

Principles for the development and use of AI

Technological disruption and AI are about to bring vast changes, some of which are hard to imagine today. As with previous disruptive technological advances, we need to be thoughtful about how we address the societal issues that these changes bring about.

Against this premise, in [‘Future Computed - Artificial Intelligence and its role in society’](#), Microsoft outlines the potential benefits and risks of AI, and emphasizes that to be trustworthy, AI must be human-centered – designed in a way that augments human ingenuity and capabilities – and guided by ethical considerations and values.

Acknowledging its crucial role and responsibility as a leading developer of AI-based solutions and services, Microsoft developed a set of principles to guide the development and deployment of AI-powered solutions within the company, and suggests their adoption by governments, society and business in general.

These six principles are designed to ensure that AI systems are fair, reliable and safe, private and secure, inclusive, transparent, and accountable.

Understanding the importance of oversight and guidance, Microsoft’s AI and Ethics in Engineering and Research (AETHER) Committee was established to integrate these six principles. Working to define best practices, resolve questions of ethics and societal implications, AETHER is central to AI development at Microsoft.

An example of an important step in complying with the 6 principles of AI development is Microsoft’s understanding and compliance with regulations already in place that target data privacy and consumer safety such as the EU’s GDPR.



Microsoft’s AI-principles outlining the values AI needs to respect