Microsoft

Responsible AI at Microsoft

Charles Verdon
Data & AI Lead
Strategic Partnership Team, Asia

AI Transformation

Opportunities

Enrich employee experiences

Reinvent customer engagement

Reshape business processes

Bend the curve on innovation

Governance is necessary for AI adoption

Barriers to AI adoption

35%
mistakes/errors with real-world consequences

29%
general loss of trust due to bias, hallucinations and inaccuracies

Ambition

Organizations understand the importance of **governance, risk, and compliance** for responsible AI adoption.

vs

Activating

Building **cross-functional teams** to identify risks and define processes, determining executive sponsorship and accountability.

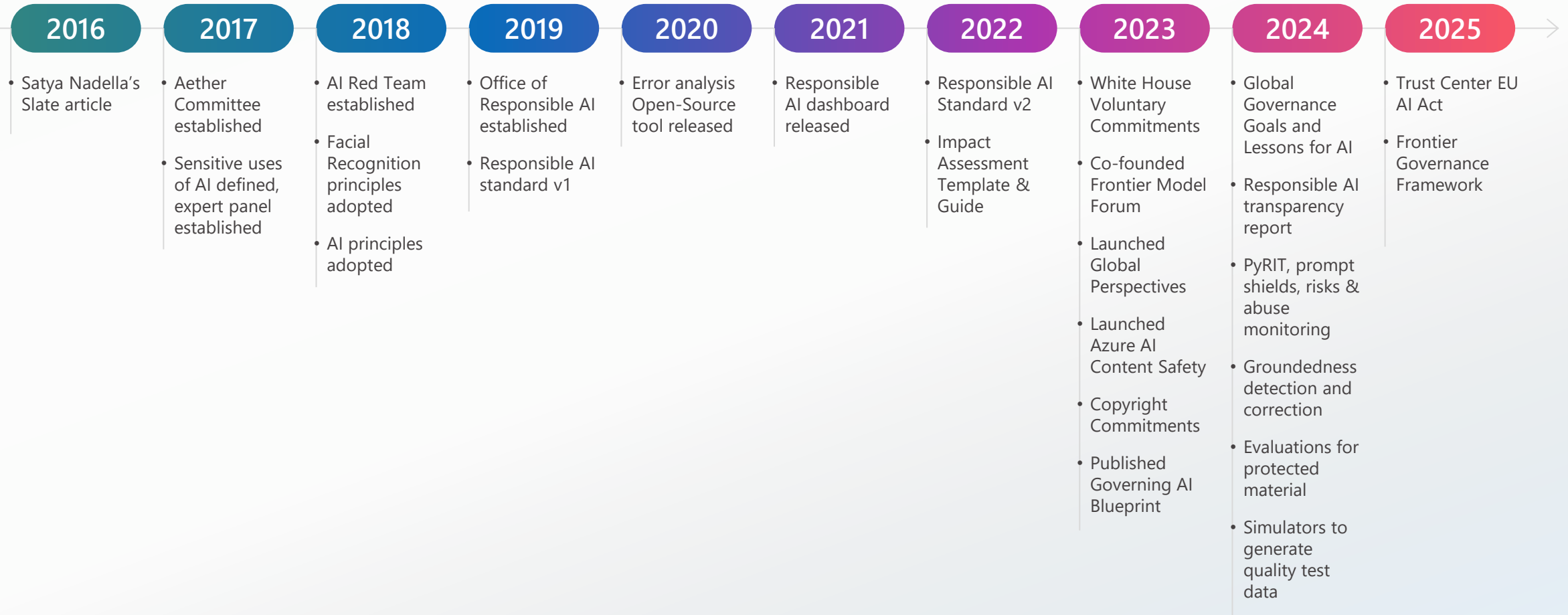## Barriers to AI development & deployment

Worried about regulatory compliance

Difficulty managing risks

Implementation challenges

Microsoft's commitment is
to build AI systems that are safe,
secure, and trustworthy **by design**

Our Responsible AI Journey

2016
- Satya Nadella's Slate article

2017
- Aether Committee established
- Sensitive uses of AI defined, expert panel established

2018
- AI Red Team established
- Facial Recognition principles adopted
- AI principles adopted

2019
- Office of Responsible AI established
- Responsible AI standard v1

2020
- Error analysis Open-Source tool released

2021
- Responsible AI dashboard released

2022
- Responsible AI Standard v2
- Impact Assessment Template & Guide

2023
- White House Voluntary Commitments
- Co-founded Frontier Model Forum
- Launched Global Perspectives
- Launched Azure AI Content Safety
- Copyright Commitments
- Published Governing AI Blueprint

2024
- Global Governance Goals and Lessons for AI
- Responsible AI transparency report
- PyRIT, prompt shields, risks & abuse monitoring
- Groundedness detection and correction
- Evaluations for protected material
- Simulators to generate quality test data

2025
- Trust Center EU AI Act
- Frontier Governance Framework

Microsoft's AI Principles

Fairness

Reliability & Safety

Privacy & Security

Inclusiveness

Transparency

Accountability

Microsoft's Responsible AI framework

Principles
Fairness • Privacy and security • Transparency
Reliability and safety • Inclusiveness • Accountability

Corporate Standard
Goals • Requirements • Practices

Implementation
Processes • Training • Tools

Oversight
Monitoring • Reporting • Auditing

How we build generative AI systems

We use a multi-layered approach to manage and mitigate risks for generative AI systems across the AI development lifecycle

Govern
Roles and responsibilities
Documented processes
Culture of risk management

Map
Impact assessment
Red teaming
Privacy and security reviews

Measure
Metrics for testing
Analysis of test results

Manage
User agency
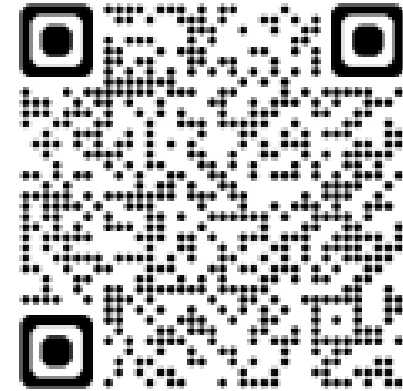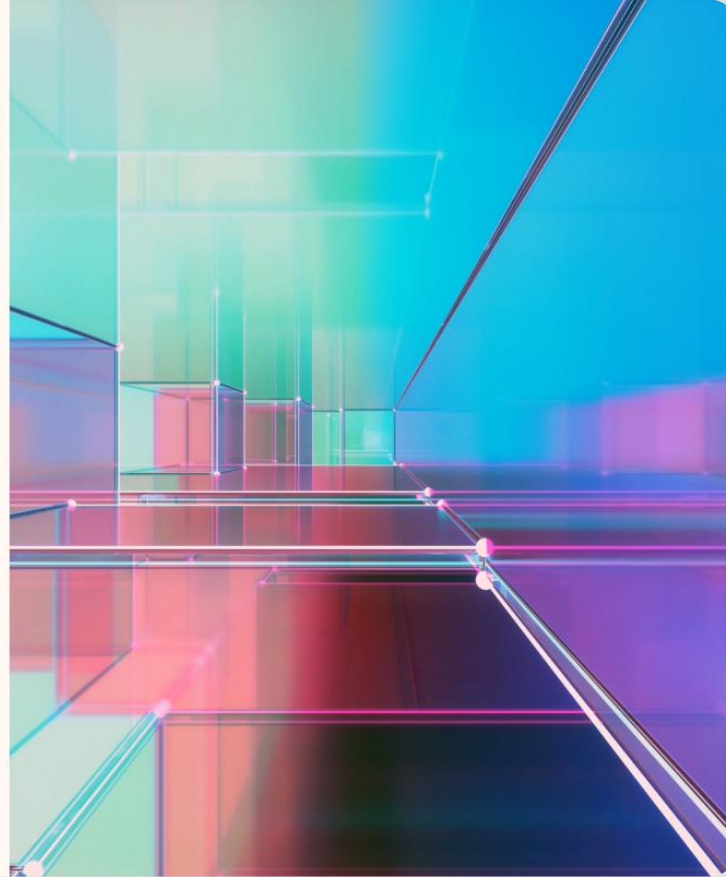Human oversight
Transparency
Ongoing monitoring

Responsible AI Transparency Report

Microsoft

Responsible AI
Transparency Report

How we build, support
our customers, and grow

May 2024

How we make decisions about releasing generative applications

We use a multi-layered approach to manage and mitigate risks for generative AI systems at every layer in our tech stack from the models to the applications they're deployed in.

Case study   **GitHub Copilot**

Govern
GitHub Copilot team works to map, measure and manage risks and complete the **RAI review process** with support from the **RAI Champ**

Map
RAI impact assessment, security review, and privacy review to map risks

Measure
Assessing performance, quality of responses, and conducting research to assess GitHub Copilot's **fitness for purpose**

Manage
Human oversight
Staying on topic
Transparency

Microsoft's existing privacy commitments apply to AI

We will keep your organization's data private

You are in control of your organization's data

Your access control and enterprise policies are maintained

Your organization's data is not shared

Your organization's data security and privacy are protected by design

Your organization's data is not used to train foundation models without your permission

Our products and solutions comply with global data protection regulations

Microsoft Responsible AI Partner Initiative

We are excited to foster a culture of collaboration and continuous learning within our partner ecosystem.

RAI Initiative partners have:

- Comprehensive AI practices
- Dedicated responsible AI resources
- Experience developing and deploying of responsible AI systems on Azure

Visit **aka.ms/RAI-Partner-Innovation**

How Microsoft is taking down AI hackers who create harmful images of celebrities and others

Thank you