



Our approach to responsible AI

AI is a defining constellation of technologies of our time and Microsoft is optimistic about what these technologies can do for people, industry, and society now and in the future. Around the world, Microsoft leaders and teams are working to ensure that our AI innovations and breakthroughs benefit people and society. We are committed to making sure AI systems are developed responsibly and in ways that warrant people's trust.

Putting principles into practice



In 2016, we led the industry in creating a set of human-centered principles to guide the ethical creation and use of AI. These principles evolved into the six outlined in [The Future Computed](#), released in 2018. These guide our end-to-end approach to AI, from its development to its deployment. Today, Microsoft is putting our principles into practice by embracing diverse perspectives, fostering continuous learning, and proactively responding as AI technology evolves.

We are operationalizing responsible AI at scale, across Microsoft. This is accomplished through Microsoft's AI, Ethics, and Effects in Engineering and Research (Aether) Committee, and our Office of Responsible AI. Aether is tasked with advising Microsoft's leadership around rising questions, challenges, and opportunities brought forth in the development and fielding of

AI innovations. The Office of Responsible AI implements our cross-company governance, enablement, and public policy work. Together, Aether and the Office of Responsible AI work closely with our engineering and sales teams to help them uphold Microsoft's AI principles in their day-to-day work. An important hallmark of our approach to responsible AI is having this ecosystem to operationalize responsible AI across the company, rather than a single organization or individual leading this work. Our approach to responsible AI also leverages our process of building privacy and security into all of our products and services from the start.

Office of Responsible AI

The Office of Responsible AI is tasked with putting Microsoft's principles into practice. It has four key functions:

Internal policy: Setting the company-wide rules for enacting responsible AI, as well as defining roles and responsibilities for teams involved in this effort.

Enablement: Readiness to adopt responsible AI practices, both within our company and among our customers and partners.

Case management: Review of sensitive use cases to help ensure that Microsoft's AI principles are upheld in our development and deployment work.

Public policy: Help to shape new laws, norms, and standards that will be needed to ensure that the promise of AI technology is realized for the benefit of society at large.

The Office of Responsible AI complements Aether's advisory function by administering, implementing, and maintaining the company's commitment to responsible AI governance.

Aether Committee

In 2016, we established the Aether Committee, which serves as an advisory role to the company's senior leadership on rising questions, challenges, and opportunities with the development and fielding of AI technologies.

Aether also provides guidance to teams across the company to ensure that AI products and services align with Microsoft's AI principles. The committee includes top talent in research, engineering, ethics, law, and policy from across Microsoft who come together to formulate recommendations on policies, processes, and best practices.

The Aether Committee commissions working groups on specific topics. Working groups to date include teams focused on sensitive uses of AI, bias and fairness of AI systems, intelligibility and explanation of AI reasoning and recommendations, reliability and safety, human-AI interaction, and engineering best practices.

Examples of efforts of the Aether committee and its working groups to date include its deliberation and input on Microsoft's decisions around sensitive uses of AI, such as applications of facial recognition, and its work on developing tools for detecting and addressing bias, recommended guidelines for human-AI interaction, and policies and methods for making AI recommendations more understandable.

Beyond guidance, Aether working groups have been points of organization, planning, engineering, and fielding of responsible AI technologies. Tools fielded to date include Data Sheets, Error Terrain Analysis, InterpretML, and FairLearn.