THE COALITION FOR HEALTH AI

BLUEPRINT FOR TRUSTWORTHY AI Webinar

April 5, 2023 12-1pm ET

CHAI

Coalition for Health AI

Agenda | April 5 , 2023

MEETING HOSTED BY:	The Coalition for Health AI *DRAFT Blueprint* https://www.coalitionforhealthai.org/papers/blueprint-for-trustworthy-ai_V1.0.pdf
12PM	**Welcome | Opening Remarks** **What is CHAI?** 1. - Brian Anderson
12:05PM	**Blueprint for an AI Bill of Rights:** *Making Automated Systems Work for the American People* - Alan E. Mislove, EOP/OSTP
12:15PM	**NAM Code of Conduct and Collaboration with CHAI** - Laura Adams, National Academy of Medicine
12:25PM	**Introduction to the DRAFT Blueprint** - Gil Alterovitz – Overview - Nicoleta Economou – Mapping NIST RMF to CHAI Assurance Standards - John Halamka – The Need of Health AI Assurance Labs - Peter Embi – Need for Maturity Framework
12:45PM	**Taking AI into the Real-World of Healthcare** - Eric Horvitz, Microsoft
12:55PM	**CLOSING REMARKS** - Michael Pencina

CHAI
Coalition for Health AI

What is CHAI?

- Brian Anderson, MITRE

CHAI

Coalition for Health AI

- Alan E. Mislove, EOP/OSTP

THE WHITE HOUSE
WASHINGTON

BLUEPRINT FOR AN AI BILL OF RIGHTS
Making Automated Systems Work for the American People

Alan Mislove
Assistant Director for Data and Democracy
White House Office of Science and Technology Policy

Blueprint for an AI Bill of Rights

Safe and Effective Systems

You should be protected from unsafe or ineffective systems.

Blueprint for an AI Bill of Rights

Safe and Effective Systems

You should be protected from unsafe or ineffective systems.

Algorithmic Discrimination Protections

You should not face discrimination by algorithms and systems should be used and designed in an equitable way.

Blueprint for an AI Bill of Rights

Safe and Effective Systems

You should be protected from unsafe or ineffective systems.

Algorithmic Discrimination Protections

You should not face discrimination by algorithms and systems should be used and designed in an equitable way.

Data Privacy

You should be protected from abusive data practices via built-in protections and you should have agency over how data about you is used.

Blueprint for an AI Bill of Rights

Safe and Effective Systems

You should be protected from unsafe or ineffective systems.

Algorithmic Discrimination Protections

You should not face discrimination by algorithms and systems should be used and designed in an equitable way.

Data Privacy

You should be protected from abusive data practices via built-in protections and you should have agency over how data about you is used.

Notice and Explanation

You should know when an automated system is being used and understand how and why it contributes to outcomes that impact you.

Blueprint for an AI Bill of Rights

Safe and Effective Systems

You should be protected from unsafe or ineffective systems.

Algorithmic Discrimination Protections

You should not face discrimination by algorithms and systems should be used and designed in an equitable way.

Data Privacy

You should be protected from abusive data practices via built-in protections and you should have agency over how data about you is used.

Notice and Explanation

You should know when an automated system is being used and understand how and why it contributes to outcomes that impact you.

Human Alternatives, Consideration, and Fallback

You should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems you encounter.

Listening to the American People

Adobe
American Civil Liberties Union (ACLU)
The Aspen Commission on Information Disorder
The Awood Center
The Australian Human Rights Commission
Biometrics Institute
The Brookings Institute
BSA | The Software Alliance
Cantellus Group
Center for American Progress
Center for Democracy and Technology
Center on Privacy and Technology at Georgetown Law
Christiana Care
Color of Change
Coworker
Data Robot
Data Trust Alliance
Data and Society Research Institute
Deepmind
EdSAFE AI Alliance

Electronic Privacy Information Center (EPIC)
Encode Justice
Equal AI
Google
Hitachi's AI Policy Committee
The Innocence Project
Institute of Electrical and Electronics Engineers (IEEE)
Intuit
Lawyers Committee for Civil Rights Under Law
Legal Aid Society
The Leadership Conference on Civil and Human Rights
Meta
Microsoft
The MIT AI Policy Forum
Movement Alliance Project
The National Association of Criminal Defense Lawyers
O'Neil Risk Consulting & Algorithmic Auditing

The Partnership on AI
Pinterest
The Plaintext Group
pymetrics
SAP
The Security Industry Association
Software and Information Industry Association (SIIA)
Special Competitive Studies Project
Thorn
United for Respect
University of California at Berkeley Citris Policy Lab
University of California at Berkeley Labor Center
Unfinished/Project Liberty
Upturn
US Chamber of Commerce
US Chamber of Commerce Technology Engagement Center A.I. Working Group
Vibrent Health
Warehouse Worker Resource Center
Waymap

58 meetings with industry and civil society

150+ emails to public email address

30+ federal departments, agencies, and components

FROM PRINCIPLES TO PRACTICE

A Technical Companion to the Blueprint for an AI Bill of Rights

Applying the Blueprint for an AI Bill of Rights

THIS FRAMEWORK DESCRIBES PROTECTIONS THAT SHOULD BE APPLIED WITH RESPECT TO ALL AUTOMATED SYSTEMS THAT HAVE THE POTENTIAL TO MEANINGFULLY IMPACT INDIVIDUALS' OR COMMUNITIES' EXERCISE OF:

RIGHTS, OPPORTUNITIES, OR ACCESS

Civil rights, civil liberties, and privacy, including freedom of speech, voting, and protections from discrimination, excessive punishment, unlawful surveillance, and violations of privacy and other freedoms in both public and private sector contexts;

Equal opportunities, including equitable access to education, housing, credit, employment, and other programs; or,

Access to critical resources or services, such as healthcare, financial services, safety, social services, non-deceptive information about goods and services, and government benefits.

A Technical Companion
to the Blueprint for an AI Bill of Rights

1 **WHY THIS PRINCIPLE IS IMPORTANT:**

This section provides a brief summary of the problems that the principle seeks to address and protect against, including illustrative examples.

2 **WHAT SHOULD BE EXPECTED OF AUTOMATED SYSTEMS:**

- The expectations for automated systems are meant to serve as a blueprint for the development of additional technical standards and practices that should be tailored for particular sectors and contexts.

- This section outlines practical steps that can be implemented to realize the vision of the Blueprint for an AI Bill of Rights. The expectations laid out often mirror existing practices for technology development, including pre-deployment testing, ongoing monitoring, and governance structures for automated systems, but also go further to address unmet needs for change and offer concrete directions for how those changes can be made.

3 **HOW THESE PRINCIPLES CAN MOVE INTO PRACTICE:**

This section provides real-life examples of how these guiding principles can become reality, through laws, policies, and practices. It describes practical technical and sociotechnical approaches to protecting rights, opportunities, and access.

Extra Protections for Data Related to Sensitive Domains

- **Some domains, including health, employment, education, criminal justice, and personal finance, have long been singled out as sensitive domains deserving of enhanced data protections.** This is due to the intimate nature of these domains as well as the inability of individuals to opt out of these domains in any meaningful way, and the historical discrimination that has often accompanied data knowledge. Domains understood by the public to be sensitive also change over time, including because of technological developments. Tracking and monitoring technologies, personal tracking devices, and our extensive data footprints are used and misused more than ever before; as such, the protections afforded by current legal guidelines may be inadequate. The American public deserves assurances that data related to such sensitive domains is protected and used appropriately and only in narrowly defined contexts with clear benefits to the individual and/or society.

- **To this end, automated systems that collect, use, share, or store data related to these sensitive domains should meet additional expectations.** Data and metadata are sensitive if they pertain to an individual in a sensitive domain (defined below); are generated by technologies used in a sensitive domain; can be used to infer data from a sensitive domain or sensitive data about an individual; or have the reasonable potential to be used in ways that are likely to expose individuals to meaningful harm, such as a loss of privacy or financial harm due to identity theft.

Coalition for Health AI's

Blueprint for Trustworthy AI Implementation Guidance and Assurance for Healthcare

BLUEPRINT FOR TRUSTWORTHY AI IMPLEMENTATION GUIDANCE AND ASSURANCE FOR HEALTHCARE

COALITION FOR HEALTH AI
DECEMBER 7, 2022

www.whitehouse.gov/ostp/ai-bill-of-rights

WH.GOV

- Laura Adams, NAM

CHAI

Coalition for Health AI

NAM/CHAI Collaboration for Trustworthy AI

April 5, 2023

The NAM's Code of Conduct (AICC) Initiative Objectives

1. **Develop a "current best practice" AI Code of Conduct for health, healthcare, and biomedical science,** honoring, building upon, and harmonizing the outstanding work to date.

2. **Align the field,** identifying the roles and responsibilities of each stakeholder at each stage of the AI lifecycle vis-à-vis the Code.

3. **Describe a national architecture** that gives rise to and supports responsible AI in health, healthcare and biomedical science.

Three Key Ways the CHAI and NAM Initiatives Intersect and are Interdependent

1. **The CHAI Blueprint will be a foundational element of and inform the AI Code of Conduct.**

   In addition, the NAM is conducting a systematic review of the existing healthcare AI guidelines, frameworks, and principles in the following areas:

   - From the published literature since 2010

   - Developed by professional medical societies

   - Issued by the federal government.

Three Key Ways the CHAI and the NAM Initiatives Intersect and are Interdependent (cont'd)

2. The NAM will be looking to CHAI for robust best practice technical implementation

3. CHAI's work on algorithmic assurance labs will provide an opportunity to test and improve the Code of Conduct and will play a very significant role in assuring that under-resourced delivery systems and small start-up innovators are not disadvantaged, and experiencing the digital divide widening even further.

The work and final deliverables of these initiatives are mutually reinforcing and coordinated to establish an AI Code of Conduct and Technical Framework for Health AI Assurance.

- Gil Alterovitz, Harvard Medical School
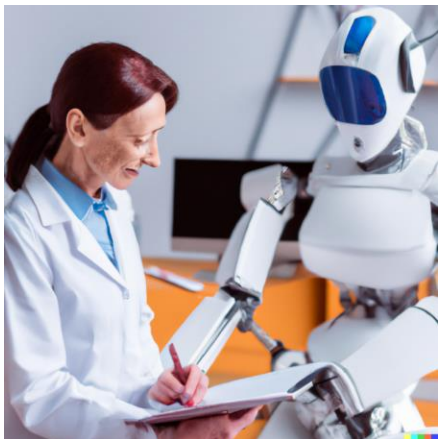
BLUEPRINT FOR TRUSTWORTHY AI IMPLEMENTATION GUIDANCE AND ASSURANCE FOR HEALTHCARE

COALITION FOR HEALTH AI

APRIL 04, 2023
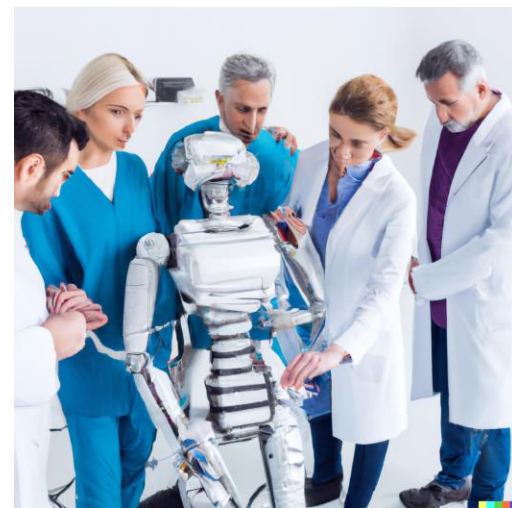
News

Elon Musk and other tech experts call for 'pause' on advanced AI systems

More than 1,000 researchers and executives sign open letter to halt what they describe as a 'dangerous' arms race

The letter follows a rush of AI launches over the past five months, including Microsoft-backed OpenAI's ChatGPT in November and this month's release of GPT-4 © Rafael Henrique press/Dreamstime

MOTHERBOARD
TECH BY VICE

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

The incident raises concerns about guardrails around quickly-proliferating conversational AI models.

By Chloe Xiang

"AI labs and independent experts should use this pause to jointly develop and implement a set of shared safety protocols for advanced AI design and development that are rigorously audited and overseen by independent outside experts."

"when Motherboard tested the [updated] platform, it was still able to share very harmful content regarding suicide…"

Introduction

- AI can accelerate clinical research and improve healthcare delivery.

- Implementation guidance focusing on health impact, fairness, ethics, and equity principles is needed to ensure AI benefits all populations.

- CHAI proposed a blueprint and an assurance approach to enable trustworthy AI in healthcare, facilitate guidelines, and foster a strong healthcare community.

Overview

- Using a common set of principles to build and facilitate the use of AI tools is necessary.

- An independent assurance lab can evaluate processes and tools submitted by health systems, tool developers, and vendors to ensure the ethical use of AI.

- An advisory body can help ensure equity and access to trustworthy health AI for patients across different health systems, including small, rural, and/or resource-constrained ones.


Health System Preparedness and Assessment

AI Tools, Trustworthiness, Transparency, etc. across the Lifecycle

Integrated Data Infrastructure for Discovery, Evaluation, and Assurance

Collaboration, Guiding Principles, and Leadership

Recommendations

- Establishing an assurance lab and advisory service infrastructure can help create an ecosystem for AI solutions in healthcare.

- The infrastructure includes at least four components: a shared definition of value, registries of tools, legal agreement templates for data sharing, and sandbox environments.

- To ensure initial buy-in, it is important to demonstrate the value proposition of AI solutions for patients and healthcare organizations.

- Structured checklists and maturity models can help assess the potential use of AI tools and the level of maturity of organizations and developers.

- Registries of AI tools can empower providers/patients, enable comparison of algorithms and analytic options, and facilitate precision medicine.

- Sandboxes for testing models locally and evaluation and monitoring platforms can ensure long-term reliability and identify bias and performance degradation.
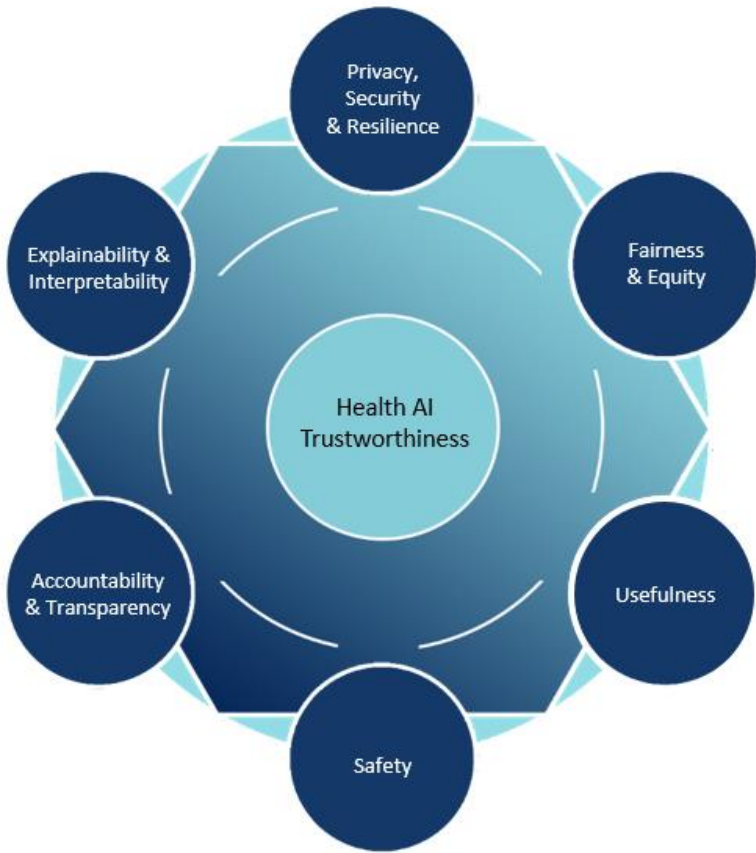
Mapping NIST RMF to CHAI Assurance Standards

-   # Nicoleta Economou, Duke Health

CHAI
Coalition for Health AI

Core Principles
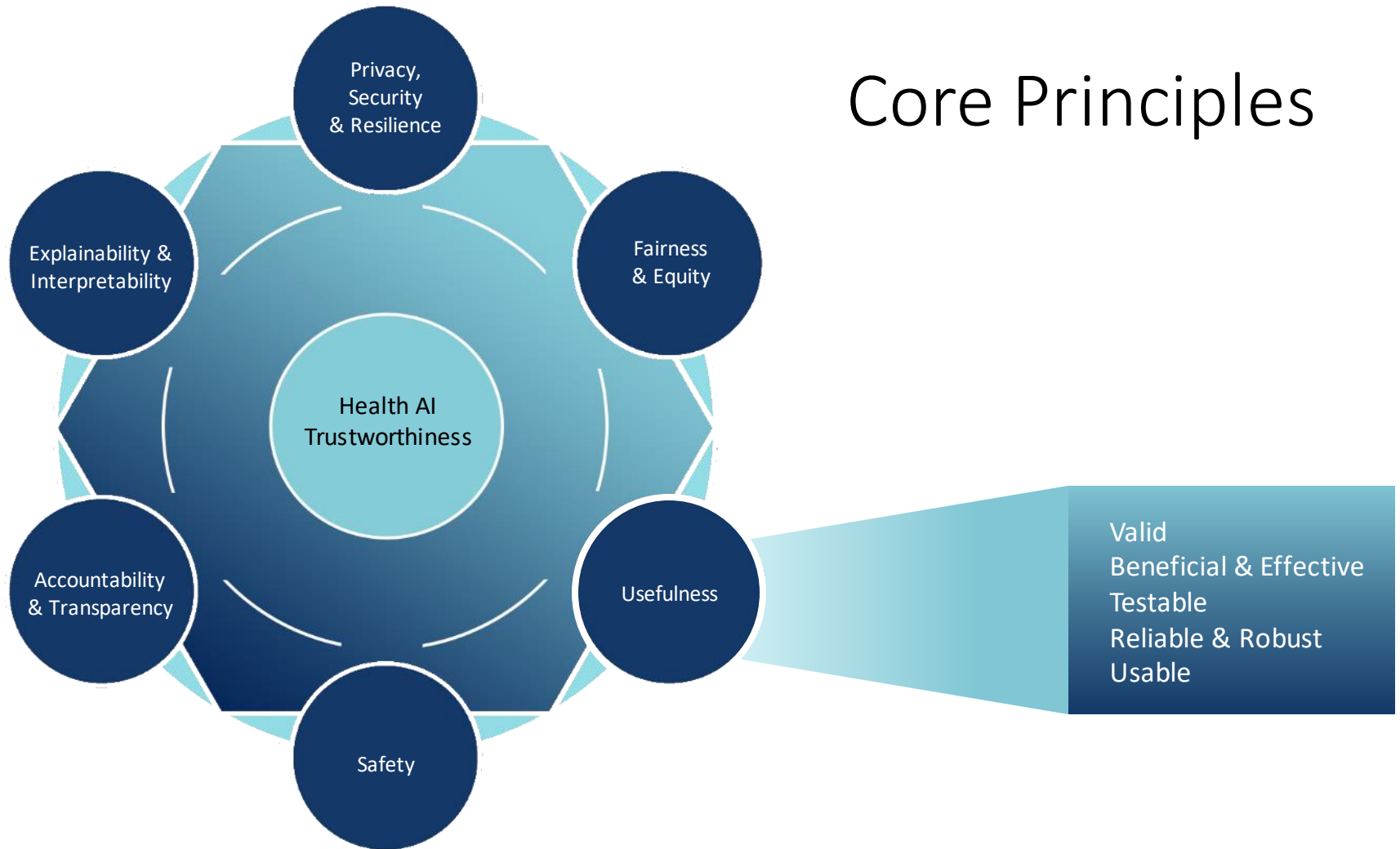
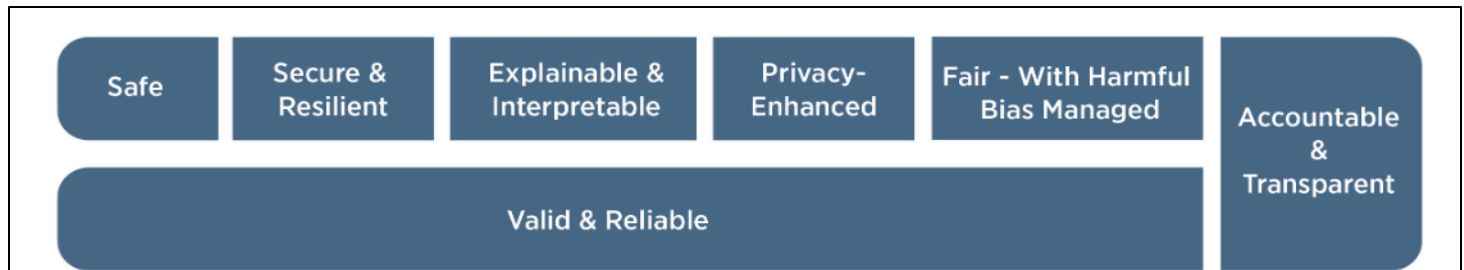www.whitehouse.gov/ostp/ai-bill-of-rights

Fig. 4. Characteristics of trustworthy AI systems. Valid & Reliable is a necessary condition of trustworthiness and is shown as the base for other trustworthiness characteristics. Accountable & Transparent is shown as a vertical box because it relates to all other characteristics.

NIST AI Risk Management Framework (1.0), 2023

https://www.coalitionforhealthai.org/

Coalition for Health AI - Join Us to be a Work Group Volunteer!

Thank you for volunteering to join us in our mission and work to generate an assurance standards guide and evaluation rubric as well as catalogue transparency information to advance the credibility and trustworthiness of health AI tools deployed today. As a SME Volunteering to join a WG we need your help to support and help generate these products.

We look forward to collaborating with you and will follow up with you regarding this effort after you provide more information below.

https://forms.office.com/g/ULirvZqx9K

-    John Halamka,
Mayo

- Peter Embi, Vanderbilt University Medical Center

CHAI
Coalition for Health AI

Taking AI into the Real-World
of Healthcare

-   Eric Horvitz, Microsoft

- Michael Pencina, Duke Health

https://www.coalitionforhealthai.org/
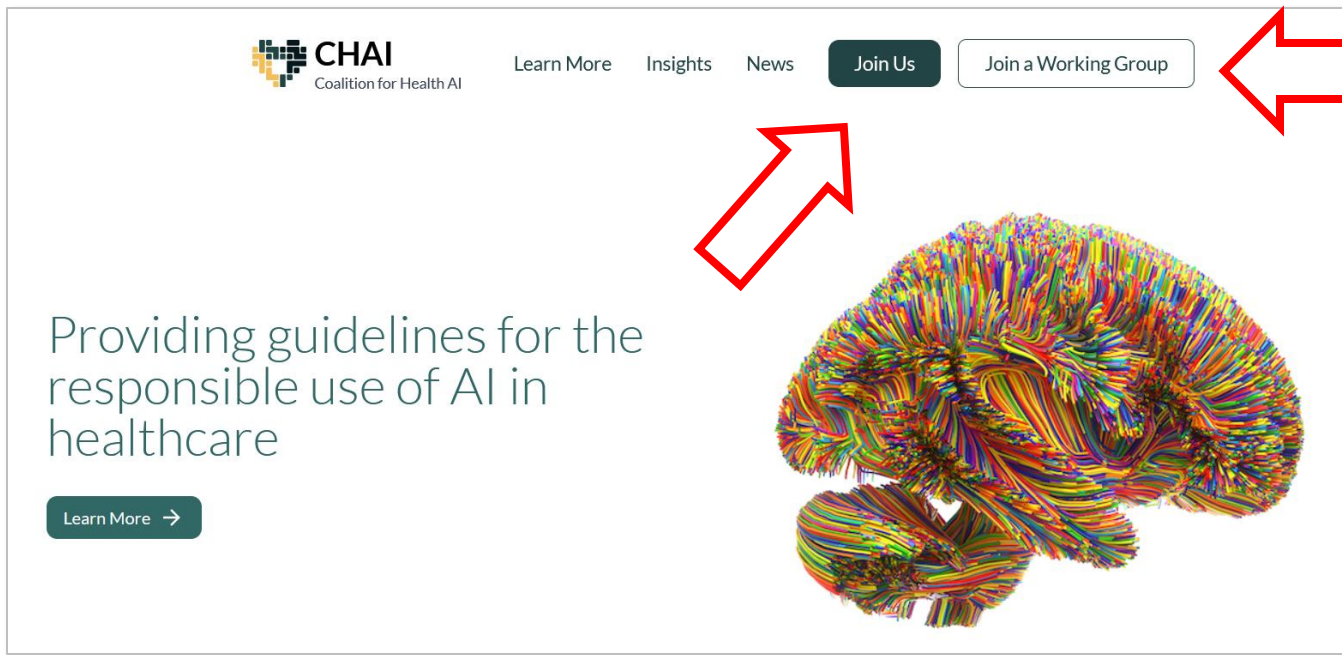
Thank You

For questions please email

us at; CHAI@MITRE.org .

Please be sure to address the

question to who you asking.

(ie: CHAI; Brian Anderson; etc.)

https://forms.office.com/g/ULirvZqx9K