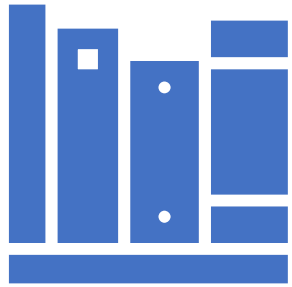
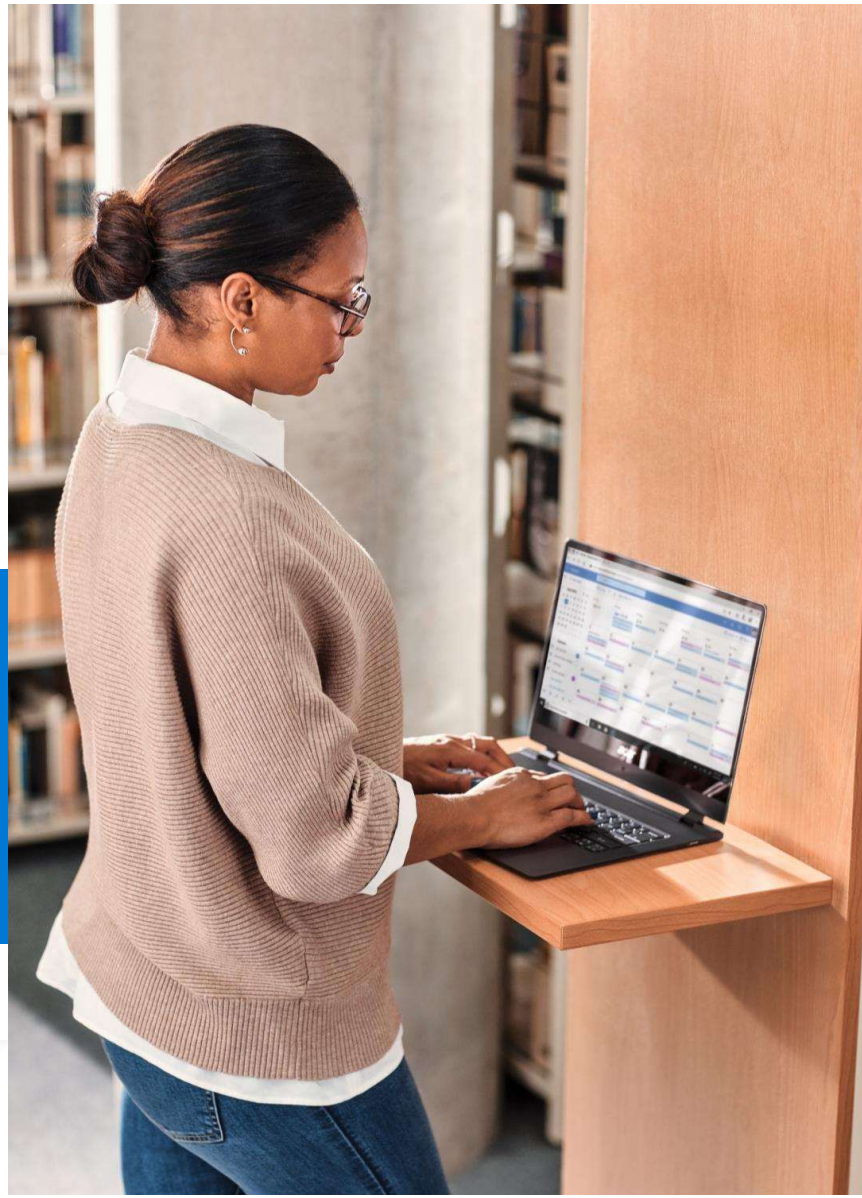


Responsible AI

Ruth Yakubu
Cloud Advocate
@RuthieYakubu

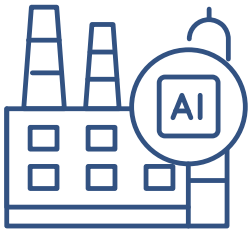
Agenda

- ❑ Why Responsible AI?
- ❑ Azure Responsible AI dashboard
- ❑ OpenAI Content Safety
- ❑ Microsoft 365 for Inclusion

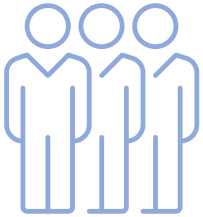


The world today…

AI **innovation** is occurring at a rapid pace



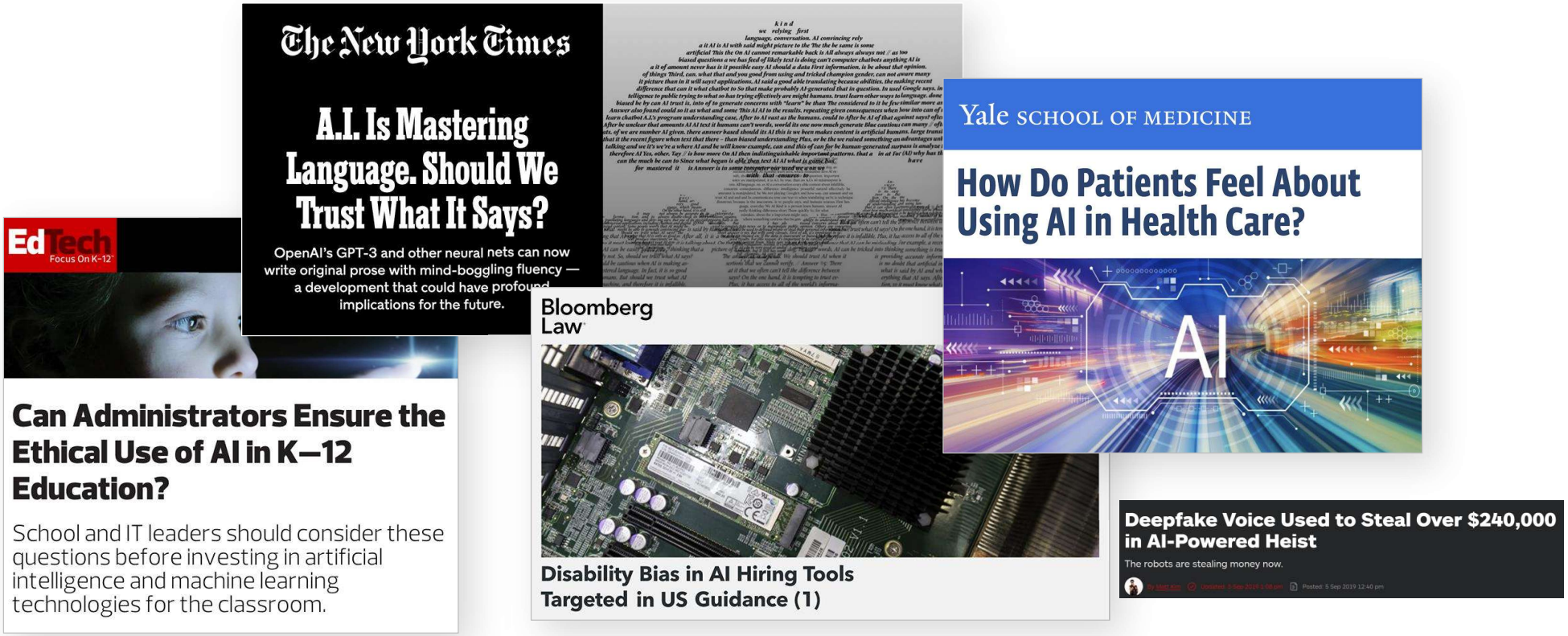
Companies are accelerating **adoption** of AI



Societal expectations are **evolving**

Governments are **regulating** AI in response

Societal expectations are increasing



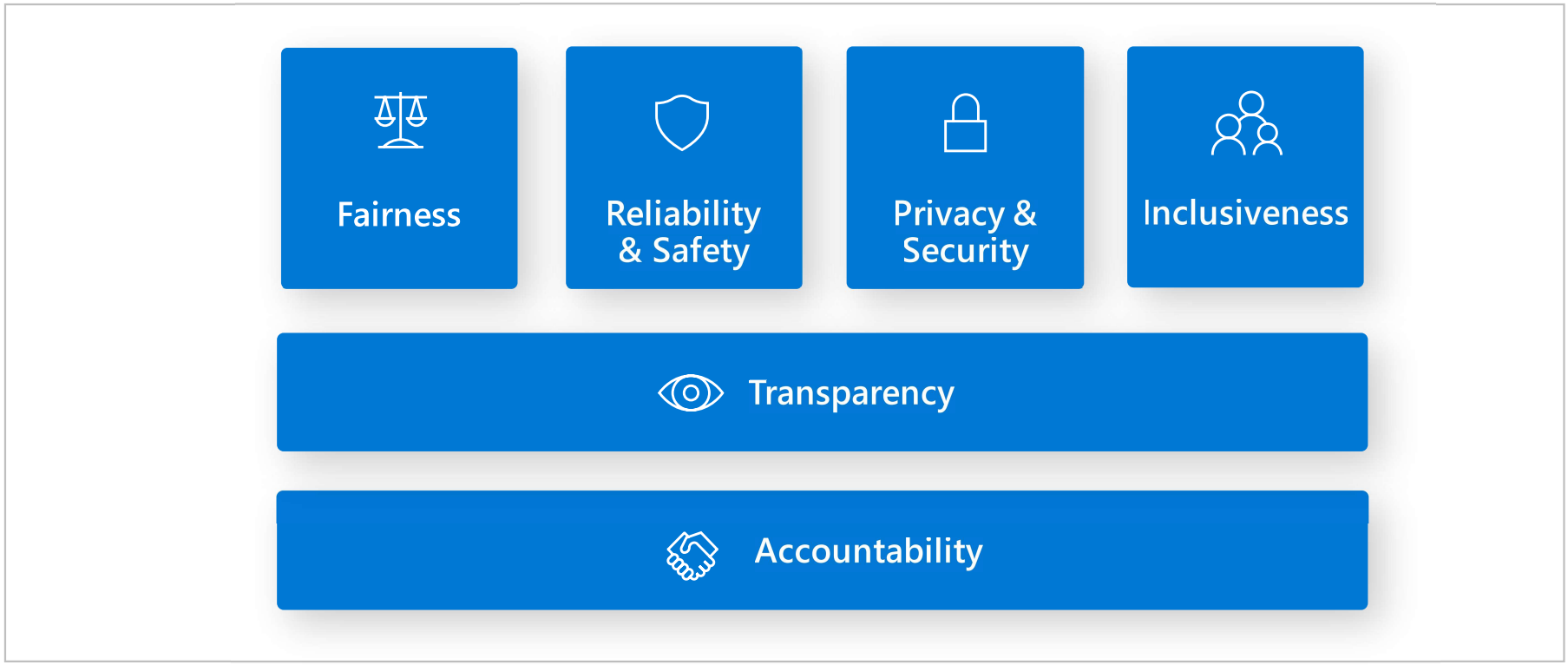
Why Responsible AI?

AI systems continue to miss expectations.

New AI breakthroughs expose new challenges.

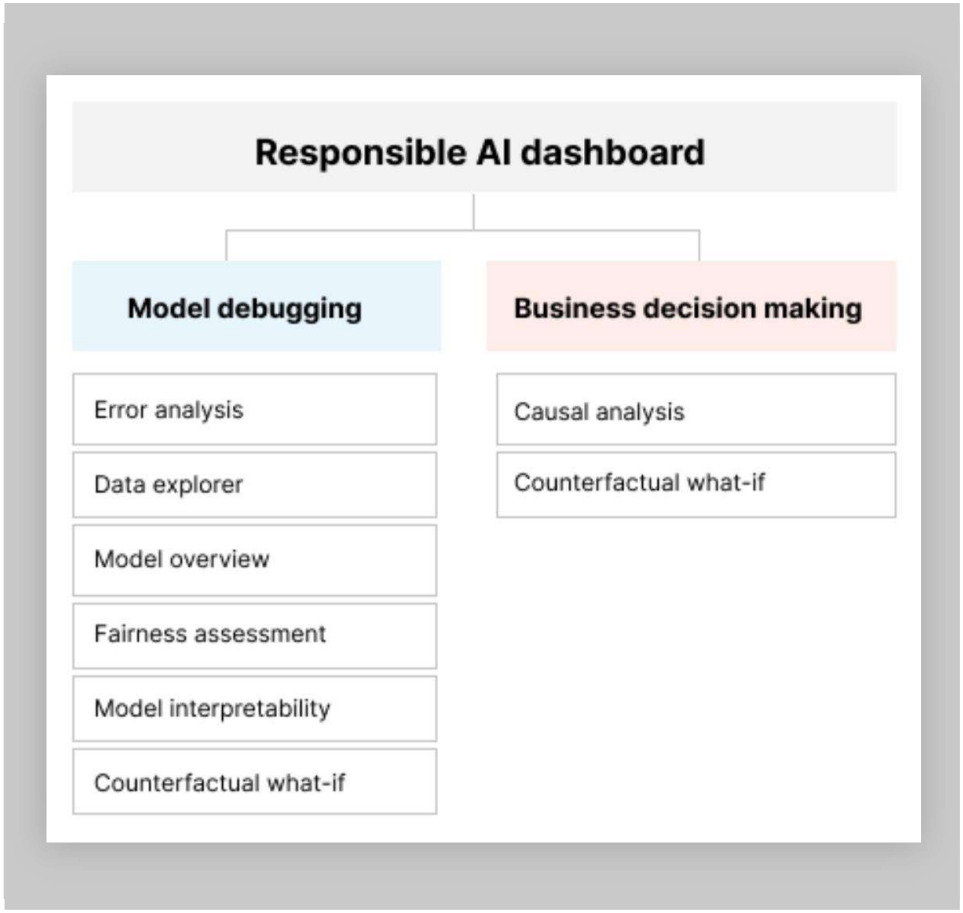
Government regulation on AI imminent.

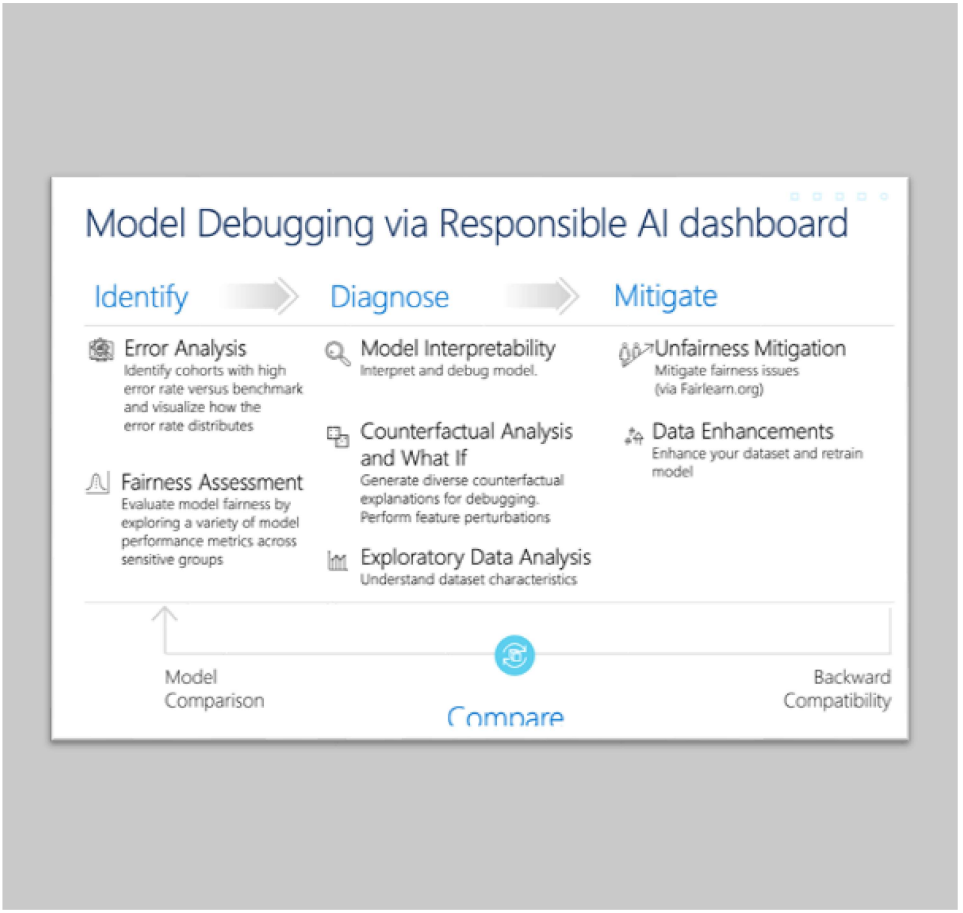
Learn Microsoft's AI principles



What does this mean to you?

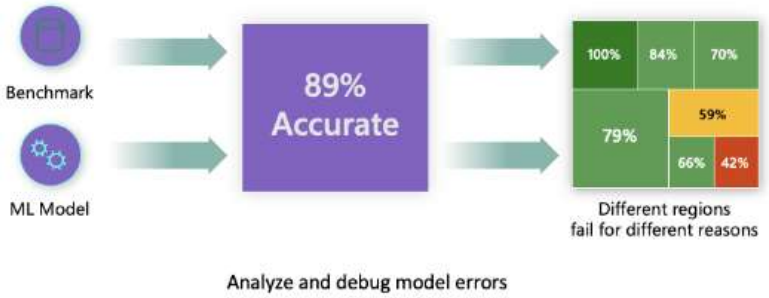
Model debugging with Responsible AI dashboard



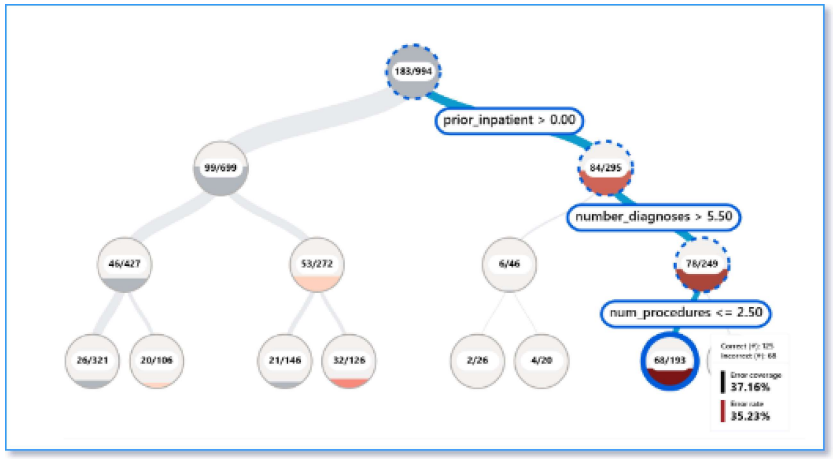


Gaps in model error identification

Error distribution

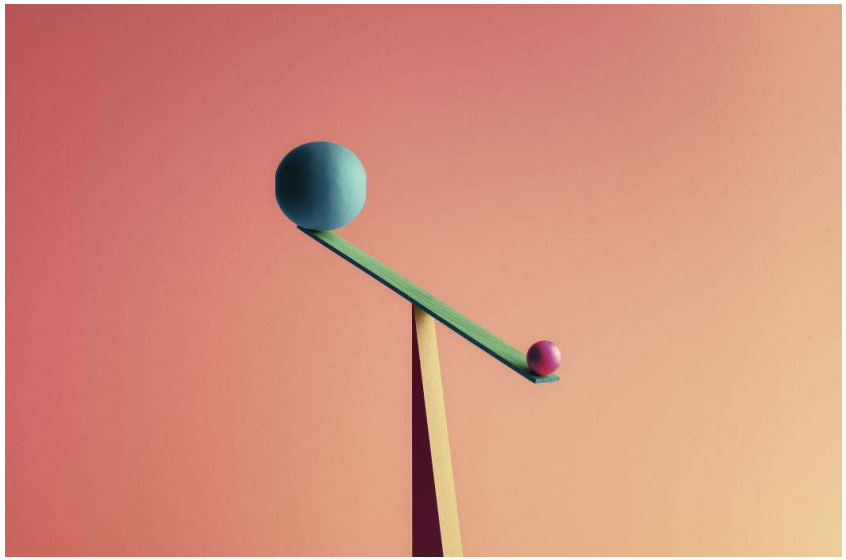


Error Analysis

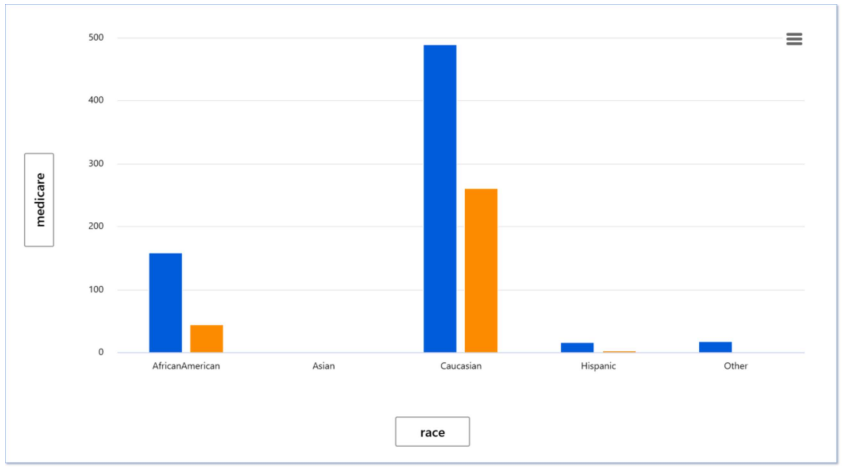


Gaps in finding data bias

DATA Representation

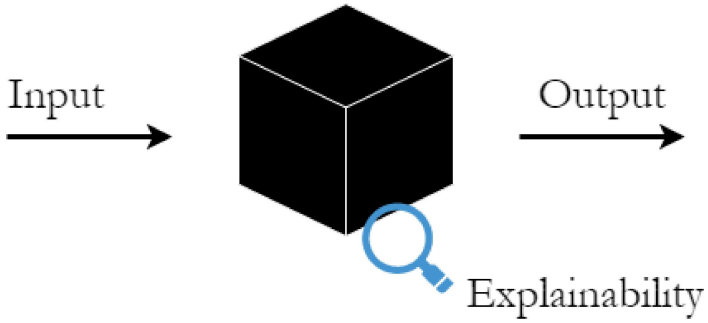


Data Analysis

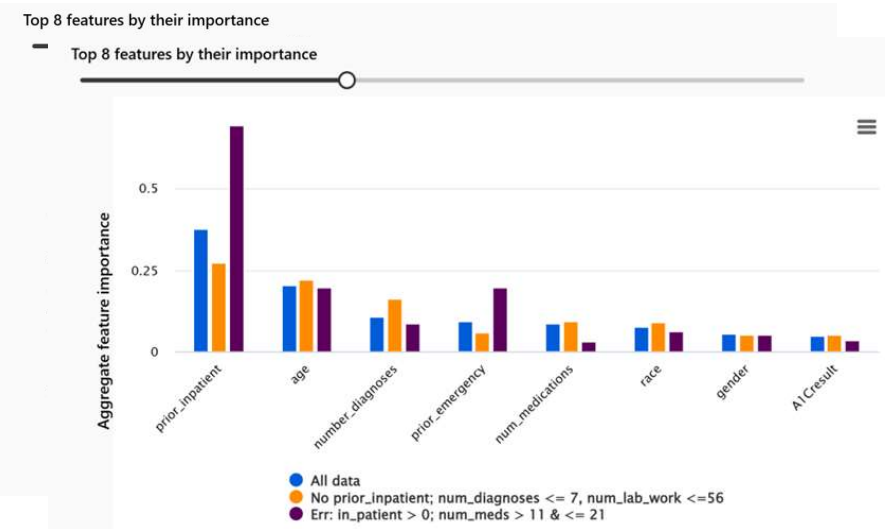


Gaps in model explainability

Blackbox



Feature Importance



Open AI Challenges to consider

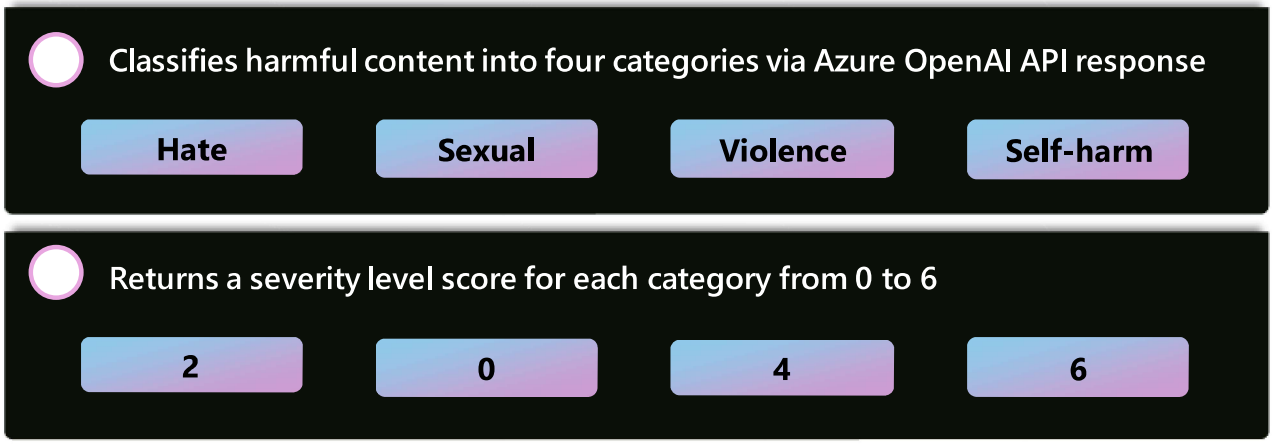
- Harmful to Society
- Bias
- Regulatory Compliance
- Content toxicity
- Relevant & Factual information
- Hallucination

Azure Content Safety

Content Filters

The Azure AI Content Safety service works alongside core models by running both the prompt and completion to detect and prevent the output of harmful content.

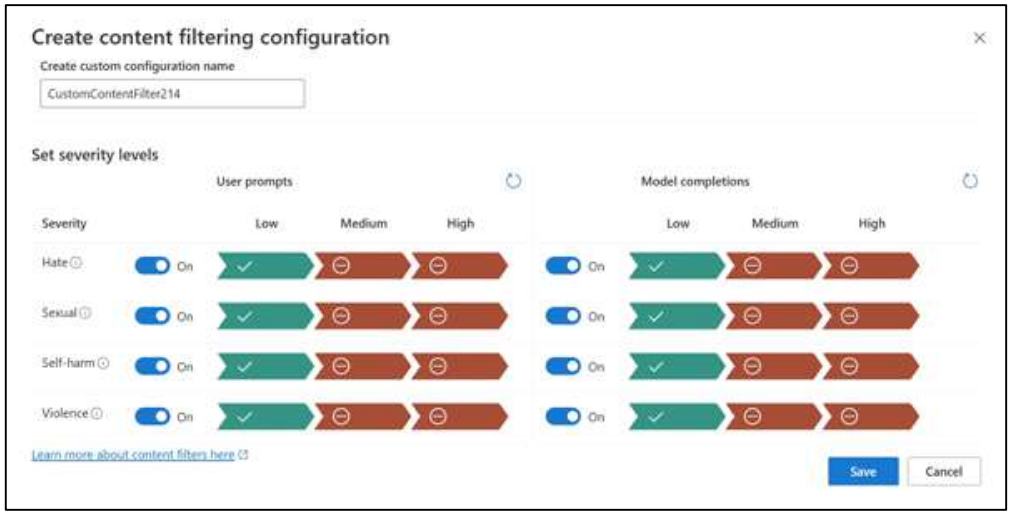
Supported languages: English, German, Japanese, Spanish, French, Italian, Portuguese, and Chinese



Configurable Azure OpenAI

Content Filters

Severity	Config for prompts	Config for completions	Description
Low, Medium, High	Yes	Yes	Strictest filtering configuration. Content detected at severity levels low, medium and high is filtered.
Medium, High	Yes	Yes	Default setting. Content detected at severity level low passes the filters, content at medium and high is filtered.
High	No	No	Content detected at severity levels low and medium passes the content filters. Only content at severity level high is filtered.

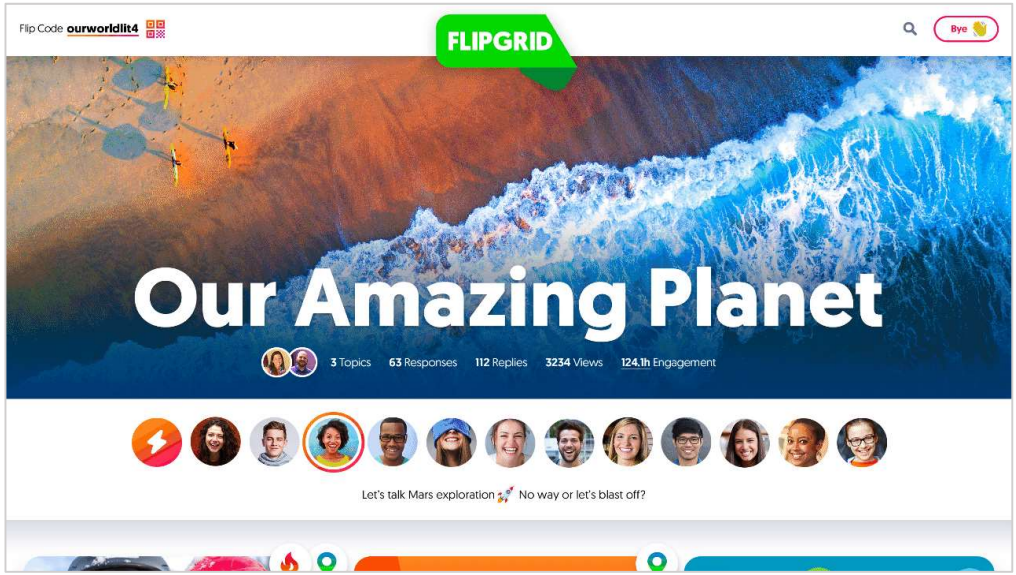


M365 for Education

- Inclusively designed to help people of all abilities with reading.
- Dyslexia, ADHD, emerging readers, non-native speakers, visual impairments
- Accessibility and learning tools, including Immersive Reader, translation, transcription, speech-to-text, and more
- Already used in Word, OneNote, Teams, Edge, PPT, SharePoint, Outlook, Forms, Flipgrid, Minecraft EDU and Lens.

Immersive Reader Capabilities

- Read Aloud and word/line highlighting (50 languages)
- Voice speed
- Reduce visual crowding
- Page colors tuned for accessibility
- Break words into syllables
- Highlight Parts of Speech (noun/verb/adj/adv)
- Line Focus
- Picture Dictionary
- Word or Document Translation (110 languages)

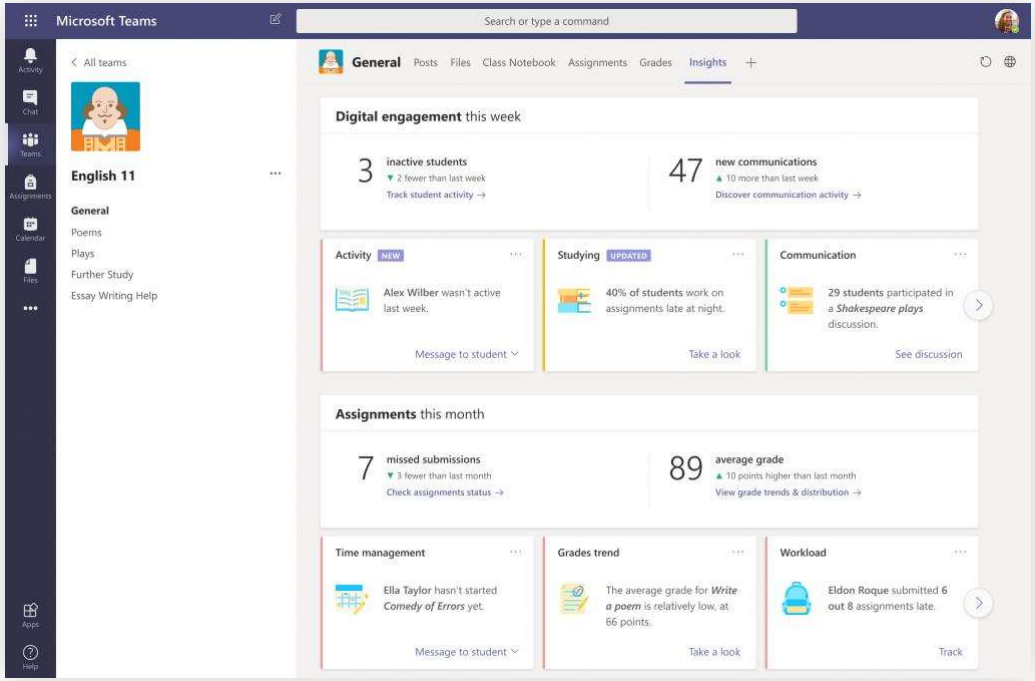


Example of the Immersive Reader in Flipgrid

Automation, AI, and analytics

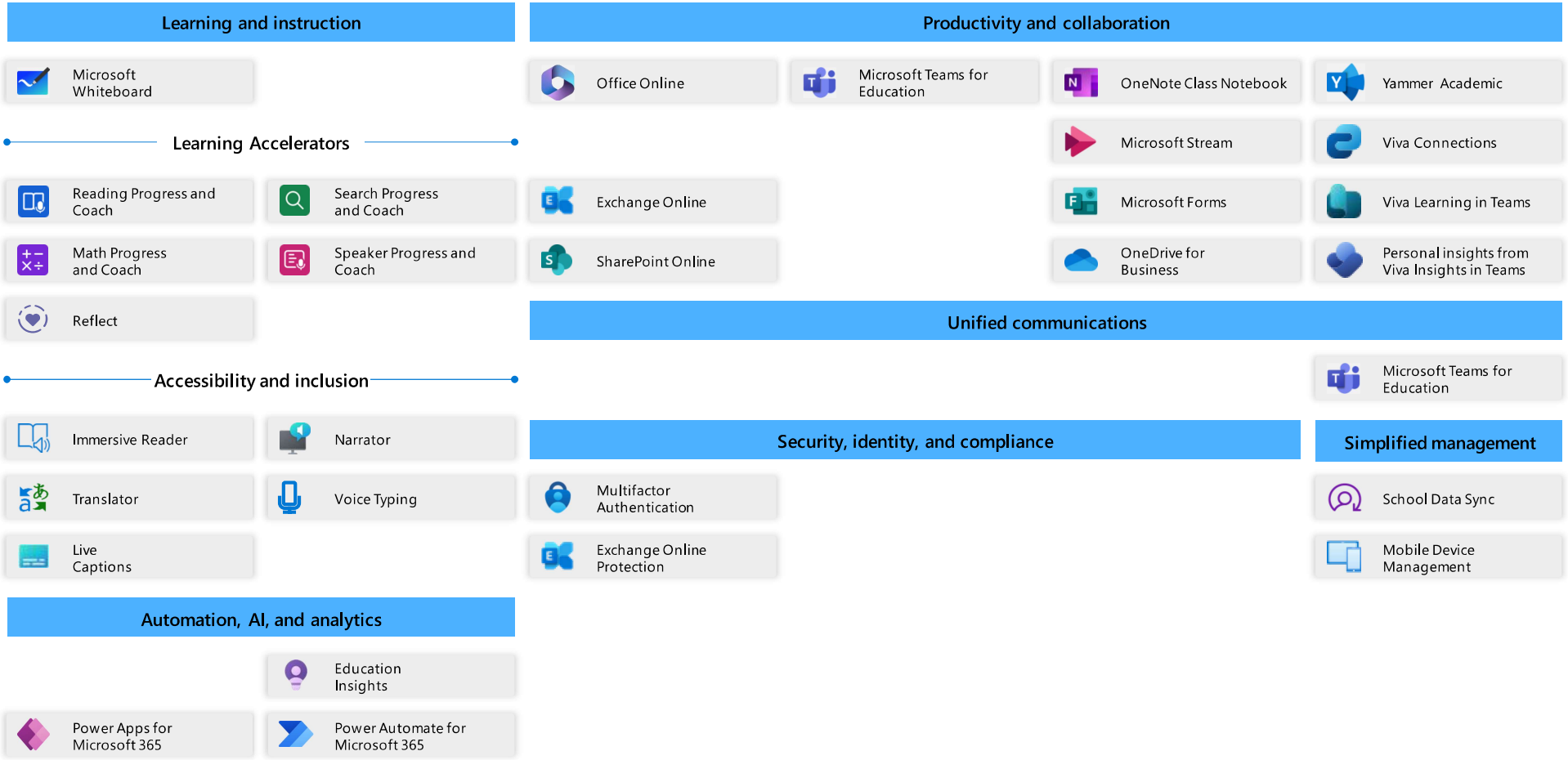
Education Insights

Educators and educational leaders have access to data and insights for individual students, multiple classrooms, and across the entire institution—to help improve learning outcomes



Education Insights

Microsoft 365 for Education



Resources

- Microsoft Learn Educator: https://aka.ms/Microsoft-Education-Center
- Azure Content Safety: https://aka.ms/Edu-ChatGPT-Content-Safety
- Responsible AI dashboard: https://aka.ms/Edu-ResponsibleAI-dashboard
- Pre-built AI: https://aka.ms/Edu-Prebuild-AI-Services

THANK YOU