

Google

AI Principles

Our approach to building beneficial AI



Our approach to developing and harnessing the potential of AI is grounded in our founding mission — to organize the world's information and make it universally accessible and useful. We believe [our approach](#) to AI must be both bold and responsible. Bold in rapidly innovating and deploying AI in groundbreaking products used by and benefiting people everywhere, contributing to scientific advances that deepen our understanding of the world, and helping humanity address its most pressing challenges and opportunities. And responsible in developing and deploying AI that addresses both user needs and broader responsibilities, while safeguarding user safety, security, and privacy.

We approach this work together, by collaborating with a broad range of partners to make breakthroughs and maximize the broad benefits of AI, while empowering others to build their own solutions.

Our approach to AI is grounded in these three principles:

1 **Bold innovation**

We develop AI that assists, empowers, and inspires people in almost every field of human endeavor; drives economic progress; and improves lives, enables scientific breakthroughs, and helps address humanity's biggest challenges. This means:

- Developing and deploying models and applications where the likely overall benefits substantially outweigh the foreseeable risks.
- Advancing the frontier of AI research and innovation through rigorous application of the scientific method, rapid iteration, and open inquiry.
- Using AI to accelerate scientific discovery and breakthroughs in areas like biology, medicine, chemistry, physics, and mathematics.
- Focusing on solving real world problems, measuring the tangible outcomes of our work, and making breakthroughs broadly available, enabling humanity to achieve its most ambitious and beneficial goals.

2 Responsible development and deployment

Because we understand that AI, as a still-emerging transformative technology, poses evolving complexities and risks, we pursue AI responsibly throughout the AI development and deployment lifecycle, from design to testing to deployment to iteration, learning as AI advances and uses evolve. This means:

- Implementing appropriate human oversight, due diligence, and feedback mechanisms to align with user goals, social responsibility, and widely accepted principles of international law and human rights.
- Investing in industry-leading approaches to advance safety and security research and benchmarks, pioneering technical solutions to address risks, and sharing our learnings with the ecosystem.
- Employing rigorous design, testing, monitoring, and safeguards to mitigate unintended or harmful outcomes and avoid unfair bias.
- Promoting privacy and security, and respecting intellectual property rights.

3 Collaborative progress, ~~together~~

We make tools that empower others to harness AI for individual and collective benefit. This means:

- Developing AI as a foundational technology capable of driving creativity, productivity, and innovation across a wide array of fields, and also as a tool that enables others to innovate boldly.
- Collaborating with researchers across industry and academia to make breakthroughs in AI, while engaging with governments and civil society to address challenges that can't be solved by any single stakeholder.
- Fostering and enabling a vibrant ecosystem that empowers others to build innovative tools and solutions and contribute to progress in the field.