



Responsible AI

Opportunity and Responsibility in the Era of Artificial Intelligence

Ivan Portilla
portilla@gmail.com
Jesus.Portilla@Colorado.edu+



Responsible AI _ Agenda



Why

What

How

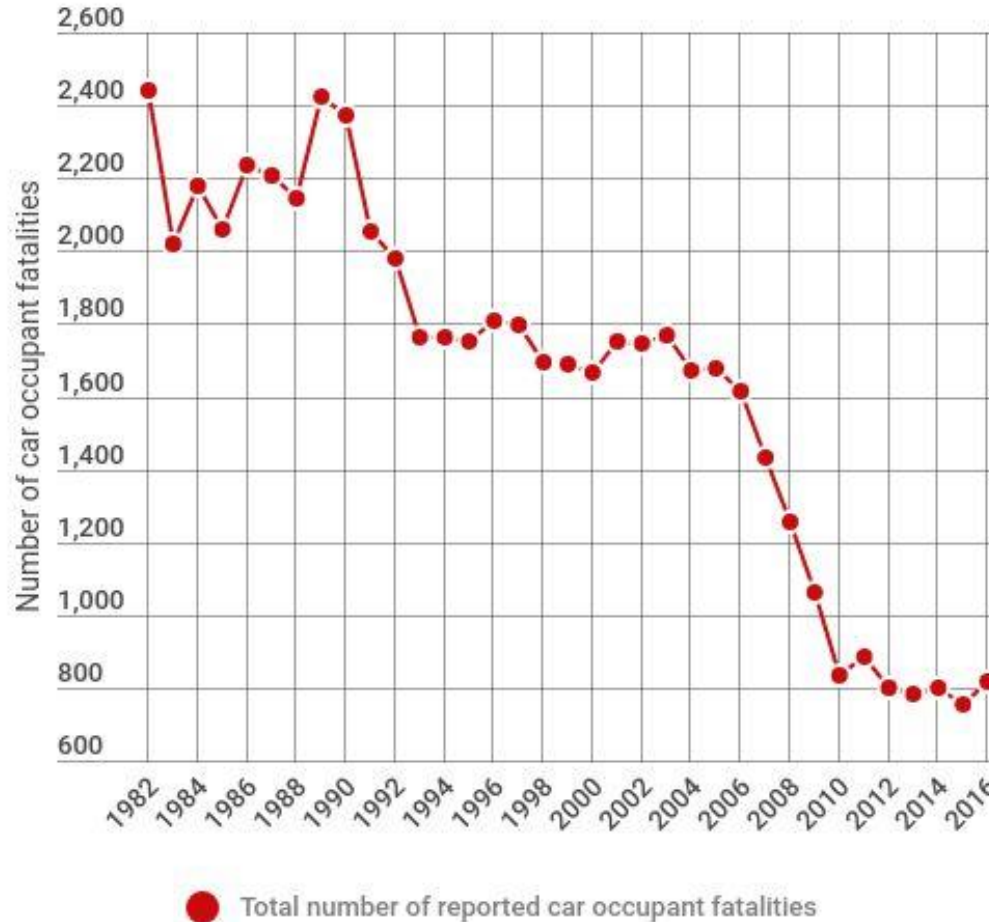


Unregulated vs Regulated Technologies

Seat Belts Save Lives

90.4%

SEAT BELT USE RATE IN 2021



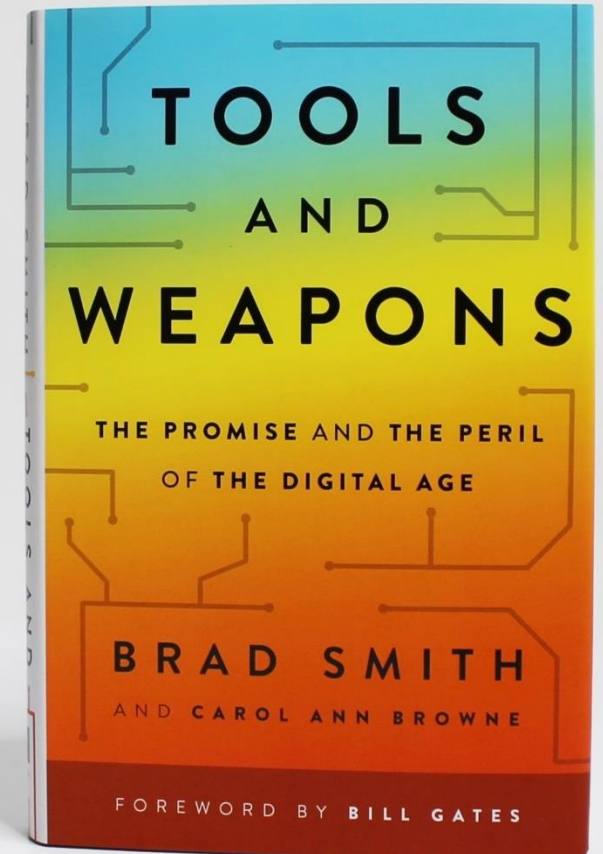
This information is based on National Statistics from the Department of Transport's 'Reported road casualties Great Britain, annual report: 2016'. Correct in January 2018.

Which?

Why responsible AI?

"The more powerful the tool, the greater the benefit or damage it can cause...Technology innovation is not going to slow down. The work to manage it needs to speed up."

Brad Smith
President and Chief Legal Officer, Microsoft



Challenges & Risks with AI

Challenge or Risk	Example
Bias can affect results	A loan-approval model discriminates by gender due to bias in the data with which it was trained
Errors may cause harm	An autonomous vehicle experiences a system failure and causes a collision
Data could be exposed	A medical diagnostic bot is trained using sensitive patient data, which is stored insecurely
Solutions may not work for everyone	A home automation assistant provides no audio output for visually impaired users
Users must trust a complex system	An AI-based financial tool makes investment recommendations - what are they based on?
Who's liable for AI-driven decisions?	An innocent person is convicted of a crime based on evidence from facial recognition – who's responsible?

Today's
debate

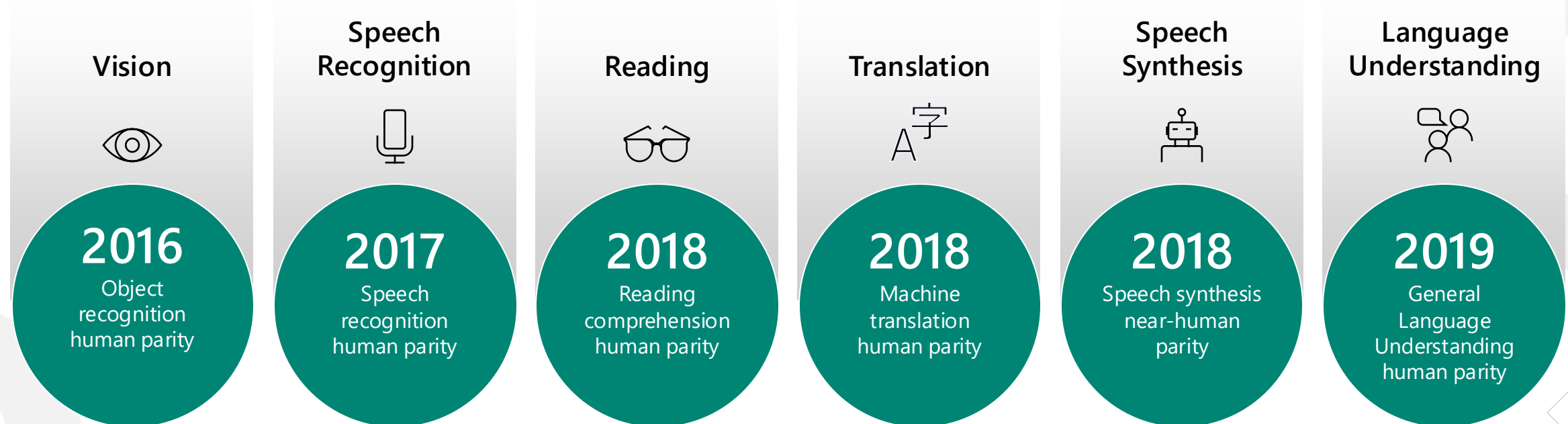
Murdered Insurance CEO Had Deployed an AI to Automatically Deny Benefits for Sick People

The algorithm in question, known as nH Predict, [allegedly had a 90 percent error rate](#).

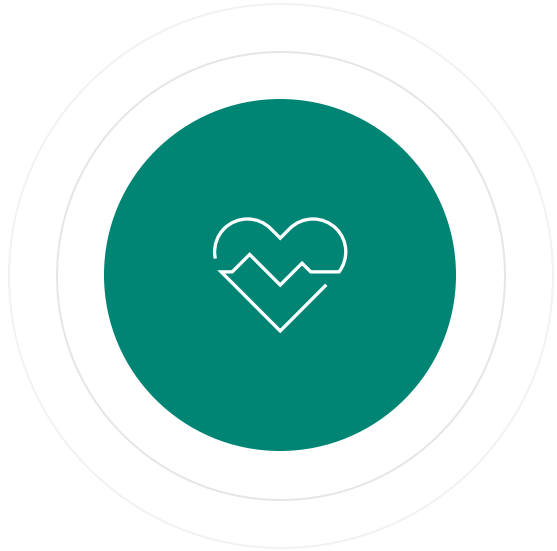


Why responsible AI?

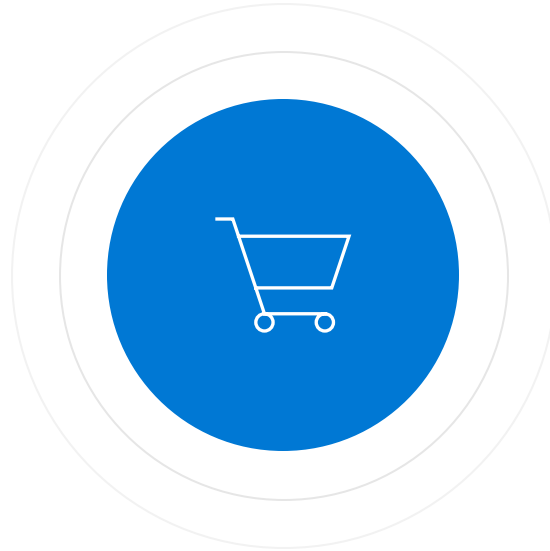
Advancements in AI are different than other technologies because of the **pace of innovation**, and its **proximity to human** intelligence – impacting us at a personal and societal level.



The Opportunities with AI



Healthcare



Retail



Financial Services



Manufacturing

Agenda _Responsible AI



Why

What

How



Microsoft AI Principles

The bottom of the slide features several abstract geometric shapes in various shades of teal and light teal. These include several elongated, rounded rectangles and one circle, all arranged in a dynamic, overlapping pattern that suggests movement and modern design.

Microsoft's AI principles



Fairness



Reliability
& Safety



Privacy &
Security



Inclusiveness



Transparency



Accountability

Microsoft's AI principles



Fairness



Reliability
& Safety



Privacy &
Security



Inclusiveness

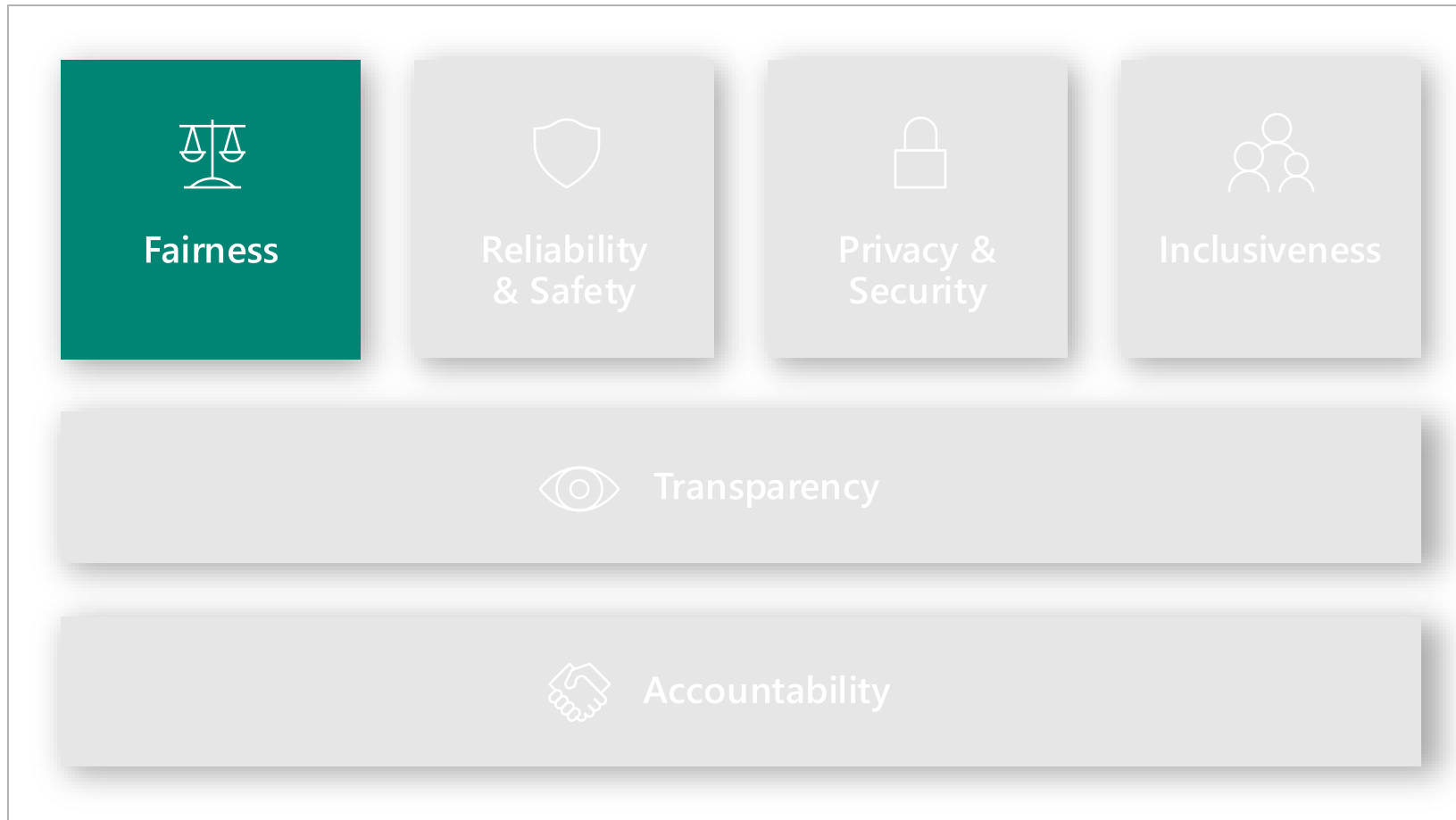


Transparency



Accountability

Microsoft's AI principles



Fairness

MENU ▾

nature

NEWS • 24 OCTOBER 2019

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

Study reveals rampant racism in decision-making software used by US hospitals – and highlights ways to correct it.

Heidi Ledford



...likely than equally ill white people to be referred to programmes that provide more personalized care. Credit: Ed Rusby/vi/redux/eyevine

Fairness



3 Low Risk

DYLAN FUGETT



10 High Risk

BERNARD PARKER

Prior Offense
1 resisting arrest
without violence

Subsequent Offenses
None



Fairness



A doctor taking care of sick children



A black doctor taking care of sick white children

Fairness

Is there gender bias in lending?

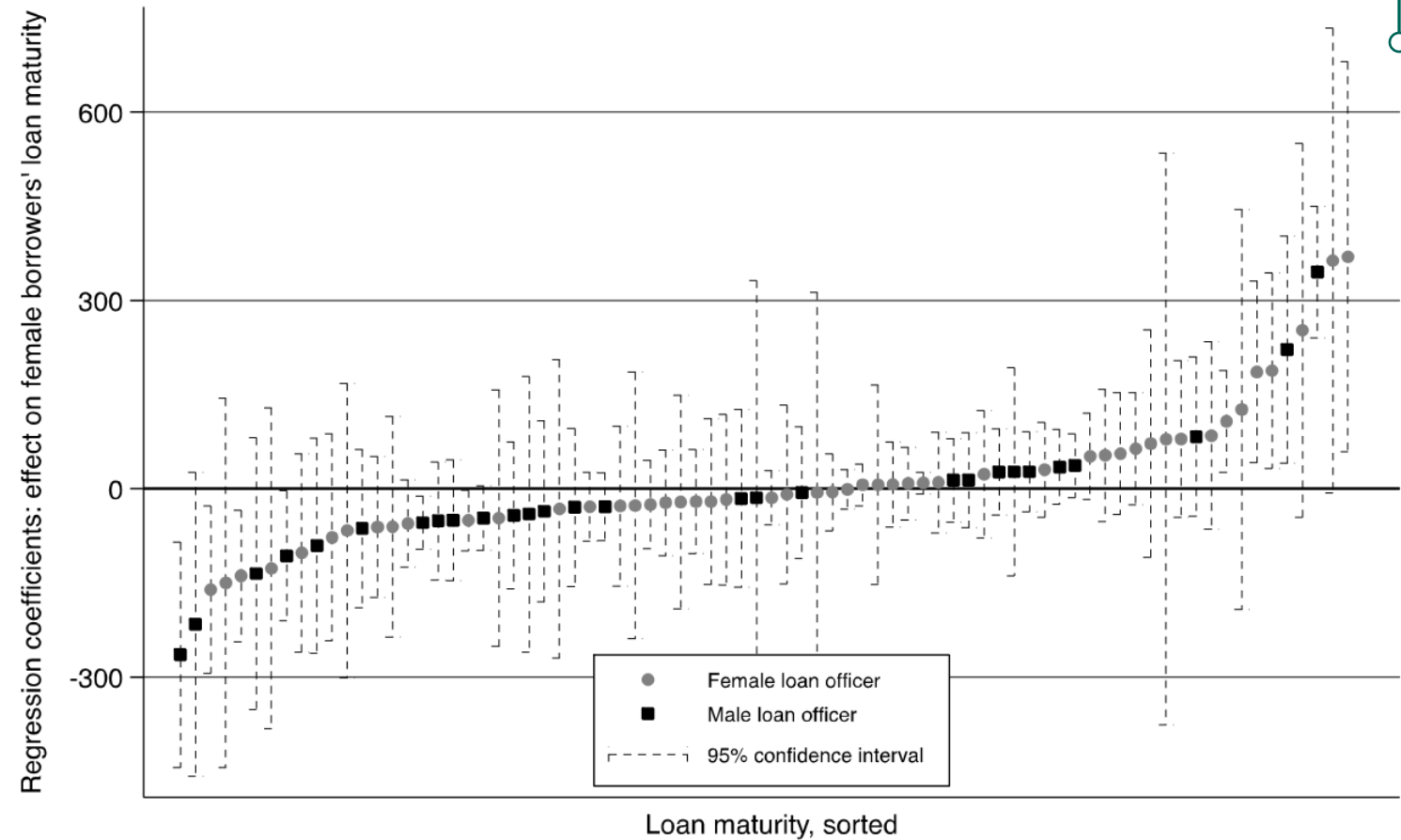
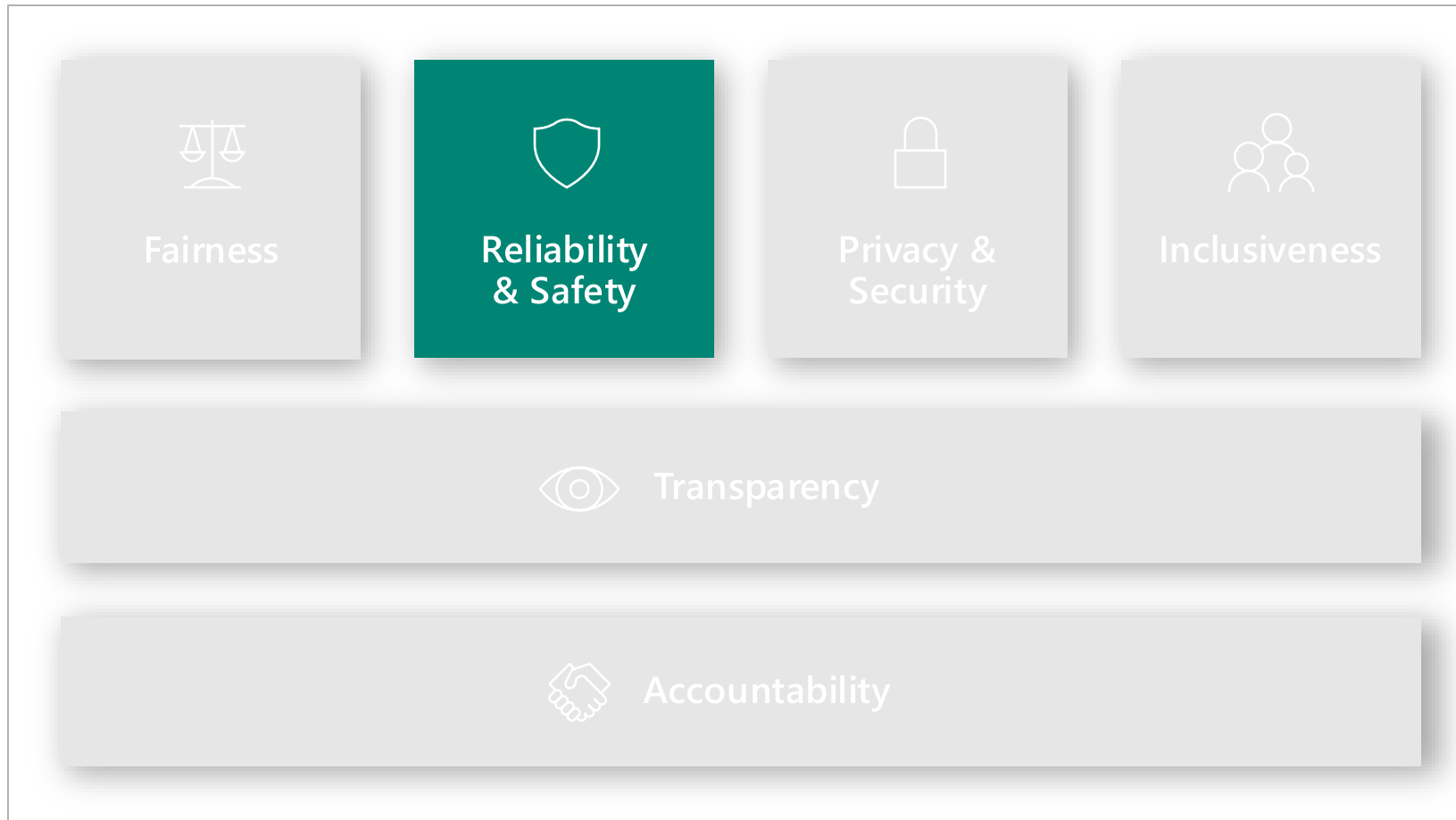


Figure 3. The figure shows the distribution of the bias on loan maturity by officer gender. Each coefficient represents an estimate of the number of extra days of loan maturity an individual officer approves for female versus male borrowers.

Source: SSRN, Sex and Credit: Is there gender bias in lending?, Beck, Behr & Madestom, April 2017

Microsoft's AI principles



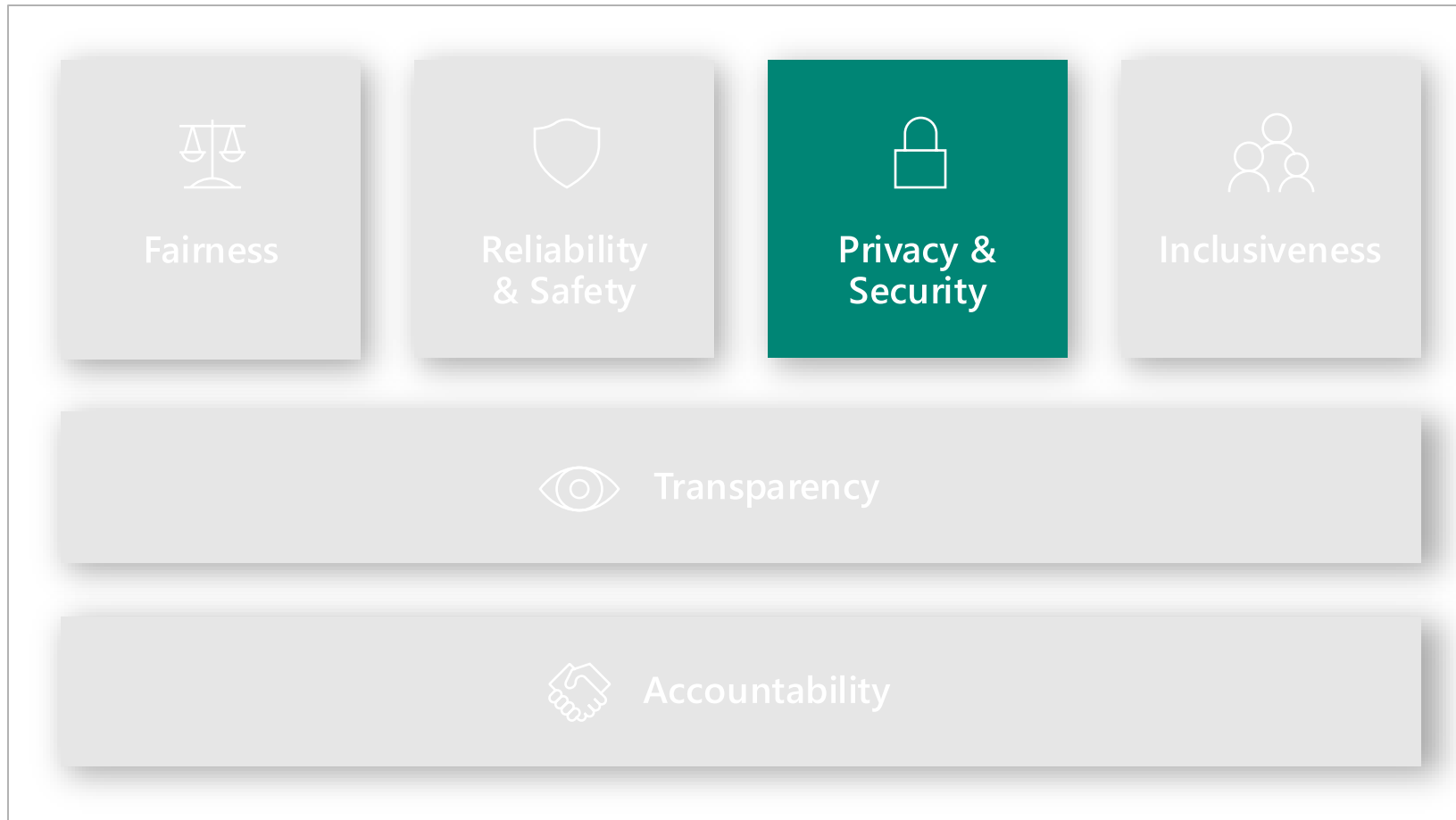
Reliability & Safety

Miles & years needed for to demonstrate autonomous vehicle reliability

Benchmark Failure Rate				
Statistical Question	How many miles (years*) would autonomous vehicles have to be driven...	(A) 1.09 fatalities per 100 million miles?	(B) 77 reported injuries per 100 million miles?	(C) 190 reported crashes per 100 million miles?
	(1) Without failure to demonstrate with 95% confidence their failure	275 million miles (12.5 years)	3.9 million miles (2 months)	1.6 million miles (1 month)
	(2) To demonstrate 95% confidence their failure to within 20% of the true rate of...	8.8 billion miles (400 years)	125 million miles (5.7 years)	51 million miles (2.3 years)
	(3) To demonstrate with 95% confidence and 80% power that their failure rate is 20% better than the human driver failure rate of...	11 billion miles (500 years)	161 million miles (7.3 years)	65 million miles (3 years)
*We assess the time it would take to complete the requisite miles with a fleet of 100 autonomous vehicles (larger than any known existing fleet) driving 24 hours a day, 365 days a year, at an average speed of 25 miles per hour.				

Source: Rand Corp. *Driving to Safety*; Kara & Paddock

Microsoft's AI principles



Privacy & Security

≡ Forbes



3,443,441 views | Feb 16, 2012, 11:02am

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did



“ We knew that if we could identify them in their second trimester, there’s a good chance we could capture them for years

”

Andrew Pole, Statistician, Target

Privacy & Security



Privacy & Security



news

Top Stories

Local

The National

Opinion

World

Canada

Calgary

Company suspends use of mall directory cameras running facial recognition software

Cadillac Fairview had been testing the technology since June, but didn't tell mall patrons

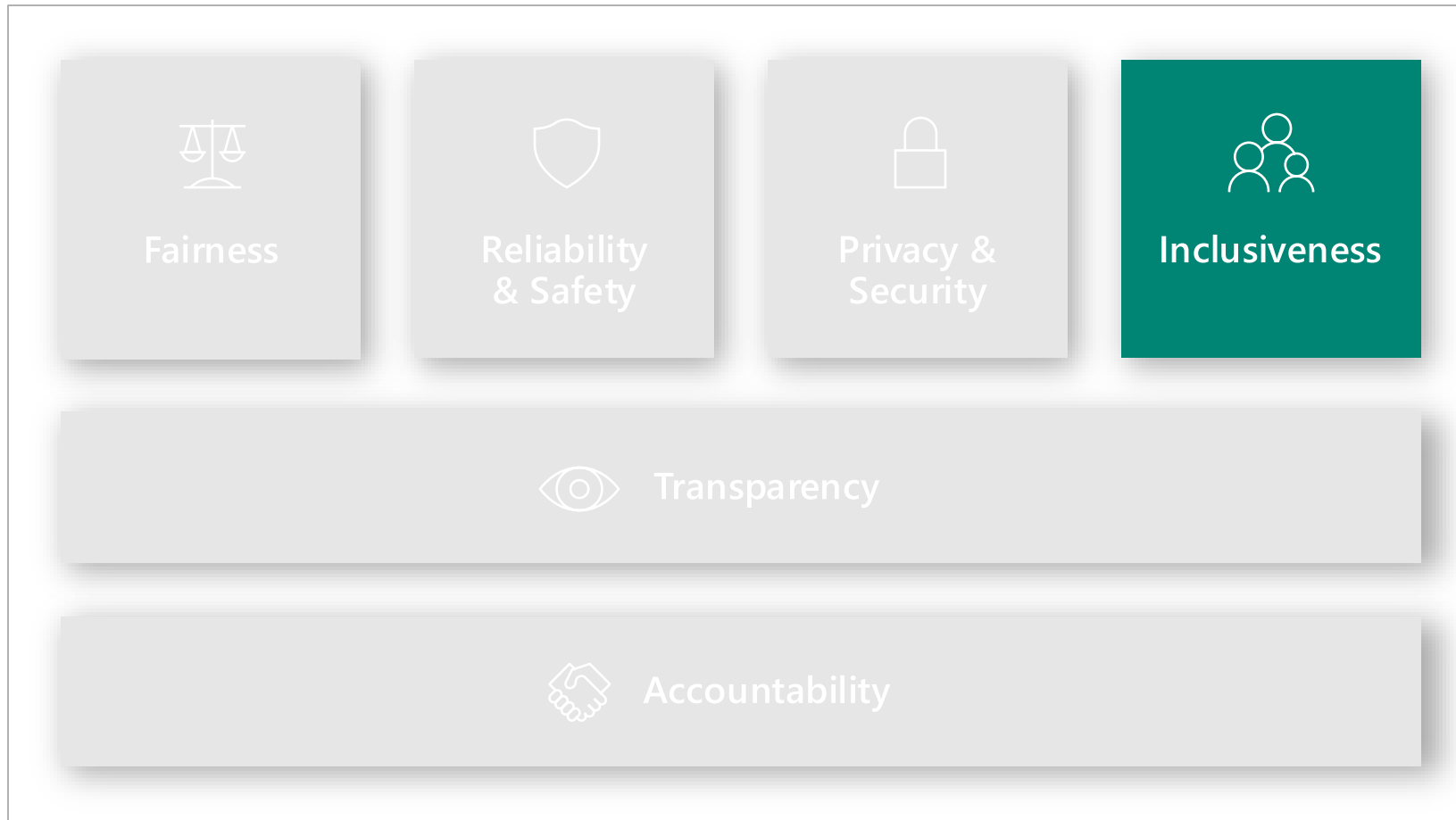


Anis Heydari · CBC News · Posted: Aug 04, 2018 6:08 PM MT | Last Updated: August 4, 2018



This mall directory at Cadillac Fairview's Chinook Centre in Calgary has a camera embedded within it, as circled in red on the left. (Anis Heydari/CBC)

Microsoft's AI principles



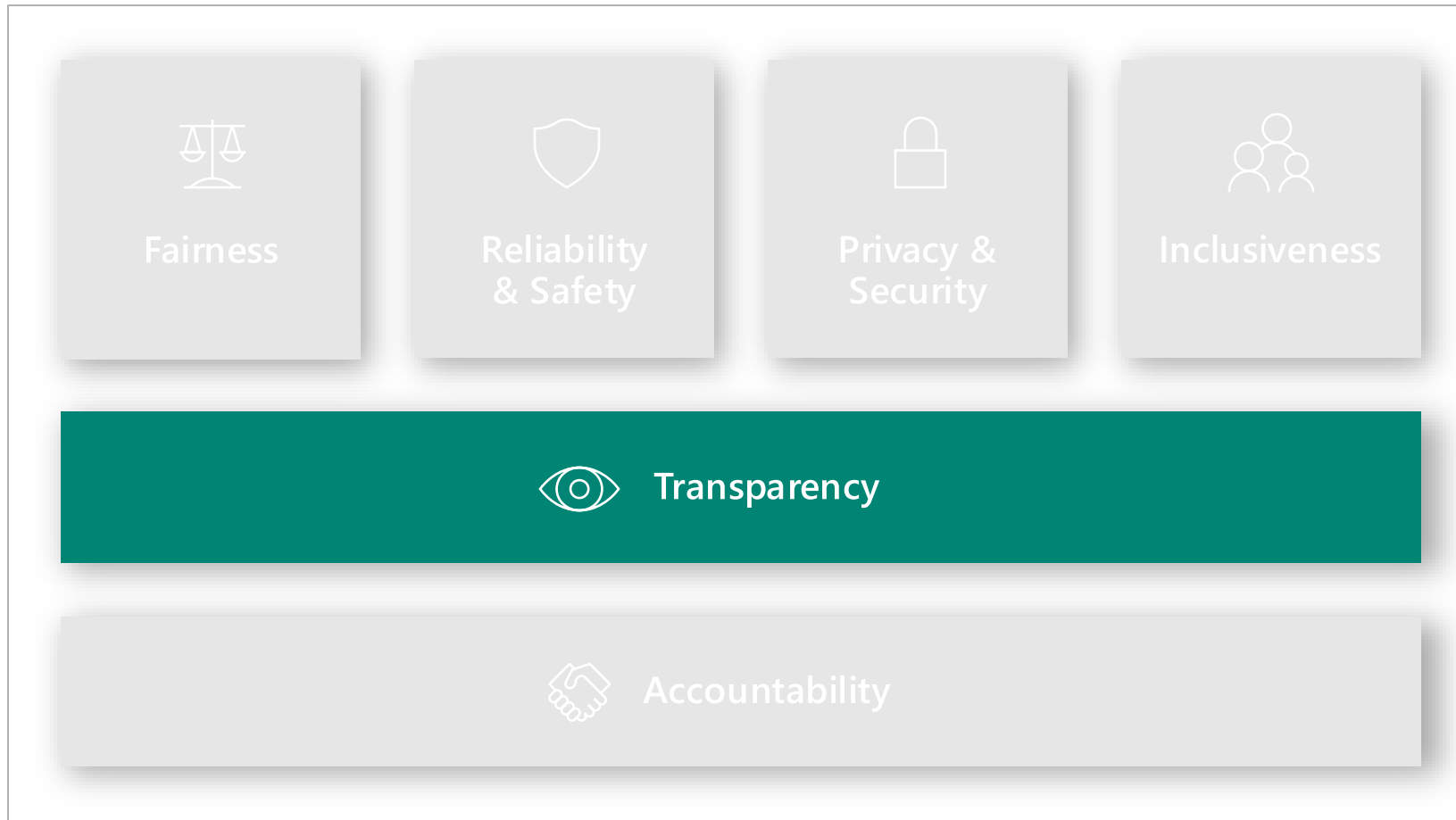
Inclusiveness



K. W. v. Armstrong plaintiff Christie Mathwig

https://www.acluidaho.org/sites/default/files/field_documents/class_action_complaint.pdf

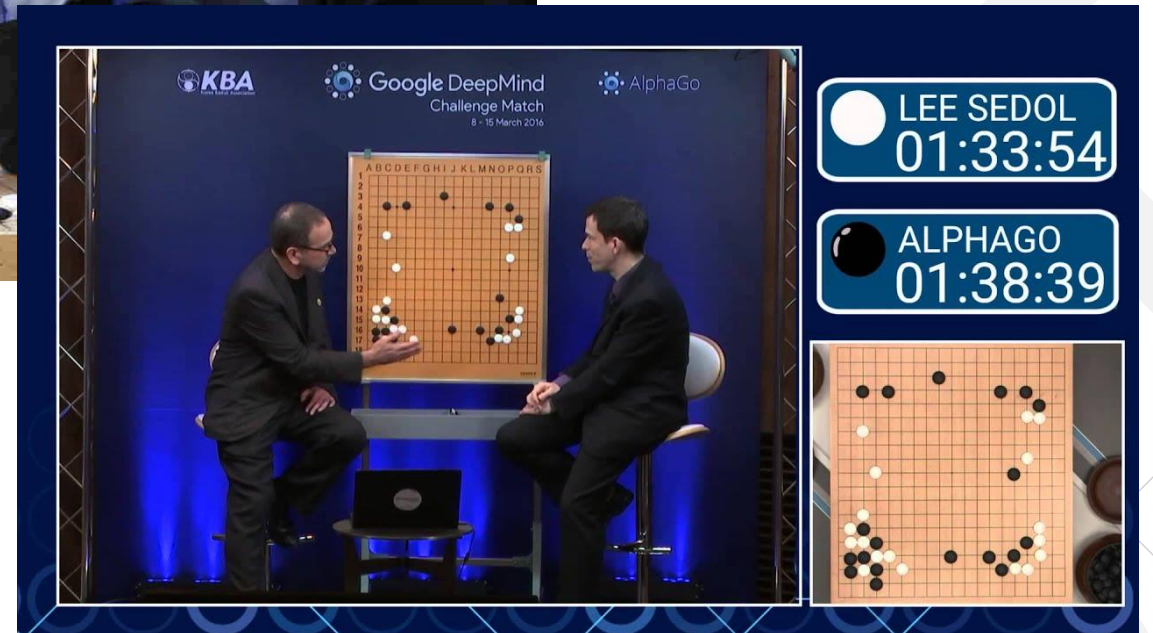
Microsoft's AI principles



Transparency



Move 37



Transparency

THE VERGE

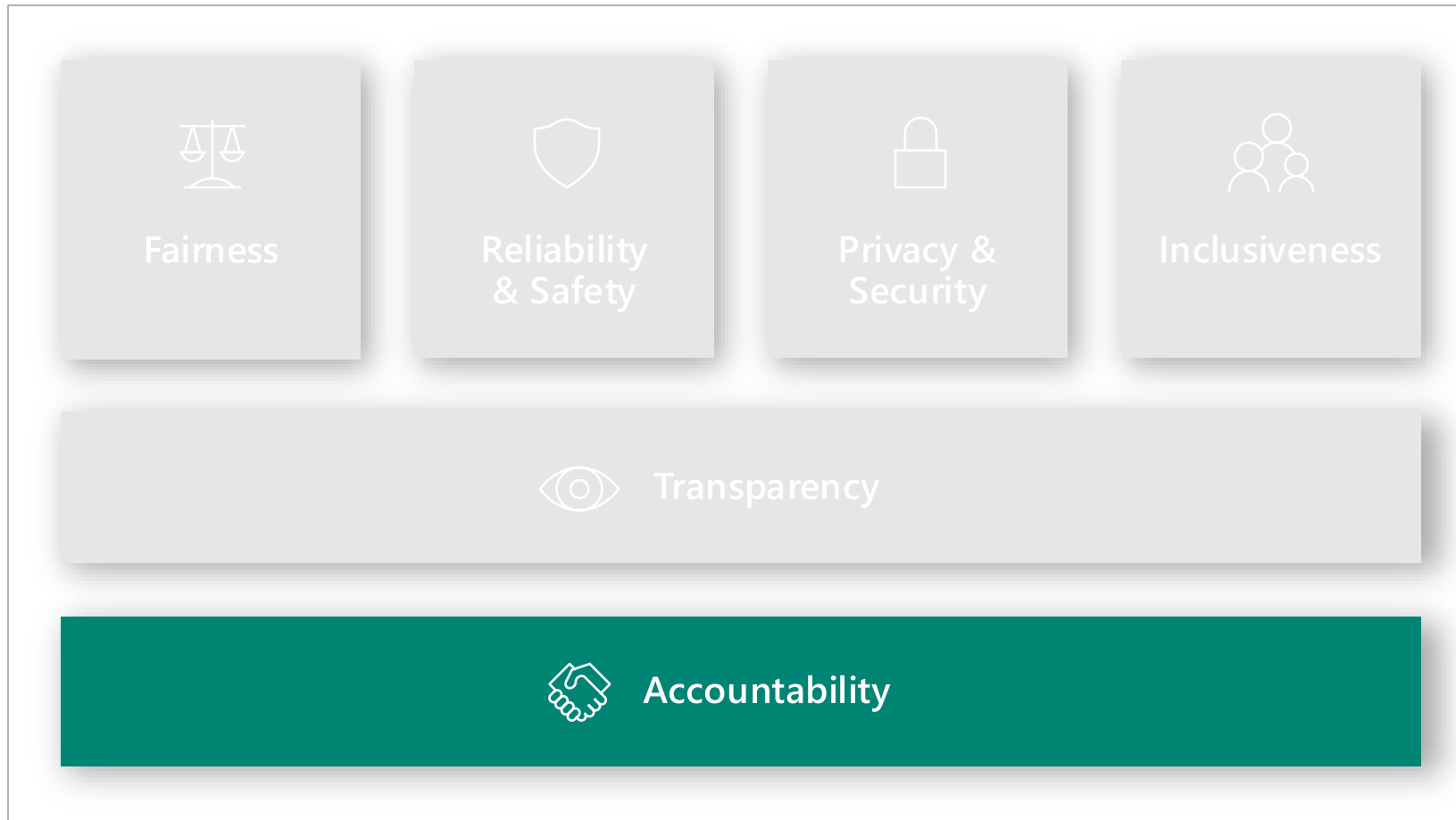
POLICY & LAW

PALANTIR HAS SECRETLY BEEN USING NEW ORLEANS TO TEST ITS PREDICTIVE POLICING TECHNOLOGY

Palantir deployed a predictive policing system in New Orleans that even city council members don't know about

By [Ali Winston](#) | Feb 27, 2018, 3:25pm EST

Microsoft's AI principles



Accountability

MIT Technology Review

Ethical Tech / AI Ethics

When algorithms mess up, the nearest human gets the blame

A look at historical case studies shows us how we handle the liability of automated systems.

by Karen Hao

May 28, 2019



Accountability

UK Official Says It's Too Expensive to Delete All the Mugshots of Innocent People in Police Databases



Sidney Fussell

4/19/18 2:30pm • Filed to: SURVEILLANCE ✓



16



4



A police officer watches a television monitor displaying a fraction of London's CCTV camera network

Photo: Daniel Berehulak (Getty)

Agenda _Responsible AI



Why

What

How

Putting Responsible AI into Practice



Responsible AI & Human Rights



Dignity of every individual



Freedom from discrimination



Freedom from invasions of privacy



Freedom of expression



Freedom of association



THE UNIVERSAL DECLARATION OF Human Rights



Article 1 — All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.

Article 2 — Everyone is entitled to all the rights and freedoms set forth in this Declaration, without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status.

Article 3 — Everyone has the right to life, liberty and security of person.

Article 4 — No one shall be held in slavery or servitude; slavery or servitude shall be prohibited in all its forms.

Article 5 — No one shall be subjected to torture or to cruel, inhuman or degrading treatment or punishment.

Article 6 — Everyone has the right to recognition as a person before the law.

Article 7 — All are equal before the law and are entitled without any discrimination to equal protection of the law. All are entitled to equal treatment of the law.

Article 8 — Everyone has the right to an effective remedy by the competent national or international authorities.

Article 9 — No one shall be subjected to arbitrary arrest, detention or exile.

Article 10 — Everyone is entitled in full equality to a fair and public hearing by an independent and impartial tribunal, in the determination of his rights and obligations.

Article 11 — 1. Everyone has the right to a fair trial. 2. No one shall be held guilty of any crime until he has been proved guilty by the competent judicial authority.

Article 12 — No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.

Article 13 — 1. Everyone has the right to freedom of movement and residence within the borders of each country. 2. Everyone has the right to leave any country, including his own, and to return to his country.

Article 14 — 1. Everyone has the right to seek and to enjoy in other countries asylum from persecution. 2. No one shall be subjected to restrictions of movement without lawful reasons.

Article 15 — 1. Everyone has the right to a nationality. 2. No one shall be arbitrarily deprived of his nationality nor denied the right to change his nationality.

Article 16 — 1. Everyone has the right to marry and to found a family, which is the basis of the society. 2. No restrictions shall be placed on the right to marry other than the freely expressed intention of the individuals. 3. The right to marry shall be exercisable only when the individuals of full age and capacity have reached the age of majority.

Article 17 — 1. Everyone has the right to own property alone as well as in association with others. 2. No one shall be arbitrarily deprived of his property.

Article 18 — Everyone has the right to freedom of thought, conscience and religion; this right includes freedom to change his religion or belief.

Article 19 — Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without restriction and to receive and impart information and ideas without restriction by border.

Article 20 — 1. Everyone has the right to peaceful assembly and to associate with others in the person or in the community. 2. No restrictions shall be placed on the right of assembly or association.

Article 21 — 1. Everyone has the right to take part in the government of his country, directly or through freely chosen representatives. 2. Everyone has the right of equal access to public service in his country. 3. The will of the people shall be the basis of the authority of government; this will shall be expressed in periodic and genuine elections which shall be by secret ballot and shall ensure the free expression of the will of the voters.

Article 22 — Everyone has the right to social security and, as a last resort, to the realization of the economic, social and cultural rights set forth in this Declaration.

Article 23 — 1. Everyone has the right to work, to free choice of employment, to just and favourable conditions of work and to protection against unemployment. 2. Everyone, without any discrimination, has the right to equal pay for equal work. 3. Everyone has the right to a fair and favourable remuneration which shall include wages and benefits of social security. 4. Everyone has the right to an adequate standard of living, which includes food, clothing, housing, medical care and social services, and the right to security of employment.

Article 24 — Everyone has the right to rest and leisure, including reasonable limitation of working hours and periodic holidays with pay.

Article 25 — Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing, medical care and social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in accordance with national and international law.

Article 26 — 1. Everyone has the right to education. 2. Education shall be free, at least in the elementary and fundamental stages. Elementary education shall be compulsory. Technical, vocational and higher education shall be made generally accessible and, where possible, free. 3. Education shall be directed to the full development of the human personality and to the strengthening of respect for human rights and fundamental freedoms. It shall promote understanding, tolerance and friendship between nations, racial or ethnic groups and religions. 4. Parents have the primary responsibility for the education of their children.

Article 27 — 1. Everyone has the right to the free enjoyment of his own material and intellectual property. 2. No one shall be arbitrarily deprived of his property.

Article 28 — Everyone is entitled to a social and international order in which the rights and freedoms set forth in this Declaration can be fully realized.

Article 29 — 1. Everyone has duties to the community in which alone the free and full development of his personality is possible. 2. These rights and freedoms shall be exercised by everyone in conformity with the laws, regulations, customs and moral principles of the country and of the community to which he belongs. 3. These rights and freedoms shall be exercised by everyone in conformity with the laws, regulations, customs and moral principles of the country and of the community to which he belongs.

Article 30 — No state, group or individual may engage in any activity aimed at the destruction of any of the rights and freedoms set forth in this Declaration.

Microsoft Global Human Rights Statement

Information technology is transforming human experiences at an accelerating pace. As technology evolves, it touches lives in new ways. The statement explains how Microsoft is committed to respecting and promoting human rights, to ensure that technology plays a positive role across the globe.

Putting responsible AI into practice

Principles

Fairness
Accountability
Transparency
Inclusiveness
Reliability & Safety
Privacy & Security

Putting responsible AI into practice



Human-AI Guidelines

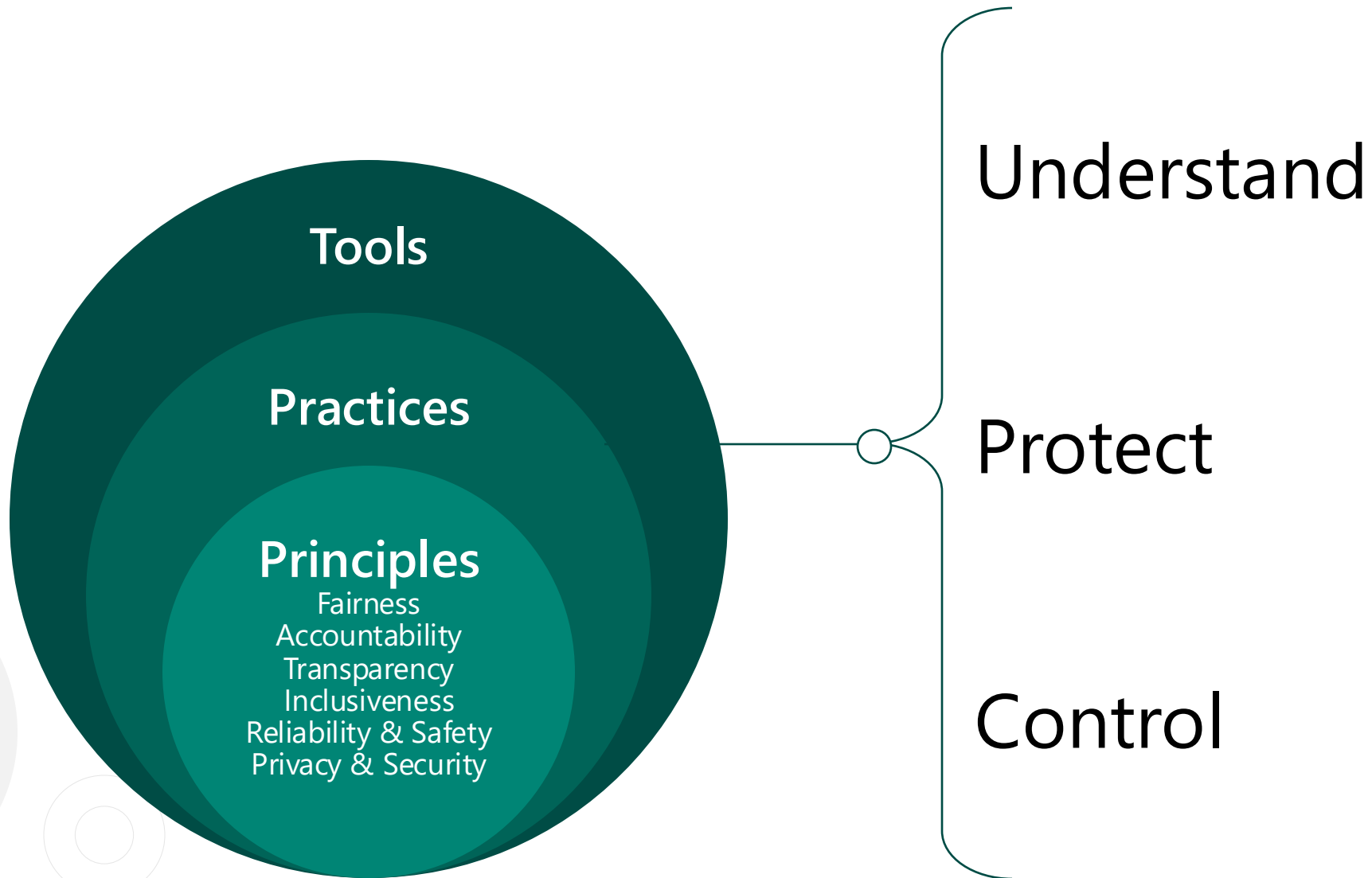
Conversational AI Guidelines

Inclusive Design Guidelines

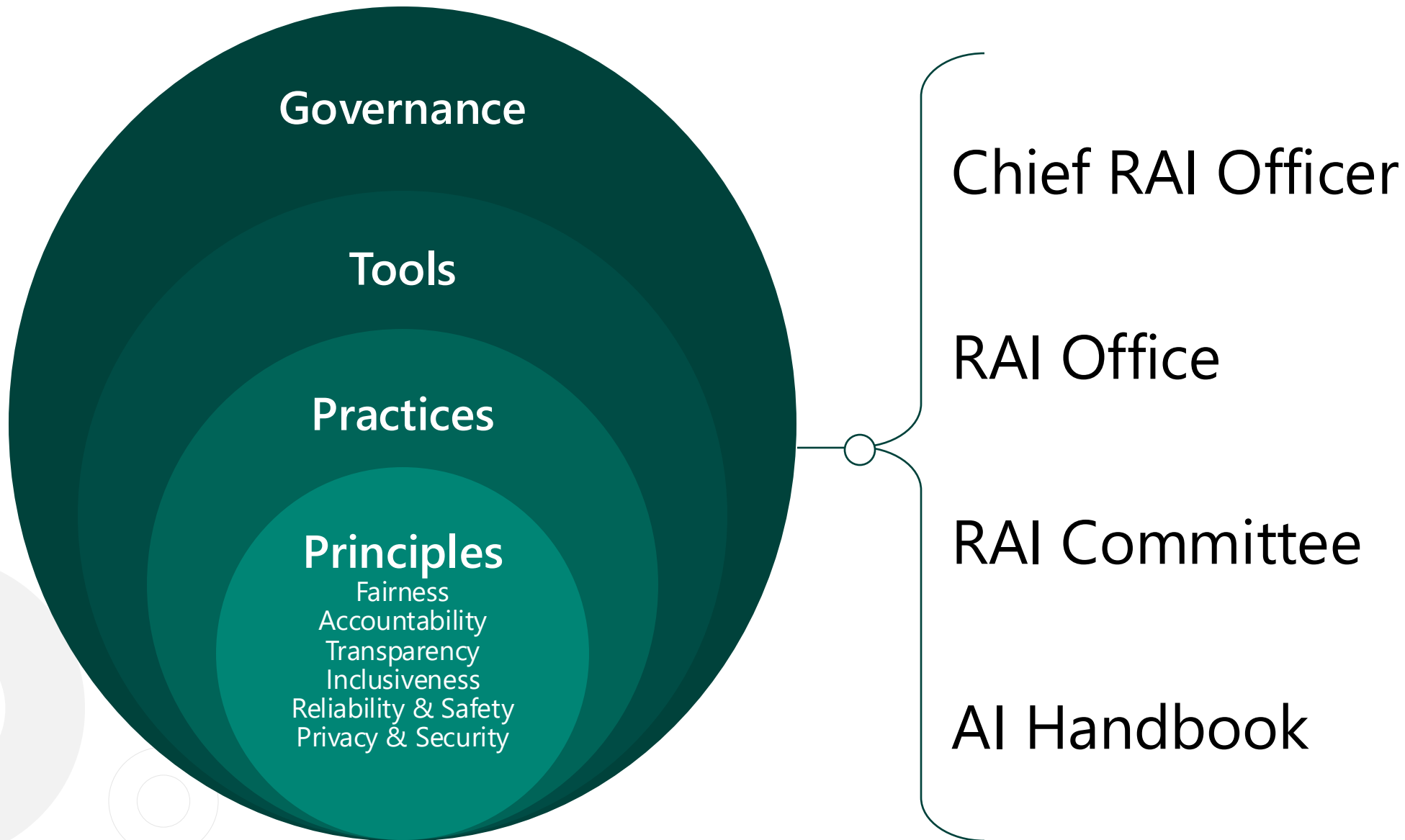
AI Fairness Checklist

Datasheets for Datasets

Putting responsible AI into practice



Putting responsible AI into practice



Responsible AI by Design

- ✓ AI is embedded in everyday life, business, government, medicine and more.
- ✓ **You** will be helping people and organizations adopt AI responsibly.
- ✓ Only by embedding ethical principles into AI applications and processes can we build systems based on trust.



Q&A

