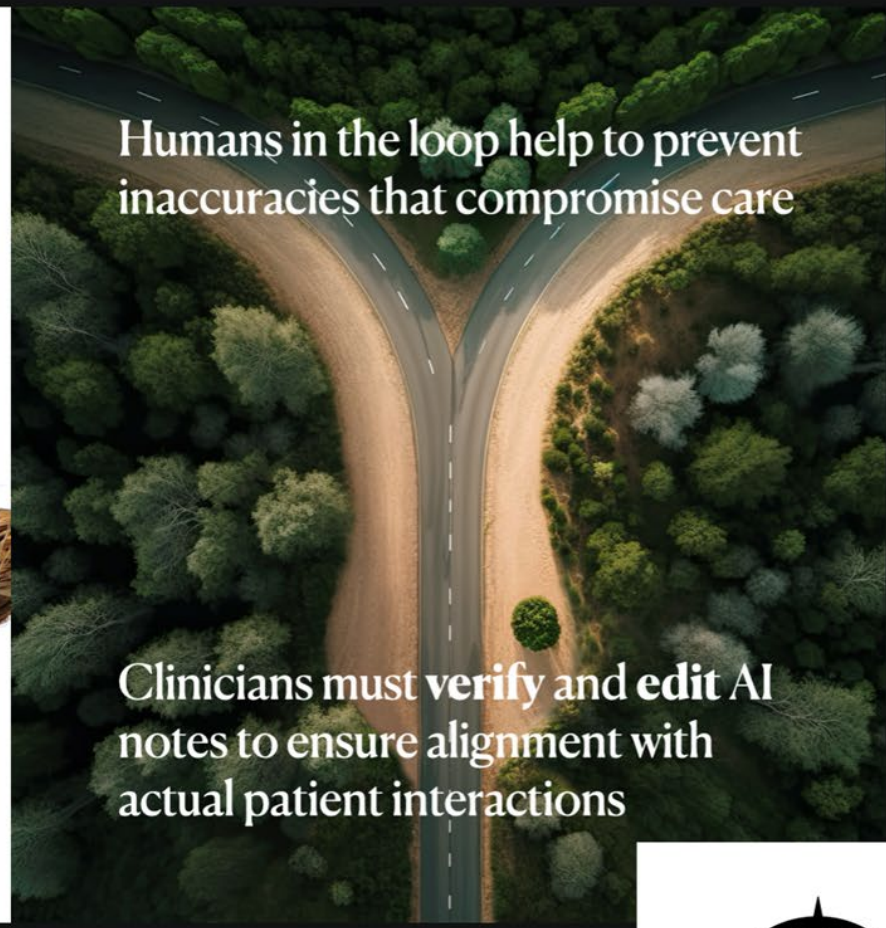
Ethical dilemma

Inaccurate AI documentation can lead to improper care, misunderstanding, or inappropriate treatments.

Example of images generated by CAN (Creative Adversarial Networks). The generated images vary from simple abstract ones to complex textures and compositions (Elgammal et a. 2017).

Erroneous EHR entries may propagate mistakes, especially in follow-up care or by other providers.

Clinical decisions may be made on false assumptions or flawed information

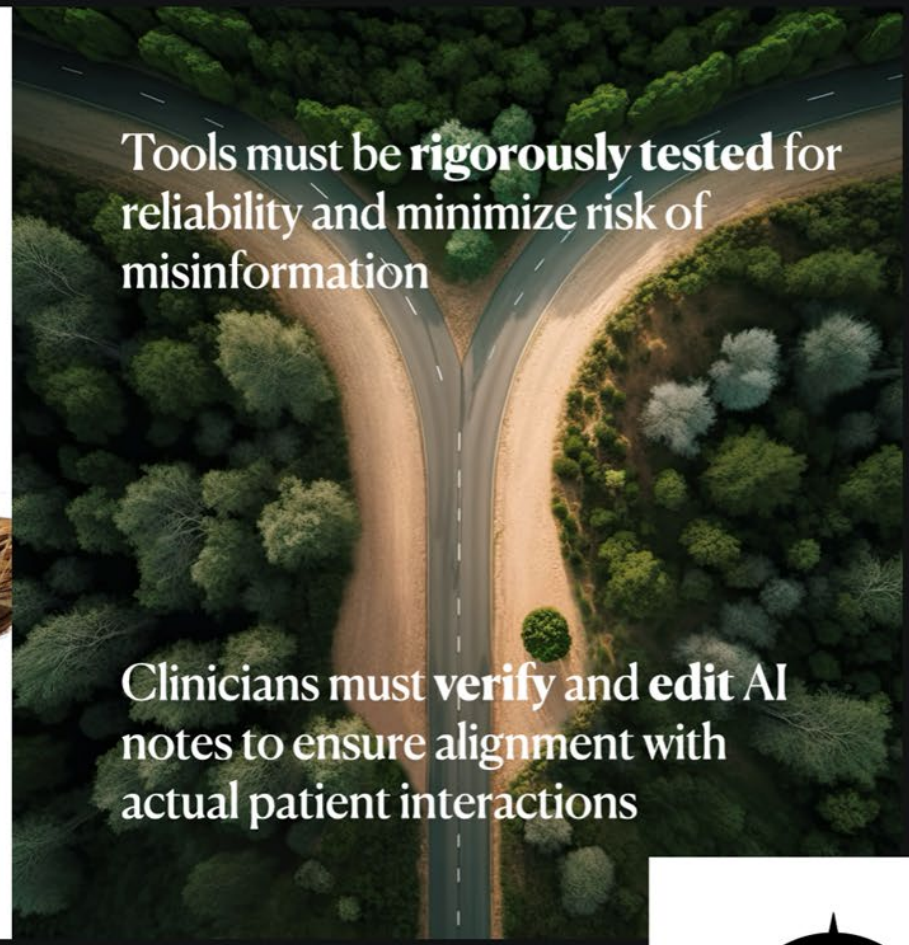


Nonmaleficence & Health AI

Preventing harm through safe, validated, and unbiased AI systems.

Misinterpretations by the AI can lead to **harmful clinical decisions.**

Undetected errors may affect future **diagnosis, treatment,** or **legal liability**



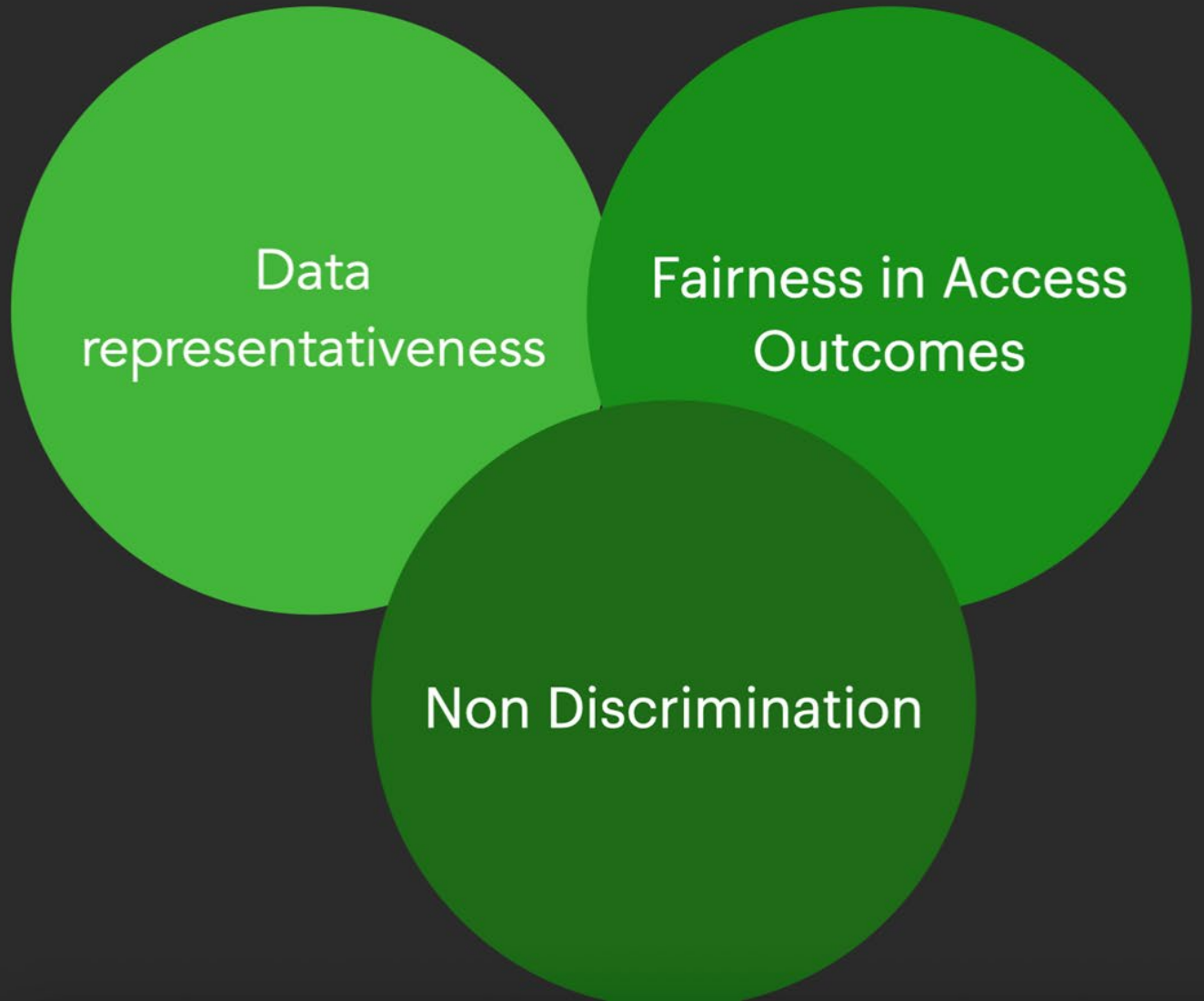
Tools must be **rigorously tested** for reliability and minimize risk of misinformation

Clinicians must **verify** and **edit** AI notes to ensure alignment with actual patient interactions

teach & transform

Justice & Health AI

Ensuring equitable access to non discriminating technology for all.



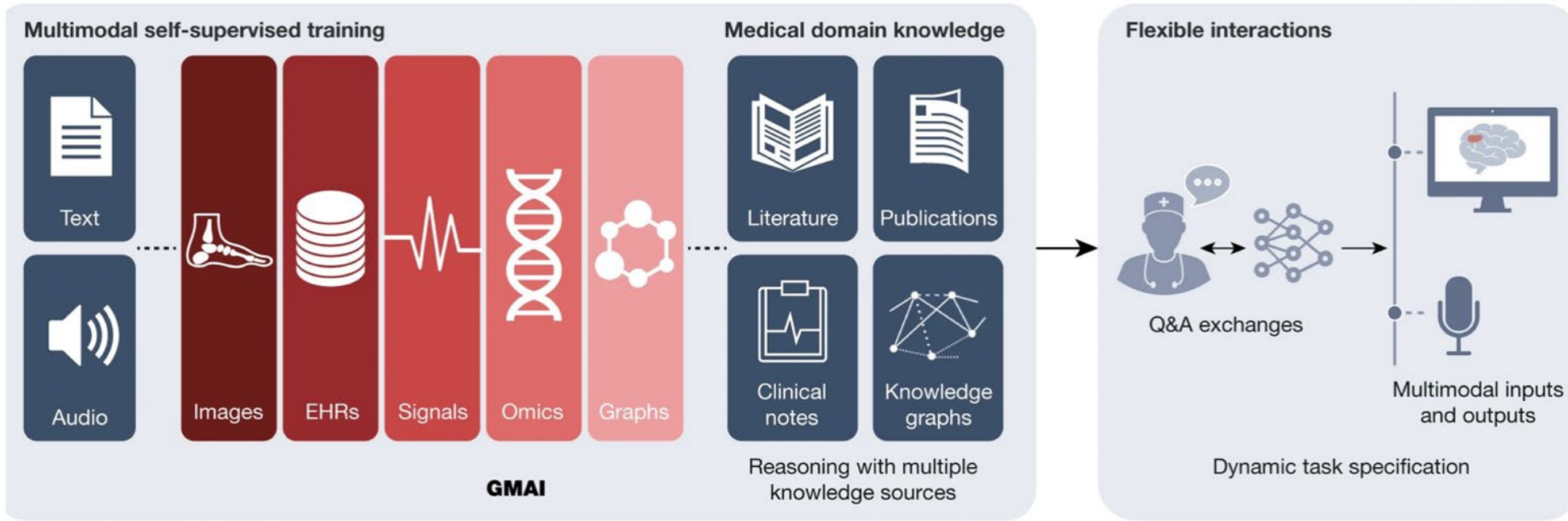
Misinterpretations AI tools developed for well-resourced settings may not function in low-income clinics.

Resource constraints might lead to AI being used more in underserved communities, raising quality concerns.

Practice data stewardship, equity in access, and ethical sourcing of patient data

Promote **public interest oversight** to ensure technology use aligns with healthcare values, not just commercial gains

Concept map of generalist medical AI & its applications



Moor, M., Banerjee, O., Abad, Z.S.H. *et al.* Foundation models for generalist medical artificial intelligence. *Nature* **616**, 259–265 (2023). https://doi.org/10.1038/s41586-023-05881-4

Ethical Conflicts

In using health AI technology in primary care

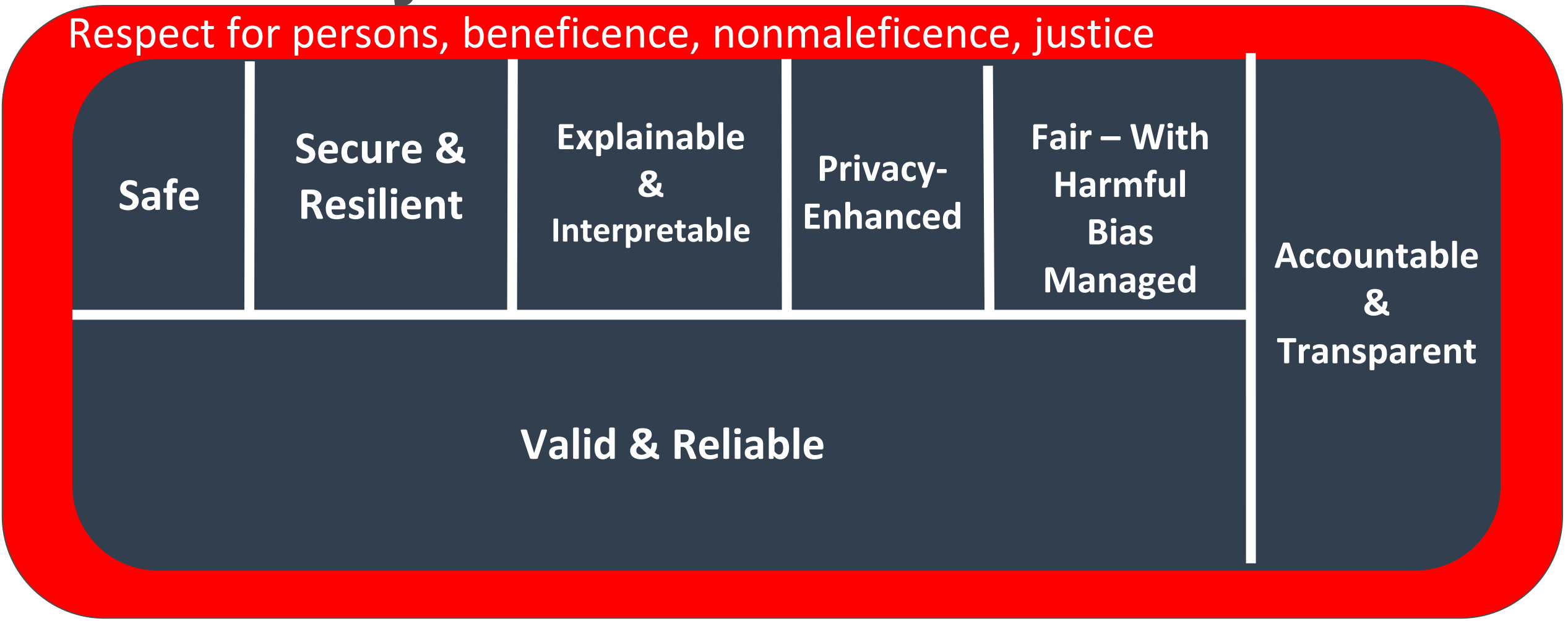
Autonomy vs. Beneficence Patient chooses a treatment that AI predicts will be ineffective.

Non-Maleficence vs. Justice: Avoiding risky AI may limit its use in underserved areas, exacerbating inequities.

Beneficence vs. Justice: High-performance AI tools are only accessible in private clinics.

Autonomy vs. Non-Maleficence: Full disclosure of AI uncertainties may overwhelm or distress patients.

Trustworthy AI



Characteristics of trustworthy AI/ML tools: Valid & Reliable is a necessary condition of trustworthiness and is shown as the base for the other characteristics. Accountable & Transparent is shown as a vertical box because it relates to all other characteristics.

Privacy

- In 2023 alone, there were 725 hacks that impacted 133 million (or 1 in 3) Americans.
- A study in 2013 showed that four anonymized data (documenting an individual's location at specific times) points are sufficient for unique identification of an individual with 95% accuracy.
- Identity theft: To commit crimes, obtain credit / goods / services, assume another person's identity in daily life, or procure medical care

The Environmental Impact of AI/ML

- Emissions associated with training one large language model (over 84 hours) were equal to 626,155 pounds of carbon dioxide or 125 round-trip flights between New York City and Beijing.
- ChatGPT uses 500 milliliters of water (approximately 16-ounce water bottle) every time you ask it a series of between 5 to 50 prompts or questions. The range varies depending on where its servers are located and the season.

Artificial intelligence technology behind ChatGPT was built in Iowa — with a lot of water

Automation Complacency and Automation Bias

Photo by Michael Berdyugin (Pexels)

Hundreds of AI tools have been built to catch covid. None of them helped.

"It's shocking," says Wynants. "I went into it with some worries, but this exceeded my fears."

Epic's widely used sepsis prediction model falls short among Michigan Medicine patients

Calculation of the Net Benefit

- The benefits of AI/ML may be muted in real-world settings
- Consider an AI/ML tool that predicts risk for decompensation in the hospital.
- When admitted, those at highest risk are proactively transferred to the Intensive Care Unit (ICU) in order to reduce mortality and length of stay.
- When the tool is used for him, Mr. G is a true positive, meaning that the tool accurately predicts that he will decompensate during this hospitalization.
- Unfortunately, no ICU beds are available. Thus, the benefit of using the tool is not ultimately realized.

Assessing the net benefit of machine learning models in the presence of resource constraints

Bias Amplification

Millions of black people affected by racial bias in health-care algorithms

Dissecting racial bias in an algorithm used to manage the health of populations

Bias in Data

AI skin cancer diagnoses risk being less accurate for dark skin – study

Characteristics of publicly available skin cancer image datasets: a systematic review

The Use of Race in Prediction Tools

Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms

New Creatinine- and Cystatin C–Based Equations to Estimate GFR without Race

- Is the need for race correction based on robust evidence and statistical analyses (e.g., with consideration of internal and external validity, potential confounders, and bias)?
- Is there a plausible causal mechanism for the racial difference that justifies the race correction?
- Would implementing this race correction relieve or exacerbate health inequities?

Race adjustments in clinical algorithms can help correct for racial disparities in data quality

Digital Divide

The Emerging AI Divide in the United States

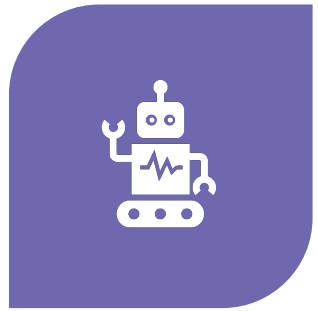
What Family Medicine Can Do

What is Explainable AI and Why is it Important?

TRUST AND TRANSPARENCY

REGULATORY COMPLIANCE: legal frameworks like the european union's general data protection regulation (GDPR) emphasize the "right to explanation," ensuring individuals can understand decisions made by automated systems.

BIAS DETECTION AND MITIGATION: XAI facilitates the identification of biases within ai models.

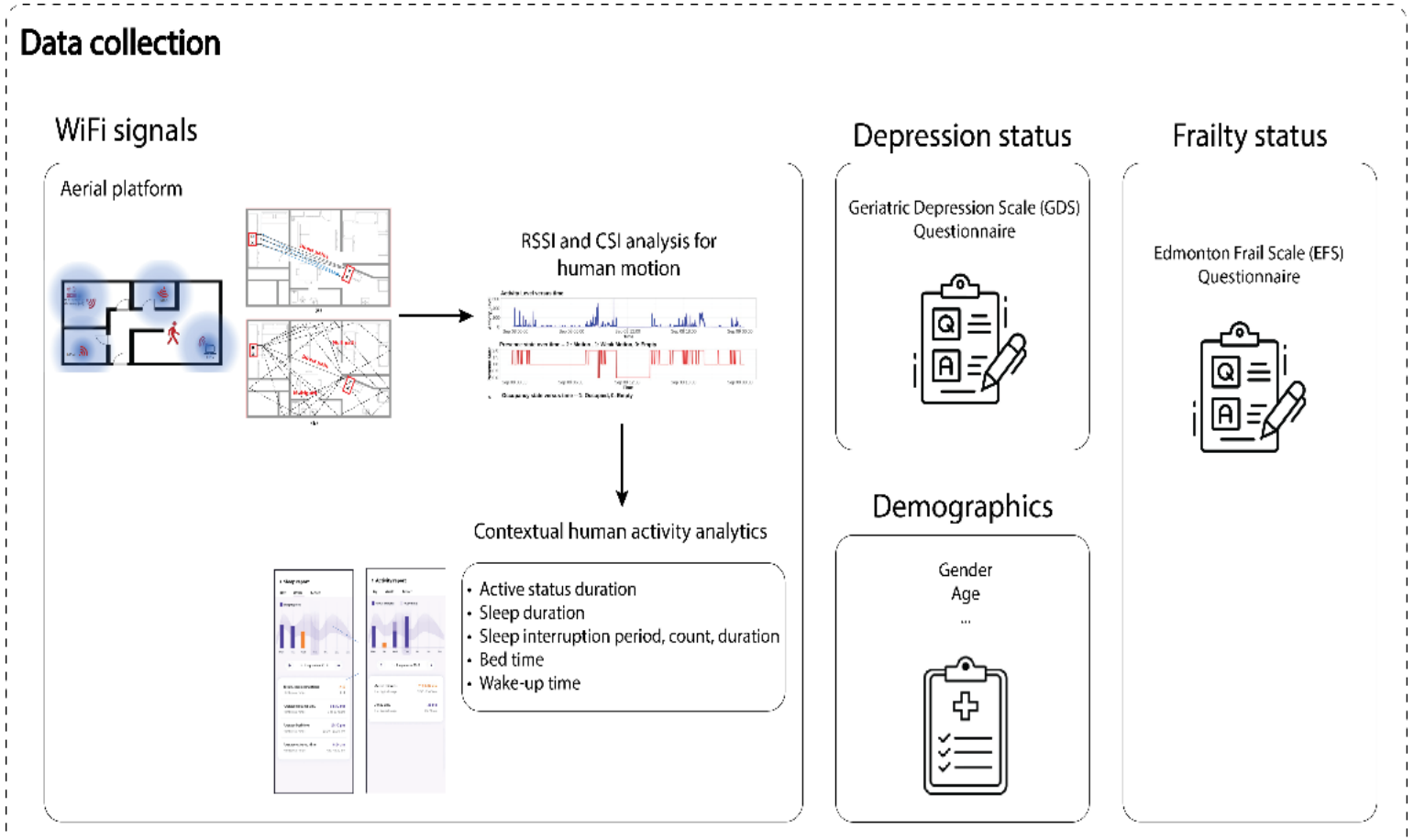


ENHANCED HUMAN-AI COLLABORATION: XAI enables users to better understand ai recommendations.

RAHIMI´S LAB

HOPE: Early detection of Depression among older adults

- **Aim**: Develop a new ML model for monitoring physical activity and classifying depression (HOPE model, published at JMIR aging)and frailty (AID model, under revision at JAMIA) in older adults
- **Sample**: Older adults (65+) in Montreal
- **Data collection**: Contextual human activity and sleep data collected over **six months** using a non-intrusive, WiFi-based motion sensors.
- **Data analysis**:
  1. New ML models (i.e., HOPE and AID) with a three-stage process.
  2. XML techniques to interpret model predictions



RAHIMI'S LAB

McGill

Économie et Innovation Québec

Mila

Aerial WIFI MOTION ANALYTICS

teach & transform

HOPE Model

JMIR Aging | Journal Information | Browse Journal

Published on 03.03.2025 in Vol 8 (2025)

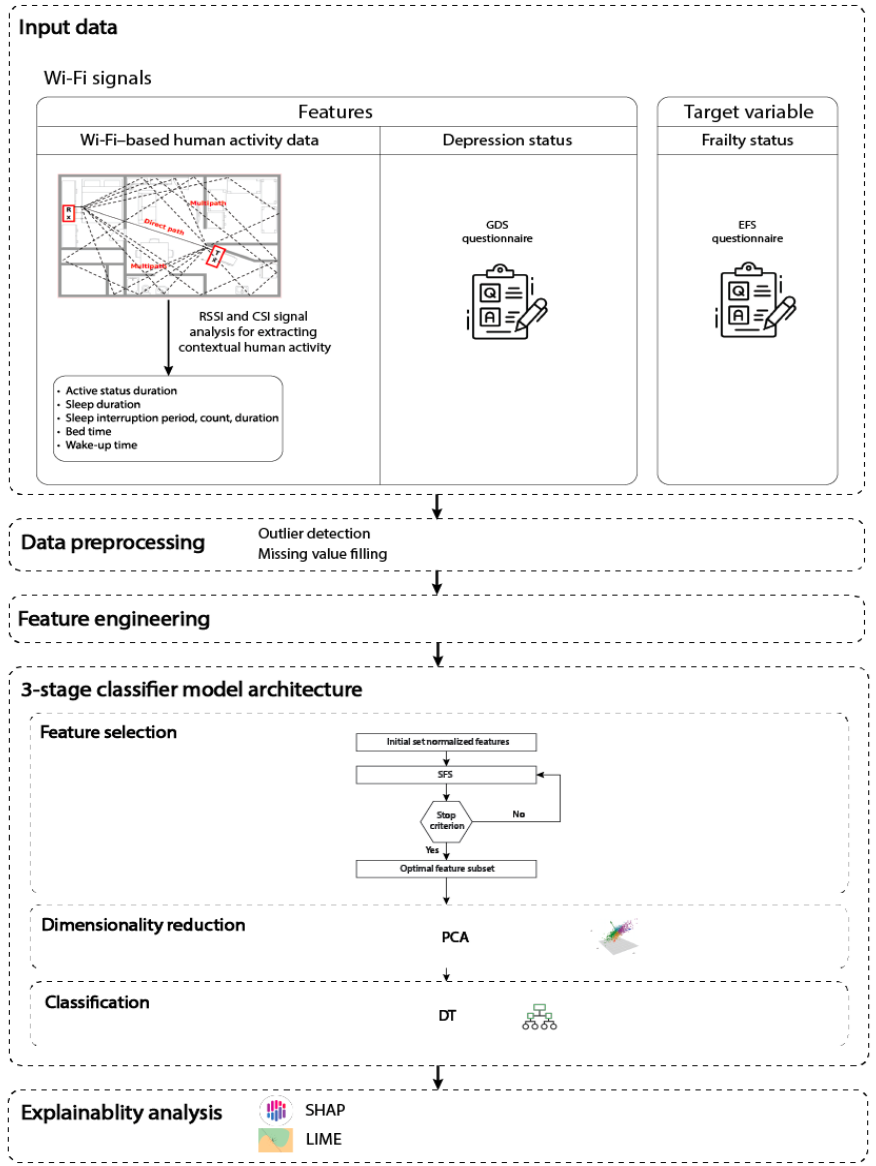
Preprints (earlier versions) of this paper are available at https://preprints.jmir.org/preprint/67715, first published October 18, 2024.

Development and Feasibility Study of HOPE Model for Prediction of Depression Among Older Adults Using Wi-Fi-based Motion Sensor Data: Machine Learning Study

Shayan Nejadshamsi[1,2,3]; Vania Karami[1,2,3]; Negar Ghourchian[4]; Narges Armanfard[1,5]; Howard Bergman[2]; Roland Grad[2]; Machelle Wilchesky[2,3,6]; Vladimir Khanassov[2]; Isabelle Vedel[2]; Samira Abbasgholizadeh Rahimi[1,2,3,7]

Table 4. Average performance across different baseline machine learning models.

Model architecture	Accuracy (%)	Sensitivity (%)	Precision (%)	F_1-score (%)
RF[a] [23,27]	12.50	5.00	5.00	N/A[b]
SVM[c] [23]	15.00	10.00	10.00	N/A
LR[d] [23]	22.50	15.00	13.34	N/A
XGBoost [34]	25.00	10.00	10.00	N/A
L1-based feature selection + DT[e] [24]	32.50	35.00	18.34	N/A
L1-based feature selection + RF [24]	22.50	15.00	13.34	N/A
L1-based feature selection + kNN[f] [24]	22.50	15.00	13.34	N/A
L1-based feature selection + NB[g] [24]	30.00	25.00	15.00	N/A
L1-based feature selection + LR [24]	37.50	45.00	25.00	N/A
L1-based feature selection + SVM [24]	25.00	10.00	10.00	N/A
Randomized LR + AdaBoost [25]	55.00	72.22	55.00	N/A
HOPE model[h]	87.50	90.00	88.34	86.00

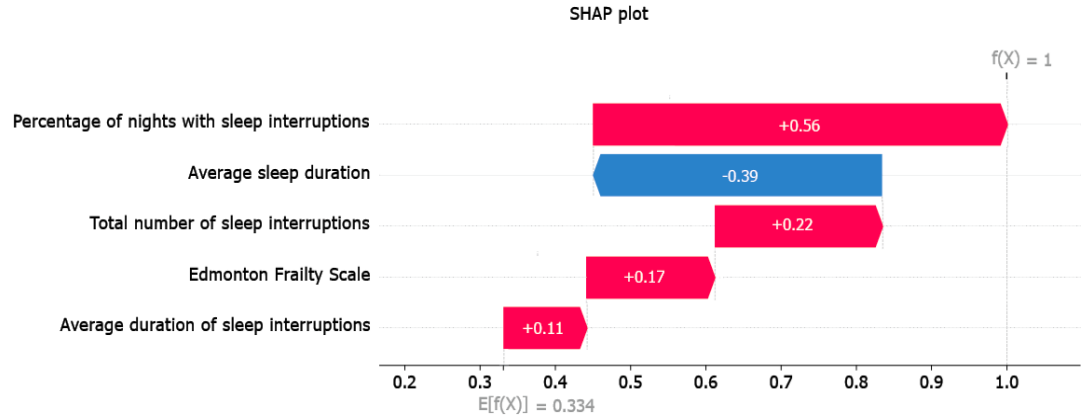


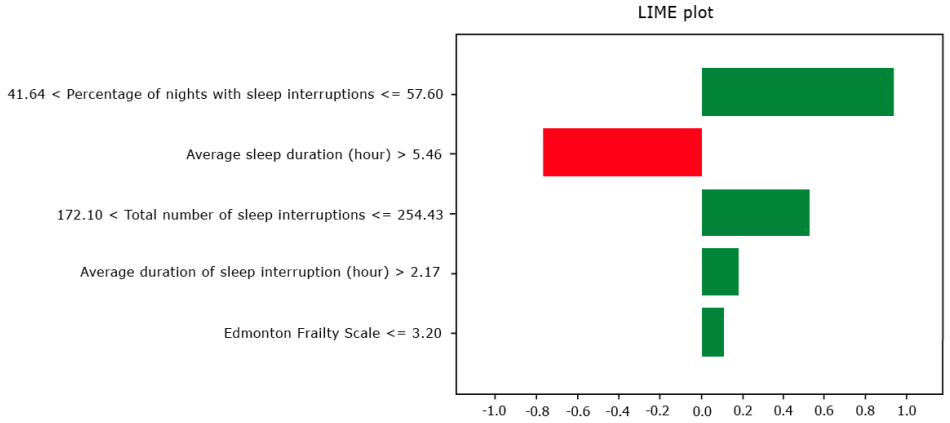
RAHIMI'S LAB

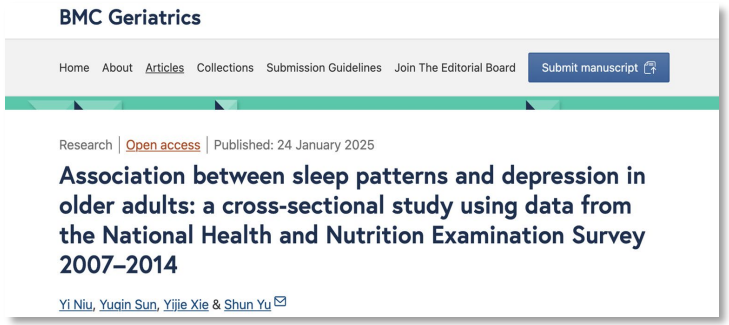
HOPE Model

Results:

- **Advantages**:
  - Non-intrusive
  - Enables continuous monitoring
- **Data features:**
  - Uses macro-level features instead of micro-level features
- **Explainability:**
  - Applied XML to identify key frailty/depression related features







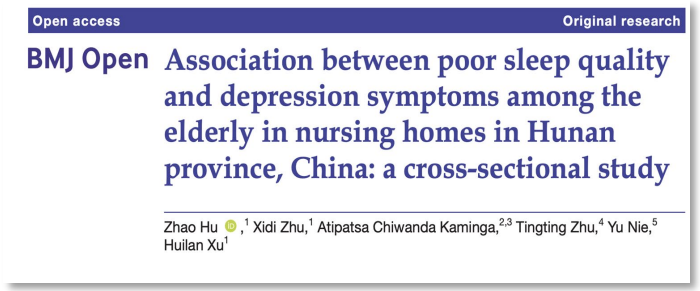
BMC Geriatrics

Home About Articles Collections Submission Guidelines Join The Editorial Board Submit manuscript

Research | Open access | Published: 24 January 2025

Association between sleep patterns and depression in older adults: a cross-sectional study using data from the National Health and Nutrition Examination Survey 2007–2014

Yi Niu, Yuqin Sun, Yijie Xie & Shun Yu



Open access Original research

BMJ Open

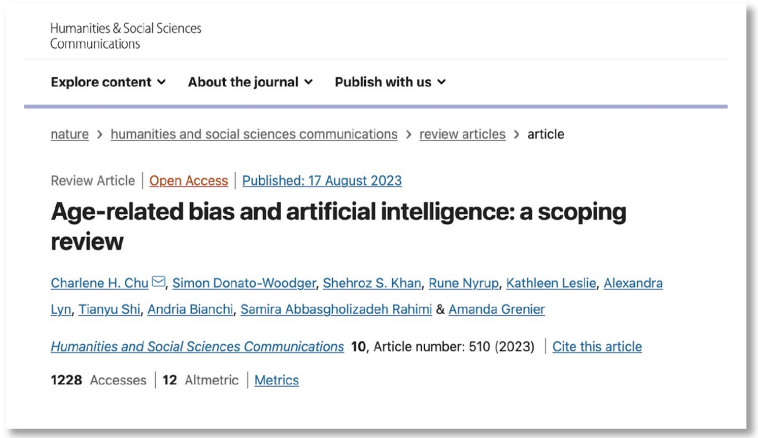
Association between poor sleep quality and depression symptoms among the elderly in nursing homes in Hunan province, China: a cross-sectional study

Zhao Hu, Xidi Zhu, Atipatsa Chiwanda Kaminga, Tingting Zhu, Yu Nie, Huilan Xu

RAHIMI'S LAB

Responsible & safe AI in health

- 2.1) XML
- 2.2) Bias in AI (age-related)
- 2.3) EDAI



CIHR IRSC
Canadian Institutes of Health Research | Instituts de recherche en santé du Canada

Fonds de recherche Santé Québec

RAHIMI'S LAB

AI & education

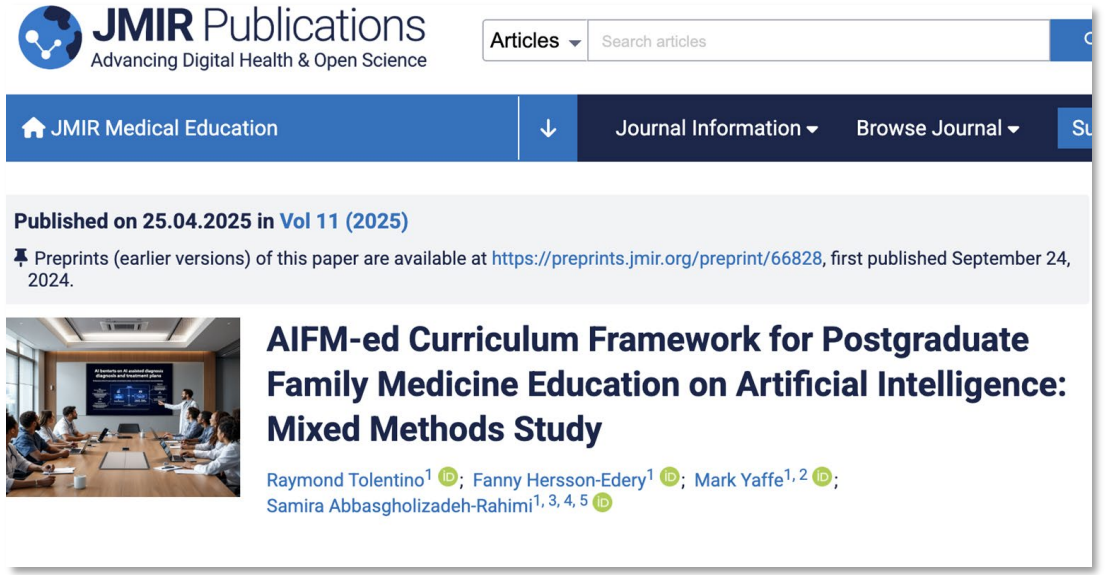
- The AIFM-ed curriculum farmwork is a structured guide developed to integrate AI competencies into postgraduate family medicine training.
- It aims to equip future family physicians with the necessary skills to utilize AI effectively in clinical practice.
- The framework emphasizes the importance of AI literacy, ethical considerations, and practical application in patient care.


Open access Original research

Family Medicine and Community Health

Performance of generative pre-trained transformers (GPTs) in Certification Examination of the College of Family Physicians of Canada

Mehdi Mousavi [1], Shabnam Shafiee,[2] Jason M Harley,[3,4,5] Jackie Chi Kit Cheung,[6,7] Samira Abbasgholizadeh Rahimi [8,9,10,11]


JMIR Publications
Advancing Digital Health & Open Science

Articles | Search articles

JMIR Medical Education | Journal Information | Browse Journal

Published on 25.04.2025 in Vol 11 (2025)

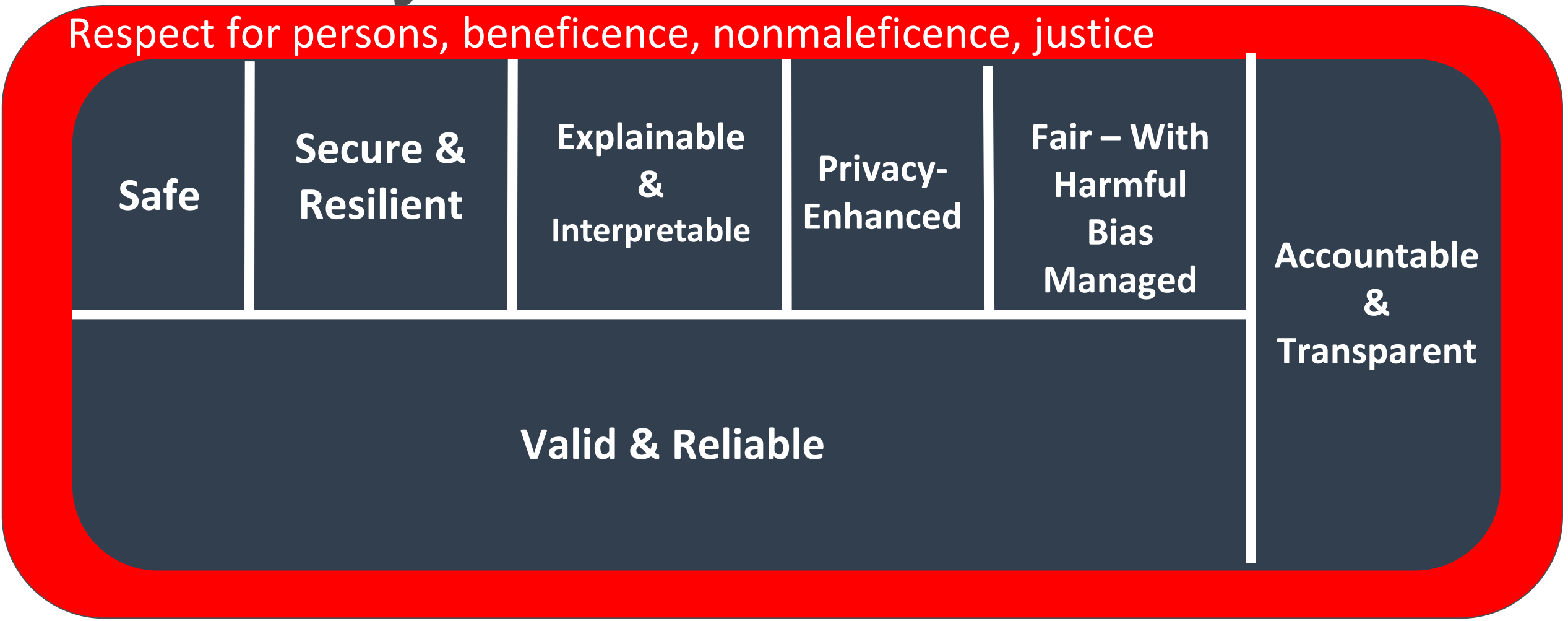
Preprints (earlier versions) of this paper are available at https://preprints.jmir.org/preprint/66828, first published September 24, 2024.

AIFM-ed Curriculum Framework for Postgraduate Family Medicine Education on Artificial Intelligence: Mixed Methods Study

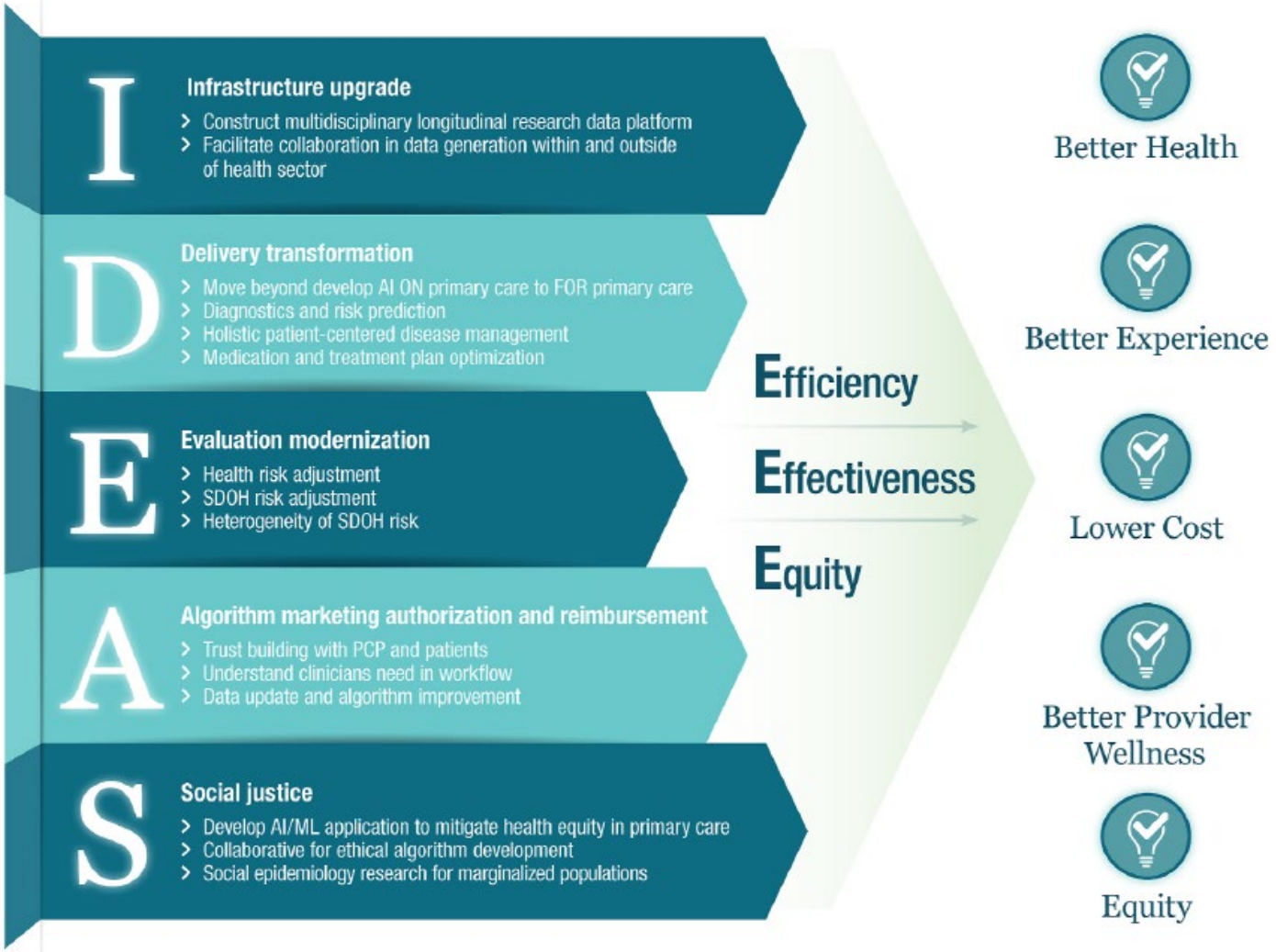
Raymond Tolentino[1]; Fanny Hersson-Edery[1]; Mark Yaffe[1,2]; Samira Abbasgholizadeh-Rahimi[1,3,4,5]

RAHIMI'S LAB

Trustworthy AI



Characteristics of trustworthy AI/ML tools: Valid & Reliable is a necessary condition of trustworthiness and is shown as the base for the other characteristics. Accountable & Transparent is shown as a vertical box because it relates to all other characteristics.



Advancing primary care with Artificial Intelligence and Machine Learning

WHO: An Ethical Approach to Use of AI/ML for Health

Ethical and transparent design of technologies

Recommendations (paraphrased)

- Stakeholders should be engaged early
- Designers should empower stakeholders
- Design values should be informed and updated by according the latest practice standards
- Continuing education and training programs should be available to designers and developers

WHO: An Ethical Approach to Use of AI/ML for Health

Demonstrate trustworthiness of AI systems to clinicians and patients

- **Recommendations (paraphrased)**
  - The public should be educated on forms of data sharing, be able to comment on socially and culturally acceptable AI, and express their concerns and expectations.
  - Training and continuing education programs should be available to assist health care professionals in understanding and adapting to use of AI/ML, learning about its benefits and risks and understanding the ethical issues raised in their use.

WHO: An Ethical Approach to Use of AI/ML for Health

Conduct ethical and social impact assessments

- **Recommendations (paraphrased)**
  - Governments should enact laws and policies that require agencies and companies to conduct impact assessments of AI/ML technologies, which should address ethics, human rights, safety, and data protection, throughout the life-cycle of an AI/ML system.

Article 27: Fundamental Rights Impact Assessment for High-Risk AI Systems

Date of entry into force: 2 August 2026

According to: Article 113

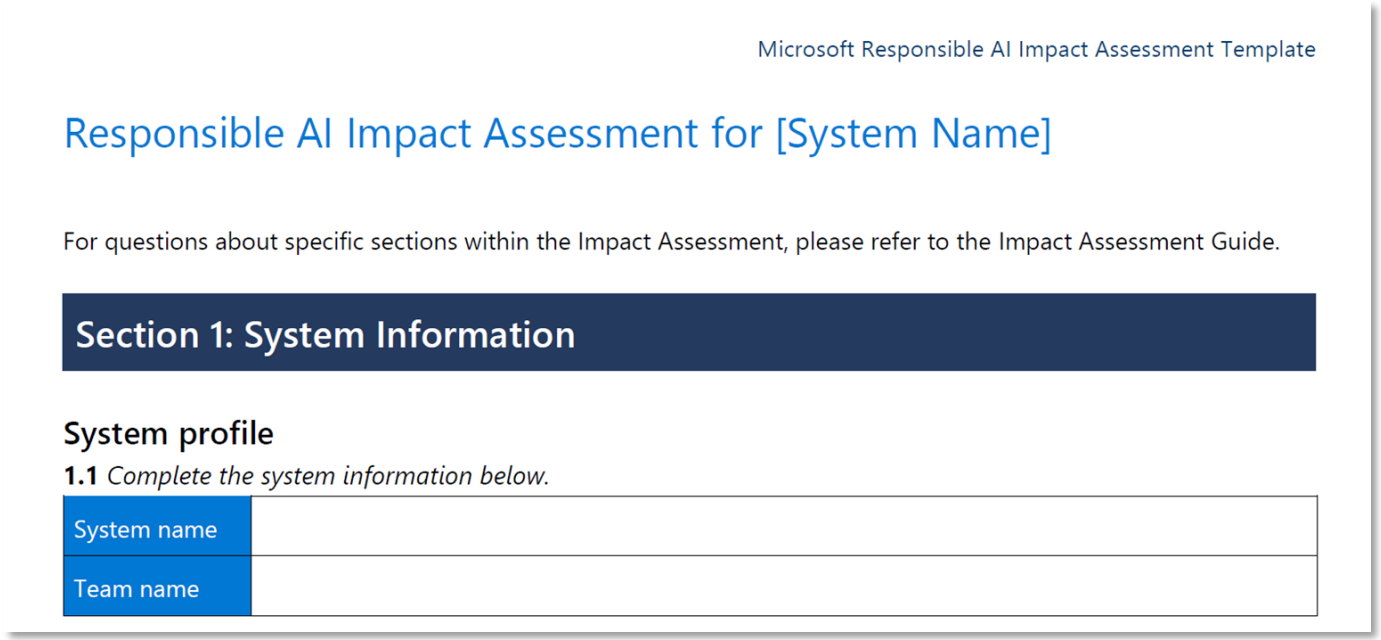
See here for a full implementation timeline.

SUMMARY +

1. Prior to deploying a high-risk AI system referred to in Article 6(2), with the exception of high-risk AI systems intended to be used in the area listed in point 2 of Annex III, deployers that are bodies governed by public law, or are private entities providing public services, and deployers of high-risk AI systems referred to in points 5 (b) and (c) of Annex III, shall perform an assessment of the impact on fundamental rights that the use of such system may produce. For that purpose, deployers shall perform an assessment consisting of:

WHO: An Ethical Approach to Use of AI/ML for Health

- Impact assessment, sample template



Microsoft Responsible AI Impact Assessment Template

Responsible AI Impact Assessment for [System Name]

For questions about specific sections within the Impact Assessment, please refer to the Impact Assessment Guide.

Section 1: System Information

System profile

1.1 *Complete the system information below.*

System name	
Team name	

WHO: An Ethical Approach to Use of AI/ML for Health

Develop a research agenda for the ethical use of AI/ML for health care

- **Pertinent research questions:**
  - For what gaps identified by health care workers and patients could AI/ML play a role in ensuring the delivery of equitable care?
  - How is AI/ML changing the relationships between health care workers and patients? Do these technologies allow clinicians to spend more time with patients, or do they make care less humane? Do specific contextual factors improve or undermine the quality of care?
  - What are the attitudes of health care workers and patients towards the use of AI/ML?
  - Do they find these technologies acceptable? Do their attitudes depend on the type of intervention, the location of the intervention, or current acceptance of these technologies?
  - Has the introduction and use of AI/ML for health exacerbated the digital divide? Or does AI/ML reduce the gap in access to care and ensure equitable access to high-quality care?
  - How can clinicians and programmers best address biases? What are the barriers to addressing biases?
  - How should governments and clinicians assess fair resource allocation for existing interventions and new technologies?
  - Can ethical design be applied specifically to AI/ML technologies for health?

Artificial Intelligence in Health, Health Care, and Biomedical Sciences

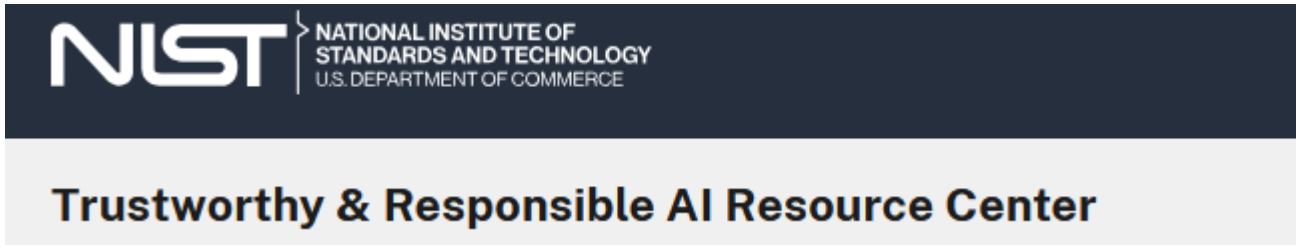
Artificial Intelligence in Health, Health Care, and Biomedical Science: An AI Code of Conduct Principles and Commitments Discussion Draft

- **Focus:** Protect and advance human health and human connection as the primary aims
- **Benefits:** Ensure equitable distribution and risk for all
- **Involvement:** Engage people as partners with agency in every stage of the life cycle
- **Workforce well-being:** Renew the moral well-being and sense of shared purpose to the health care workforce
- **Monitoring:** Monitor and openly and comprehensively share methods and evidence of AI/ML's performance and impact on health and safety
- **Innovation:** Innovate, adopt, collaboratively learn, continuously improve, and advance the standard of clinical practice
- The goal is that all decisions associated with, and actions taken, to develop and deploy AI/ML in the health sector will be consistent with these Commitments to develop and foster trust

Collective Responsibility

- AI/ML governance is a system of rules, practices, processes, and technological tools that are employed to ensure an organization's use of AI/ML technologies aligns with the organization's strategies, objectives, and values; fulfills legal requirements; and meets principles of ethical AI/ML followed by the organization

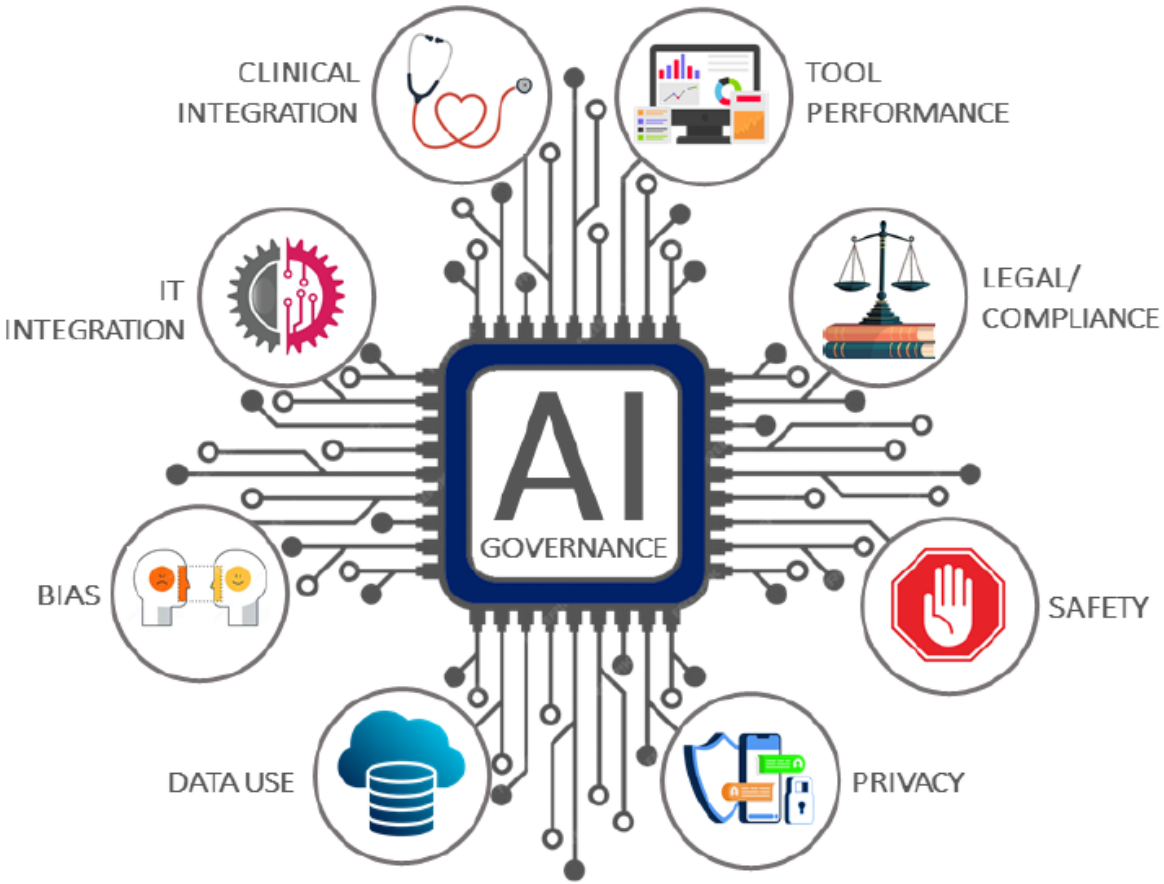
Governance



- Policies, processes, procedures and practices across the organization related to the mapping, measuring and managing of AI/ML risks are in place, transparent, and implemented effectively
- Accountability structures are in place so that the appropriate teams and individuals are empowered, responsible, and trained for mapping, measuring, and managing AI risks
- Workforce diversity, equity, inclusion, and accessibility processes are prioritized in the mapping, measuring, and managing of AI/ML risks throughout the lifecycle
- Organizational teams are committed to a culture that considers and communicates AI/ML risk
- Processes are in place for robust engagement with relevant AI actors
- Policies and procedures are in place to address AI/ML risks and benefits arising from third-party software and data and other supply chain issues

AI Governance in Health Systems

Aligning Innovation, Accountability, and Trust



Take Home Messages

- Development and use of AI/ML have ethical consequences for patients and populations
- AI/ML can challenge core ethical principles of respect for persons, beneficence, nonmaleficence and justice
- Like any technology, AI/ML has the potential to worsen health disparities if ethical principles are not appropriately considered during AI/ML development and use

  Practices such as the equitable distribution of AI/ML, transparent architecture, proactive impact assessment, and participatory design can support the development of trustworthy AI/ML.
- If you use AI/ML in your practice, it's essential to establish the appropriate policies, processes, procedures, and practices to ensure effective governance.

Poll Question

After attending this webinar, how has your view of ethical AI in family medicine changed?

- I feel more confident about the topic
- I’m more aware of ethical challenges
- I still have concerns/questions
- My view hasn’t changed much
- I plan to explore this further