Responsible
Artificial Intelligence
Institute

Advancing Trusted AI

ICRP Annual Meeting - Toronto, Canada

RAI Institute Session

Implementing RAI in Healthcare: Foundations, Practices, and Assessments

May 14, 2024

CCRA
ACRC

ICRP International Cancer Research Partnership

Who We Are

The Responsible AI Institute (RAI Institute) is dedicated to enabling successful responsible AI efforts in organizations.

Independent not-for-profit

Unparalleled subject matter expertise

Member–driven

Community–focused

Leveraging international best practices

Work directly with policy makers and regulators

Community of Responsible AI Leaders

Members

Community of Responsible AI Leaders

Ecosystem Collaborators

Academia and Government

The ‘what’ and ‘why’ behind RAI

Why is AI Everywhere?

AI went mainstream

Generative AI and the proliferation of easy to use tools like ChatGPT

It's hyped, but not that much

It is the technological revolution of this era. We are in the dial up stage of the internet

Democratization Means Scale

The increased use of AI tools means that more things related to AI are – some good, some bad

Legislation and Regulation

Conversations around guardrails, regulation and legislation are accelerating to address the need

It's Emotional!

AI makes us fundamentally question what makes us human and our fear of being obsolete

AI will have a positive impact
But there is fear that …

It hurts people

- Incorrect identification through facial recognition
- Biased predictions for jail sentences, insurance rates, access to credit, or health treatments

Wrongfully Accused by an Algorithm

Wrongfully Accused by an Algorithm, NYT

It doesn't see us

- Gender and other minority bias for hiring practices
- Lack of recognition by computer for people with disabilities, women, people of colour, or other groups

A computer program used for bail and sentencing decisions was labeled biased against blacks

Machine Bias, ProPublica

It's used without our knowledge or permission

- Contact tracing apps for health
- Complex terms of service agreements
- Automated purchases through voice recognition systems

Broken promises: How Singapore lost trust on contact tracing privacy

MIT Technology Review

Harm Happens in Different Ways

Harm types	Definitions	Example
Unintended	Harms arising from AI systems behaving in unintended ways.	• Automated decisioning tools discriminate against certain populations for mortgage lending or property rentals. • Lack of calibration in automated lending models limiting credit to homosexual couples.
Intentional	Harms arising from intentional trade-offs, risk taking, or malicious uses of AI.	• Calibration of a hiring models ignores maternity or paternity leaves in their analysis and discards candidates for employment gap. • Microtargeting during elections greatly shifts public opinion and election results to favor a candidate.
Systemic	Unintended consequences from the deployment of technology that shape the broader environment and society.	• Job advertisements are presented to and attract only particular demographics. • Lack of credit history for minority/low-income groups leads to inequality in lending and increases the wealth gap. • AI-driven layoffs disproportionately affect minority workers more frequently, leading to increased job displacement.

eClinicalMedicine
Part of THE LANCET *Discovery Science*

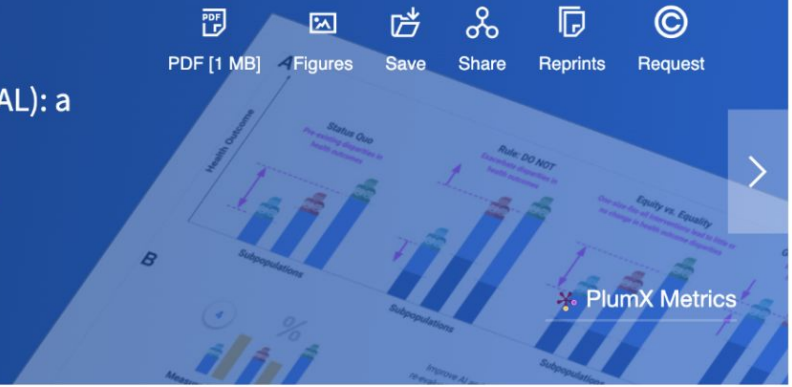
ARTICLES | ONLINE FIRST, 102479

Health equity assessment of machine learning performance (HEAL): a framework and dermatology AI model case study

Mike Schaekermann • Terry Spitz • Malcolm Pyles • Heather Cole-Lewis • Ellery Wulczyn • Stephen R. Pfohl • et al. Show all authors • Show footnotes

Open Access • Published: March 14, 2024 • DOI: https://doi.org/10.1016/j.eclinm.2024.102479

PDF [1 MB] Figures Save Share Reprints Request



Summary
Keywords
Introduction
Methods
Results

Information for Authors

Summary

Background

Artificial intelligence (AI) has repeatedly been shown to encode historical inequities in healthcare. We aimed to develop a framework to quantitatively assess the performance equity of health AI technologies and to illustrate its utility via a case study.

Review Article | Open Access | Published: 01 June 2020

Sex and gender differences and biases in artifi intelligence for biomedicine and healthcare

Davide Cirillo ✉, Silvina Catuara-Solarz, Czuee Morey, Emre Guney, Laia Subirats, Simona Annalisa Gigante, Alfonso Valencia, María José Rementeria, Antonella Santuccione Chadh Mavridis

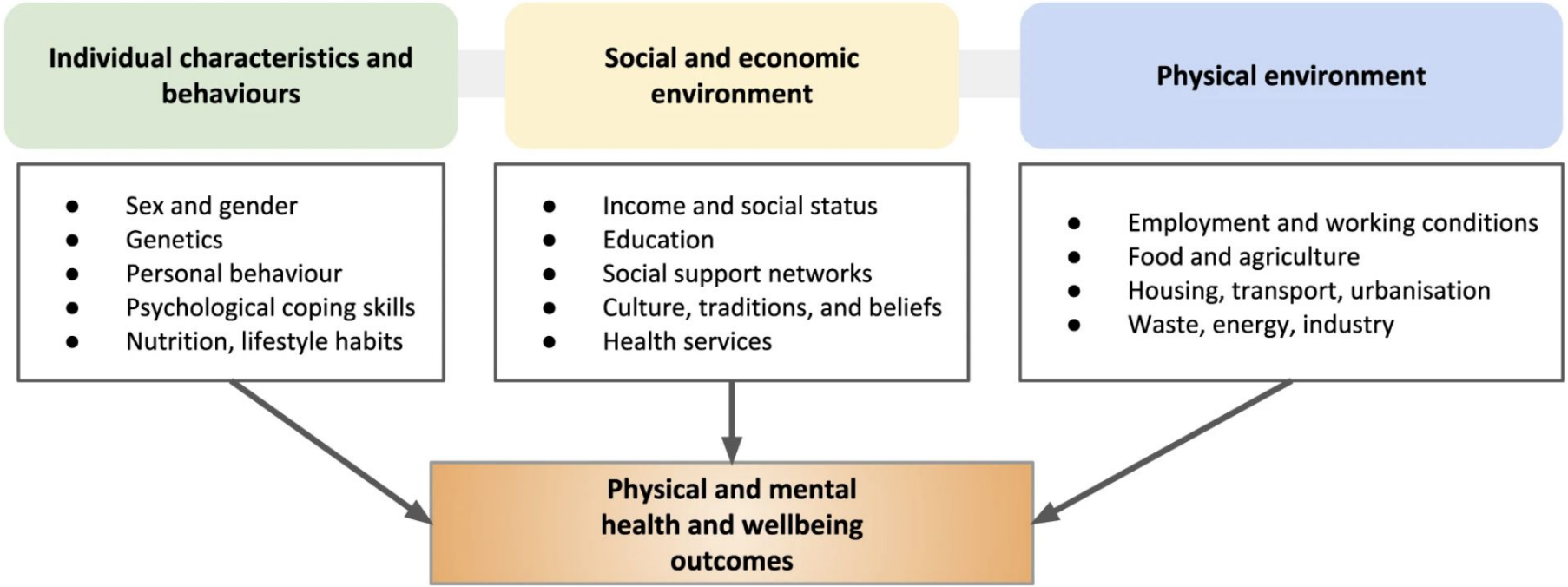
npj Digital Medicine **3**, Article number: 81 (2020) | Cite this article

14k Accesses | **7** Citations | **151** Altmetric | Metrics

Abstract

Precision Medicine implies a deep understanding of inter-individual differences i disease that are due to genetic and environmental factors. To acquire such unde there is a need for the implementation of different types of technologies based c intelligence (AI) that enable the identification of biomedically relevant patterns, f progress towards individually tailored preventative and therapeutic interventions

Determinants of physical and mental health and wellbeing



EDITORIAL · 09 DECEMBER 2020

Accounting for sex and gender makes for better science

The European Commission is set to insist on steps that will make research design more inclusive.

transparent

Pronunciation /trɑːnˈspar(ə)nt/ /trɑːnˈspɛːr(ə)nt/ /tranˈspɛːr(ə)nt/ /tranˈspar(ə)nt/

ADJECTIVE

1 (of a material or article) allowing light to pass through so that objects behind can be distinctly seen.

'transparent blue water'

'fine transparent fabrics'

+ More example sentences + Synonyms

2 Easy to perceive or detect.

'the residents will see through any transparent attempt to buy their votes'

'the meaning of the poem is by no means transparent'

+ More example sentences + Synonyms

2.1 Having thoughts or feelings that are easily perceived; open.

'you'd be no good at poker—you're too transparent'

+ More example sentences

2.2 (of an organization or its activities) open to public scrutiny.

'if you had transparent government procurement, corruption would go away'

+ More example sentences + Synonyms

3 *Computing*

(of a process or interface) functioning without the user being aware of its presence.

transparent

Pronunciation /traːnˈspar(ə)nt/ /traːnˈspɛːr(ə)nt/ /tranˈspɛːr(ə)nt/ /tranˈspar(ə)nt/

ADJECTIVE

1 (of a material or article) allowing light to pass through so that objects behind can be distinctly seen.

'transparent blue water'

'fine transparent fabrics'

+ More example sentences + Synonyms

2 Easy to perceive or detect.

'the residents will see through any transparent attempt to buy their votes'

'the meaning of the poem is by no means transparent'

+ More example sentences + Synonyms

2.1 Having thoughts or feelings that are easily perceived; open.

'you'd be no good at poker—you're too transparent'

+ More example sentences

2.2 (of an organization or its activities) open to public scrutiny.

'if you had transparent government procurement, corruption would go away'

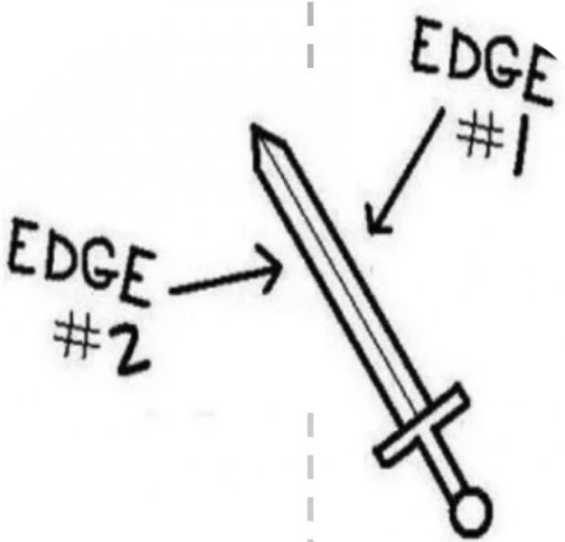
+ More example sentences + Synonyms

3 *Computing*

(of a process or interface) functioning without the user being aware of its presence.

TRUST

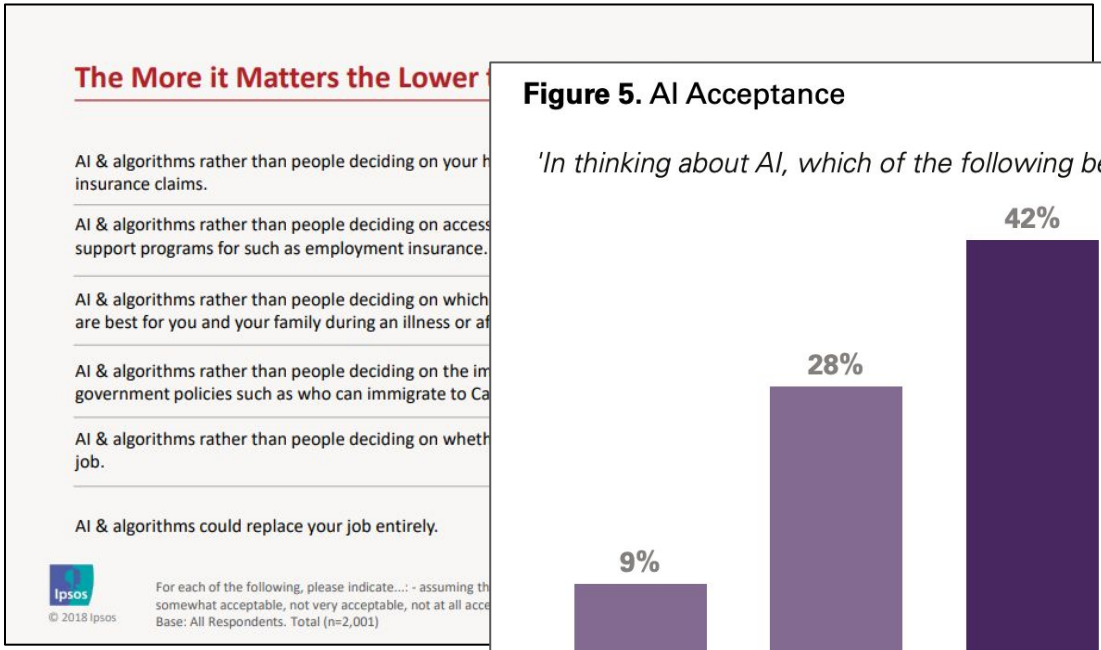
AI could play a significant role in **mitigating** existing care **inequalities.**



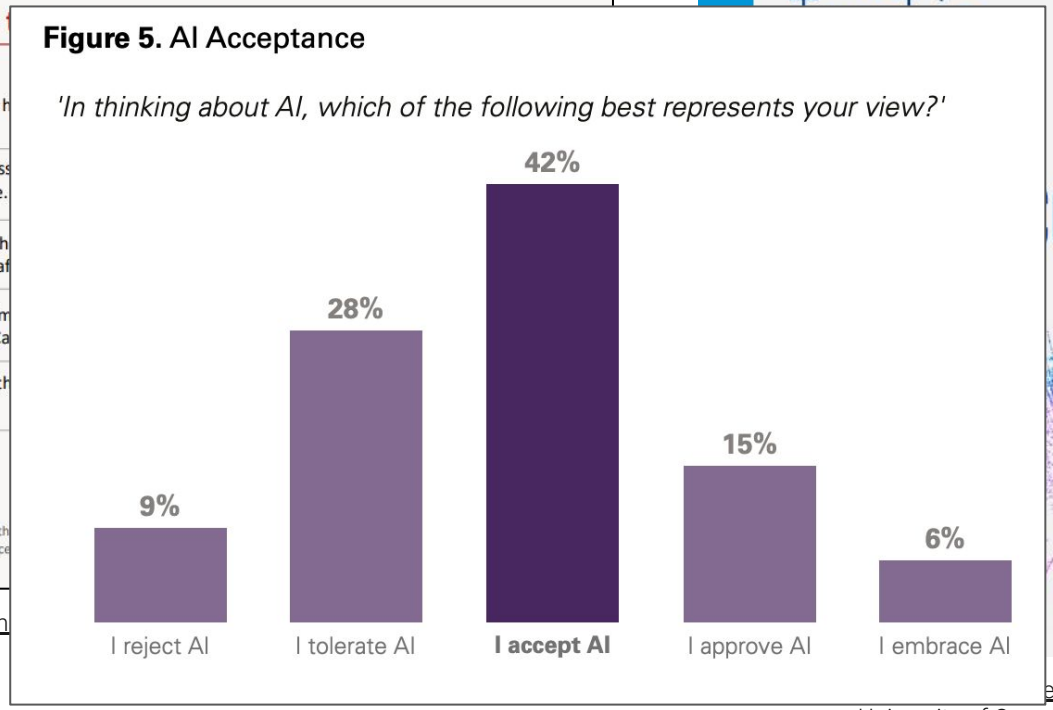
AI can **exacerbate** & **perpetuate** existing biases for sex and gender.

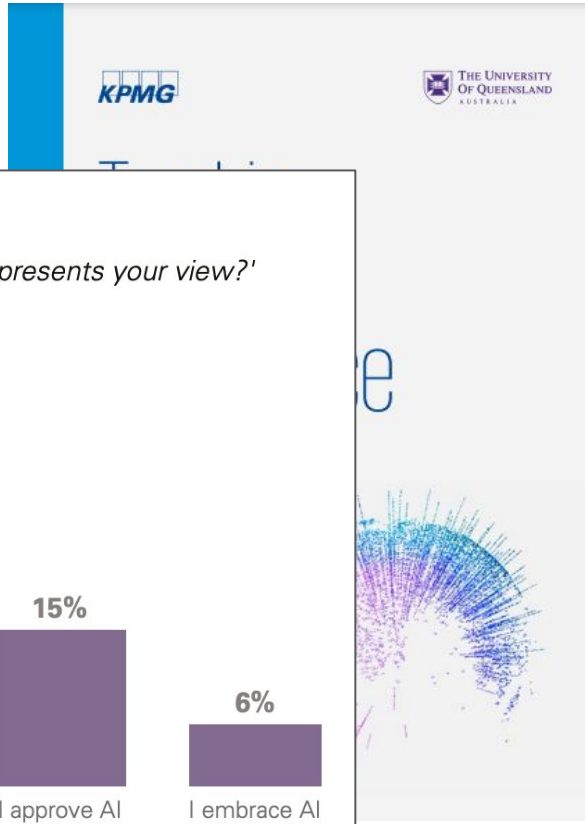
TRUST

Public Trust in AI is Low



Tech for Good: A Canadian Perspective, In

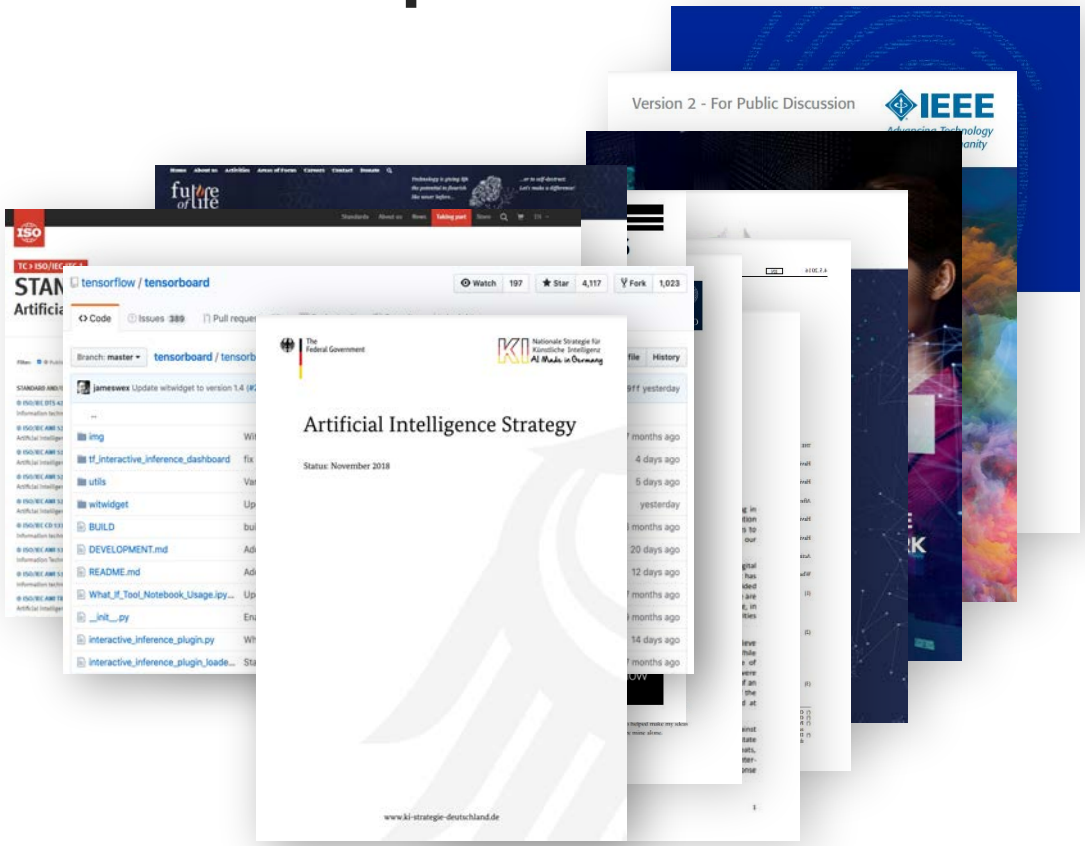




e, KPMG and the University of Queensland

Evolving RAI Best Practices Landscape

- There has been a significant demand from companies, academics, and governments for understanding RAI.
- **Soft law** instruments - reports, research, principles, documents, and tools - have emerged in response.
- While lots of good advice is readily available, it is difficult to navigate **what to do** and **how to implement**.



Evolving RAI Legislation Landscape

- Governments globally have formalized national, regional and global approaches to AI, including accompanying policies and legislation.
- These hard law instruments are care **binding and enforceable** within the scoped jurisdictions.
- Failure to recognize **constraining forces** of AI for effective regulation.

European Parliament
2019-2024

TEXTS ADOPTED

P9_TA(2024)0138

Artificial Intelligence Act

European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))

(Ordinary legislative procedure: first reading)

The European Parliament,

– having regard to the Commission proposal to Parliament and the Council (COM(2021)0206),

– having regard to Article 294(2) and Articles 16 and 114 of the Treaty on the Functioning of the European Union, pursuant to which the Commission submitted the proposal to Parliament (C9-0146/2021),

– having regard to Article 294(3) of the Treaty on the Functioning of the European Union,

– having regard to the opinion of the European Central Bank of 29 December 2021[1],

– having regard to the opinion of the European Economic and Social Committee of 22 September 2021[2],

– having regard to the provisional agreement approved by the committees responsible under Rule 74(4) of its Rules of Procedure and the undertaking given by the Council representative by letter of 2 February 2024 to approve Parliament's position, in accordance with Article 294(4) of the Treaty on the Functioning of the European Union,

– having regard to Rule 59 of its Rules of Procedure,

– having regard to the joint deliberations of the Committee on Internal Market and Consumer Protection and the Committee on Civil Liberties, Justice and Home Affairs under Rule 58 of the Rules of Procedure,

– having regard to the opinion of the Committee on Industry, Research and Energy, the Committee on Culture and Education, the Committee on Legal Affairs, the Committee

[1] OJ C 115, 11.3.2022, p. 5.
[2] OJ C 517, 22.12.2021, p. 56.

Safe, Secure,
opment and
elligence
AL ACTIONS
nstitution and the laws
follows:
extraordinary
se has the potential to
nore prosperous,
rresponsible use could
n, bias, and
le competition; and
and realizing its
s. This endeavor
nt, the private sector,
werning the
is therefore advancing
doing so. The rapid
the United States to
my, and society.
ho build it, the people
believe that the power
eativity, diversity, and
rived in past eras of
ain in this moment. We
urity, and opportunity

Defining Responsible AI

”

Responsible AI practices can help align the decisions about AI system design, development, and uses with **intended aim and values**. Core concepts in responsible AI emphasize **human centricity**, s**ocial responsibility**, and **sustainability**.

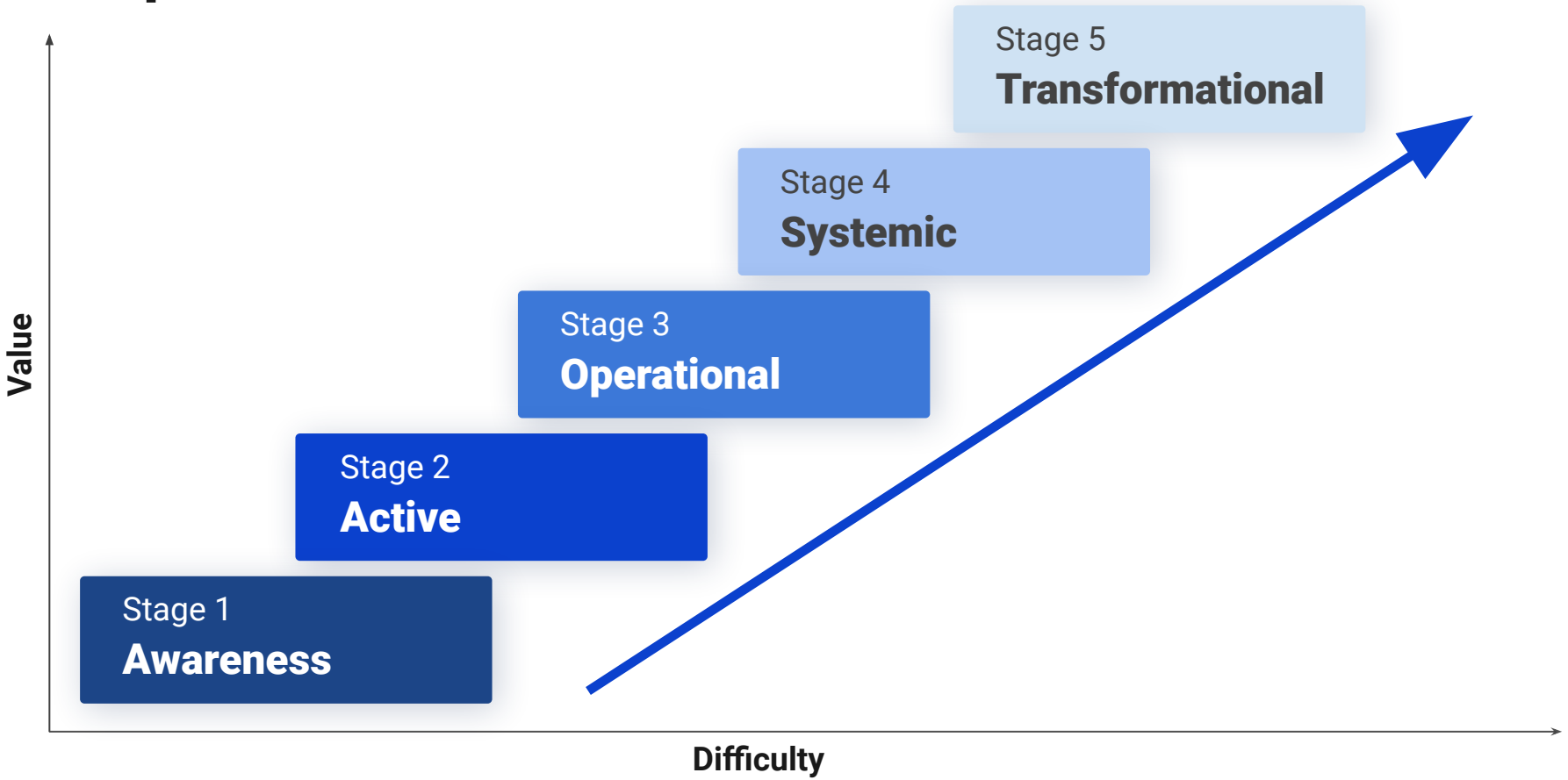
NIST AI RMF

Key Action Items and Considerations

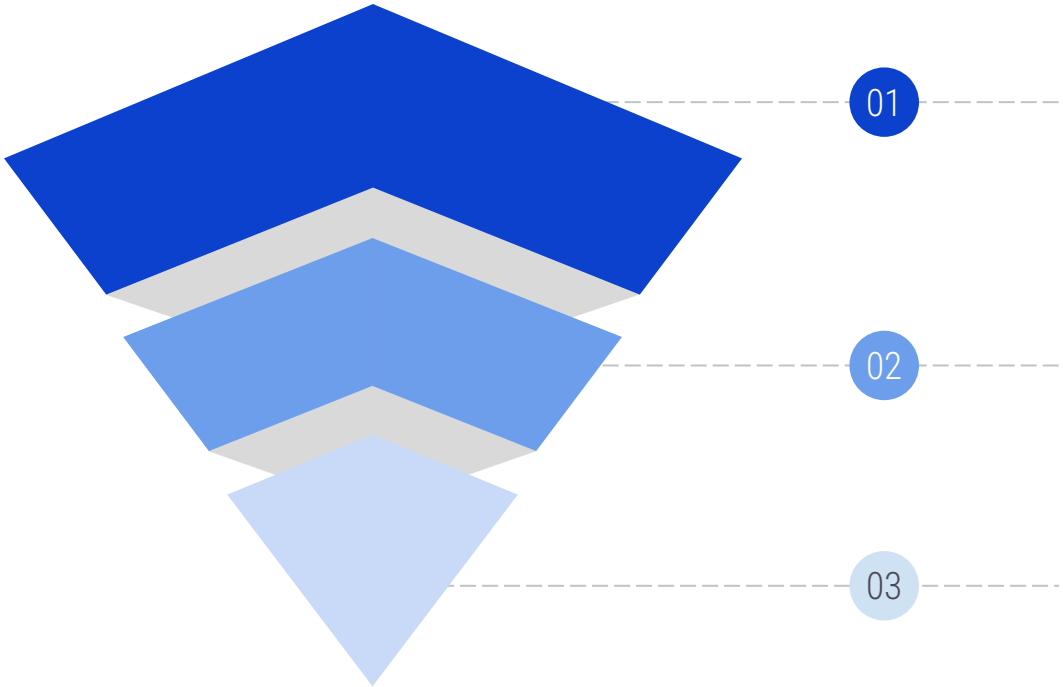
”

No **AI values** without accompanying **procedures**.
No accompanying procedures without **metrics and controls**.
No effective metrics and controls without adequate **assurance**.

Responsible AI is a Practice



Building a Resilient Responsible AI Practice



Corporate Principles/ Values

- Commitment to Responsible Innovation
- ESG
- Operating Principles

AI Principles

- RAI Principles
- Framework

Implementation

- Strategies
- Measurements
- Tactics
- Workstreams

Ways RAI Institute Conformity Assessment Is Used

1st SELF-ATTESTATION OF CONFORMITY TO THE RESPONSIBLE AI INSTITUTE CERTIFICATION FRAMEWORK

2nd INDEPENDENT ATTESTATION OF CONFORMITY TO THE RESPONSIBLE AI INSTITUTE CERTIFICATION FRAMEWORK

RESPONSIBLE AI INSTITUTE · CERTIFIED ·

First Party
For low risk systems

Evaluate

Internal self-assessment of AI system

Best used in the **design phase** as a health check or check list

Second Party
For medium risk systems

Validate

Second party assessment of AI system with robust recommendations report

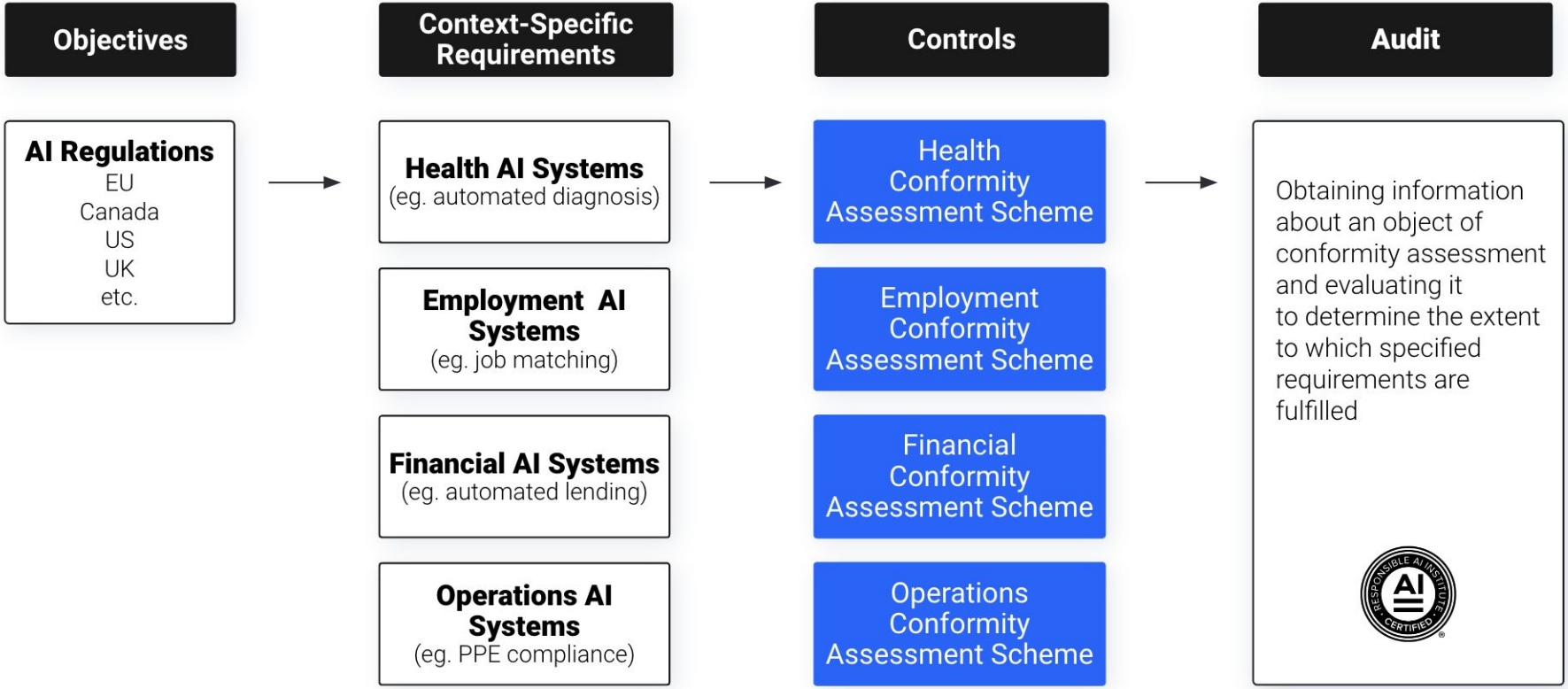
Delivered via **workshop** by RAI Institute or Delivery Partner

Third Party
For high risk systems

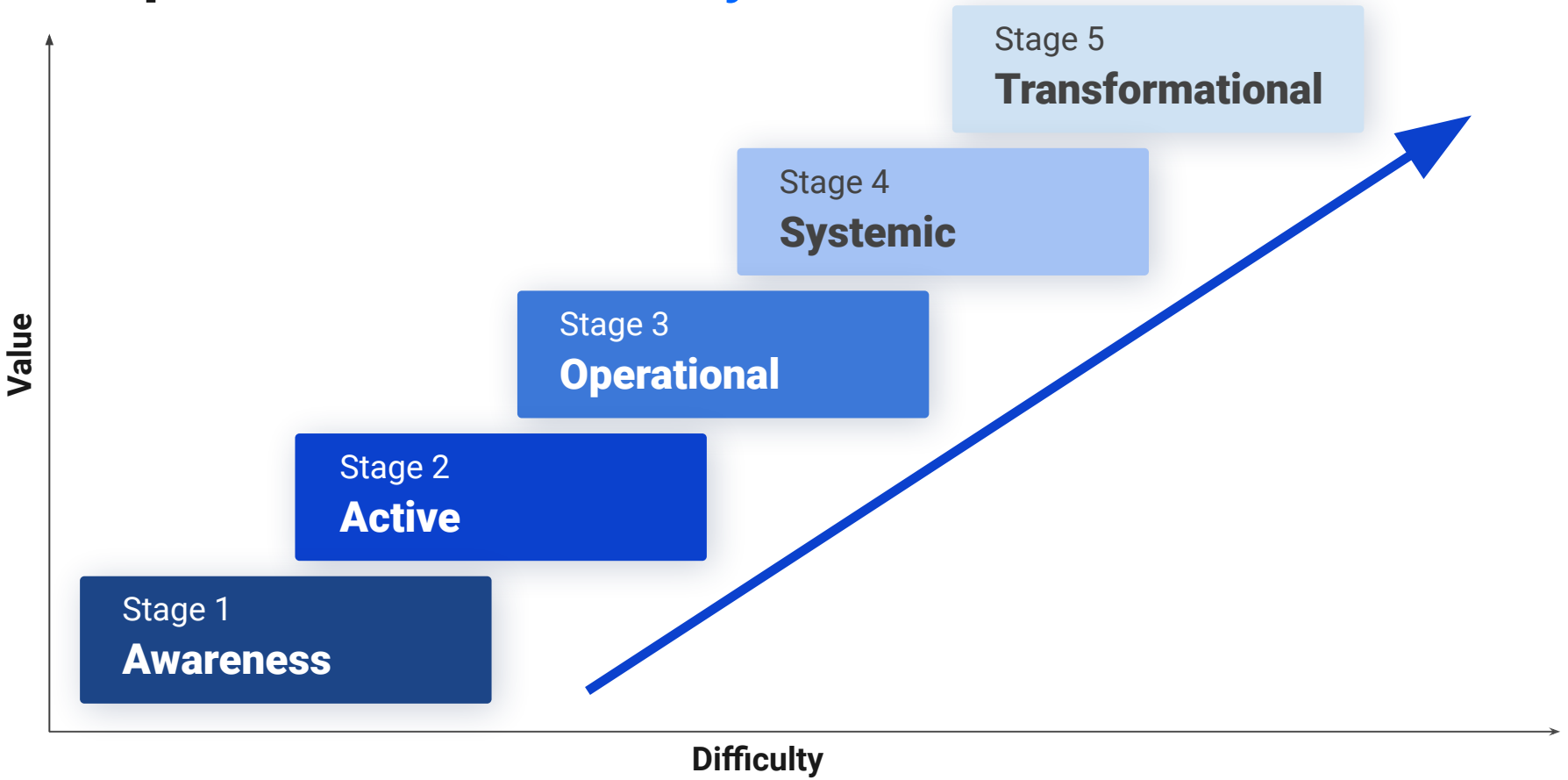
Certify

Formal review of AI system by accredited Certification Body leading to a RAI Institute Certification mark

Conformity Assessment Development Process



Responsible AI is a Journey



Here Is What Leaders Are Doing

Assessing

They are conducting assessments when building, buying or supplying AI

Tracking

They closely track regulation and are proactively preparing

Scaling

Best practices and processes are expanded and consistent

Communicating

They communicate and leverage their commitment to all stakeholders

Responsible Artificial Intelligence Institute

Advancing Trusted AI

CCRA ACRC

ICRP International Cancer Research Partnership

Thank you!

Questions?

For more updates, connect with the Responsible AI Institute at:

www.responsible.ai

Responsible AI Institute

@ResponsibleAI