



Why we focus on AI (and to what end)

It is an exciting time in the development of 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—and it is shaped by our commitment to improve the lives of as many people as possible. It is our view that AI is now, and more than ever, critical to delivering on that mission and commitment.

We believe that AI is a foundational and transformational technology that will provide compelling and helpful benefits to people and society through its capacity to assist, complement, empower, and inspire people in almost every field of human endeavor. It has the potential to contribute to tackling some of society’s most pressing challenges and opportunities, from the everyday to the more creative and imaginative.

As an information and computer science company, we aim to and have been at the forefront of advancing the frontier of AI through our path-breaking and field-defining research to develop more capable and useful AI. From this research and development, we are bringing breakthrough innovations into the real world to assist people and benefit society everywhere through our infrastructure, tools, products, and services, as well as through enabling and working with others to benefit society. We are also pursuing innovations that will help to unlock scientific discoveries and to tackle humanity’s greatest challenges and opportunities. Many of our innovations are already assisting and benefiting people (in some cases billions of people), communities, businesses, and organizations, and society broadly—with more such innovations still to come.

At the same time, we understand that AI, as a still-emerging technology, poses various and evolving complexities and risks. Our development and use of AI must address these risks. That’s why we as a company consider it an imperative to pursue AI responsibly. We are committed to leading and setting the standard in developing and shipping useful and beneficial applications, applying ethical principles grounded in human values, and evolving our approaches as we learn from research, experience, users, and the wider community.

We also believe that getting AI right—which to us involves innovating and delivering widely accessible benefits to people and society, while mitigating its risks— must be a collective effort involving us and others, including researchers, developers, users (individuals, businesses, and other organizations), governments, regulators, and citizens. It is critical that we collectively earn public trust if AI is to deliver on its potential for people and society. As a company, we embrace the opportunity to work with others to get AI right.

We are convinced that the AI-enabled innovations we are focused on developing and delivering boldly and responsibly are useful, compelling, and have the potential to assist and improve lives of people everywhere—this is what compels us. And we are excited about what lies ahead in 2023 and beyond as we get ready to share some new innovative experiences!

– James Manyika, Jeff Dean, Demis Hassabis, Marian Croak and Sundar Pichai

Our perspective, focus and principled approach (in 5 parts)

The following outlines how we think about and approach AI in five key parts. Each part will evolve as our innovations progress and as we learn more from research, experience, users, and the wider community.

1 Why we're developing AI

We believe that AI, including its core methods such as machine learning (ML), is a foundational and transformational technology. AI enables innovative new uses of tools, products, and services, and it is used by billions of people every day, as well as businesses, governments, and other organizations. AI can assist, complement, empower, and inspire people in almost every field, from everyday tasks to bold and imaginative endeavors. It can unlock new scientific discoveries and opportunities, and help tackle humanity's greatest challenges—today and in the future.

As [many have highlighted](#), we believe that AI has the potential to benefit people and society through its capacity to:

- Make information more useful and available to more people, everywhere, often helping overcome barriers including [access](#), disabilities and [language](#)
- Assist people and organizations to make decisions, solve problems, be more [productive](#) and creative in their daily and work lives
- Enable innovation that leads to [new, helpful products and services](#) for people, organizations, and society more broadly
- Help tackle current and [pressing real world challenges](#), such as public health crises, natural disasters, [climate change](#), and sustainability
- Help [identify](#) and [mitigate societal biases and structural inequities](#) (e.g. socio-economic, socio-demographic and regional inequities)
- Enable [scientific](#) and [other breakthroughs](#) to address humanity's greatest future opportunities and [challenges](#) (e.g. [medical diagnosis](#), drug discovery, climate forecasting)

The foundational nature of AI means that AI will also power and transform existing infrastructure, tools, software, hardware, and devices—including products and services not normally thought of as AI. Examples in our case that are already being transformed by AI include Google Search, Google Maps, Google Photos, Google Workspace, and Android phones. It will significantly enhance their usefulness and multiply their value to people. It will also lead to new categories of assistive tools, products, and services, often with breakthrough capabilities and performance made possible only through AI. This includes more powerful and inclusive language translators, conversational AI and assistants, generative and multi-modal AI, and driverless cars. And this is just the beginning.

2 To what end?

As Google and Alphabet, our goal is to bring to users useful innovations made possible by AI that benefit people and society. Advancing the state of the art helps us expand and progress AI capability to deliver innovations that can assist and improve the lives of many, while generating sustaining value that enables us to continue investing in transformative innovations.

We are pursuing and delivering on this aspiration in several ways:

Advance the state of the art for more capable AI that can benefit people and society.

- **Lead foundational and field-defining breakthrough AI research** to generate AI that is more capable and assistive across a variety of tasks. Examples of our contributions that have helped advance the field and have been leveraged by many at and beyond Google include [Transformers](#), [Word2Vec](#), [Sequence to Sequence Learning](#), [Federated Learning](#), [Model Distillation](#), [Diffusion Models](#), [Deep Reinforcement Learning](#), [Neural Nets with Tree Search](#), [Self-learning Systems](#), [Neural Architecture Search](#), [Autoregressive Models](#), [Networks with External Memory](#), [Large Scale Distributed Deep Networks](#), [Tensor Processing Units](#)
- **Use AI to make breakthrough progress in science and other areas** where we aim to advance scientific and engineering progress. Examples of our widely-acknowledged breakthroughs in AI and science that can benefit all of humanity include: [mapping nearly all known proteins](#), [predicting the function of proteins](#), [mapping a piece of the brain in neuroscience research](#), [discovering faster algorithms](#), [advances in quantum computing](#) and [physics](#), including [innovating in nuclear fusion](#)

Bring useful and transformational AI-powered infrastructure, products, and services to more people, businesses, organizations (large and small), and economies everywhere.

- **Build state of the art AI infrastructure** that is secure and easy to use, including compute (e.g. [Tensor Processing Units](#) and [Colab](#)) and widely-used software frameworks (e.g. [TensorFlow](#), [Jax](#), [Android ML](#) and [Private Compute](#)). Make this AI infrastructure available (with many open source tools) to millions of developers, students, and researchers in various organizations throughout the world.
- **Apply our AI advances to our core products and services** to make step-change improvements, innovations, and new experiences that enhance and multiply the usefulness and value of all our core products and services for billions of people across [Google Search](#), [Google Photos](#), [Google Maps](#), [Google Workspace](#), hardware devices (e.g. Pixel and Nest), and for those with disabilities via accessibility applications ([Android Voice Access](#), [Live Transcribe](#))
- **Develop new AI-powered products, services, and new experiences** for:
 - [Consumers](#) with assistive tools like Google Translate, Google Lens, Google Assistant, [Project Starline](#), speech-to-text, text suggestions and summarization, and generative human-assistive capabilities across [many creative](#) and productivity endeavors. To enhance all products that use speech, we've developed a Universal Speech Model capable of 400 languages and we are committed to building a model to support the 1000 most-spoken languages, improving access for billions of people

- Businesses and organizations of all sizes, across industries and regions, with examples like [contact center assistive agents](#), tools such as [AutoML](#), [Vertex AI](#), [Cloud TPUs](#), [Glass Enterprise](#), and assistive products for coding, design, and more still to come
- Sectors that can most benefit from AI, from manufacturing to life sciences (e.g. [diabetic retinopathy](#), imaging diagnostics, [DeepVariant](#)), directly and indirectly through our partners and collaborators
- **Grow and enable a large AI ecosystem** of developers and partners to build and bring more AI applications to more users, sectors, and regions of the world, for example through our provision of [tools](#), APIs, and in some cases through co-development and [co-deployments](#) of useful innovations
- **Use AI to create new category-defining businesses and companies** that are only possible through the power of AI in a variety of fields from driverless cars ([Waymo](#)) and drug discovery (e.g. [Isomorphic Labs](#)) to robotics (e.g. [Intrinsic](#))

Work with and enable many others beyond Google to apply AI to help tackle society's greatest challenges and opportunities today and in the future.

- **Collaborate with others around the world** to apply AI to society's most pressing challenges such as natural disasters, public health crises, climate change, and sustainability
 - Examples include [AI for Social Good](#), [AI for the UN Sustainable Development Goals](#), [Data Commons](#), wildfire alerts, and [flood forecasting](#) in more than 20 countries around the world so far.
- **Expand and enable the field of AI** by sharing major breakthroughs and related artifacts (e.g. [papers](#), open-source releases, and [datasets](#) such as [AlphaFold](#) protein datasets) and engaging in research collaborations. We also make tools widely available to students and educators (e.g. [Google Scholar](#) and [Colab](#), regularly used by millions of learners), including free access to leading-edge ML computation hardware for scientists (e.g. [TPU Research Cloud](#)), help build their capacity (e.g. with our partnership with the [National Science Foundation](#)), and share best practices (e.g. on safety) with other researchers.

Apply our AI innovations to continually improve Google itself as a leading organization, and to share what we learn with other organizations, including customers, governments, and non-governmental organizations.

Key areas of focus include:

- **Leveraging AI to achieve industry-leading safety and cybersecurity** across all our products and services
- Applying AI to **improve our own productivity and operations** across all functions
- Using AI to help **realize our company's bold ambitions in climate and sustainability** (e.g. [energy efficiency in our data centers](#))

We are compelled by the progress we are making across all the above, as well as our impact to date, in some cases benefitting billions of people. However, we believe still more opportunities for useful and beneficial impact lie ahead.

3 Our understanding of the complexities and risks

As with any transformational technology, [AI comes with complexities and risks](#), and these will change over time. As an early-stage technology, its [evolving capabilities and uses](#) create potential for misapplication, misuse, and unintended or unforeseen consequences. We are taking a proactive approach to understand the evolving complexities and risks as AI advances, deployment grows, and use expands, while continuing to learn from users and the wider community.

Many of the complexities and risks have been highlighted through our [own work](#) and that of many [others](#). Such risks become manifest when AI:

- Does not perform as intended (e.g. safety, quality, accuracy)
- Relies on data that is not used appropriately and responsibly
- (Is) Deployed [unsafely](#) (e.g. too early or with insufficient testing)
- (Is) Misapplied or used in harmful ways by its developers or users
- Creates or amplifies negative societal [biases](#) and [harms](#) (e.g. discrimination)
- Creates or worsens cybersecurity risks
- Creates or worsens [information hazards](#) (e.g. lack of groundedness, [non-factuality, misinformation](#))
- Creates the impression of having capabilities it does not actually have (e.g. interpreting conversational AI to have human traits)
- Creates or worsens inequality or other [socio-economic harms](#) including in the [labor markets](#)

We recognize the harms that these failures can cause, especially for different communities and contexts across the globe, and it is critical to invest in mitigating the above risks to increase trust, ensure safe and inclusive user experiences, and enable AI to fully benefit people and society.

4 Our approach to Responsible AI

Given its risks and complexities, we believe that we as a company must pursue AI responsibly. As leaders in AI, we must lead not only in state-of-the-art AI technologies, but also in state-of-the-art responsible AI—its innovation and implementation. In 2018, we were one of the first companies to articulate [AI Principles](#) that put beneficial use, users, safety, and avoidance of harms above business considerations, and we have pioneered many best practices, like the use of model and data cards, now widely used by others. More than words on paper, we apply our AI Principles in [practice](#). Doing so—along with continual research and review of our approaches—is critical.

Our responsible AI approach includes the following:

- **Focus on AI that is useful and benefits users and society.** Prioritize AI R&D, applications, and uses that assist and benefit people and society. Ensure resource and environmental sustainability throughout R&D.
- **Intentionally apply our [AI Principles](#)** (which are grounded in beneficial uses and avoidance of harm), processes, and governance to guide our work in AI, from research priorities to productization and uses. Continually interpret and update these principles and processes as we learn more and as specific issues arise. We provide regular updates on the [progress on our AI Principles](#).
- **Apply the scientific method to AI R&D** with research rigor, peer review, readiness reviews, and responsible approaches to providing access and to the externalization and use of our innovations. Set benchmarks and measure performance and progress on different factors of responsible AI. Create innovative [tools](#) (e.g. for safety) to keep pace with AI technologies. Continuously perform adversarial and related forms of testing. Through these processes, we take a differentiated and careful approach to access and deployment of novel systems such as [LaMDA](#), [PaLM](#), and Waymo.
- **Collaborate with multidisciplinary experts**, including social scientists, ethicists, and other teams with socio-technical expertise (e.g. our Responsible AI team focused on research, product, and engineering and our Responsible Innovation team focused on products, business, and policy). Work with researchers, developers, and users in areas of societal importance (e.g. [CS Research Mentorship Program](#), research grants, and collaborations).
- **Listen, learn and improve** based on feedback from developers, users, experts, governments, and representatives of affected communities (e.g. [AI Test Kitchen](#), [Crowdsourcing](#)), and involve [human raters to evaluate AI models](#).

- **Conduct regular reviews** of our AI research and application development, including use cases (e.g. our [Advanced Technology Review Council](#)). Provide transparency on learnings (e.g. [PAIR guidebook](#)). Engage with others (e.g. governments) to provide the benefits of our experiences as they shape approaches to concerns and risks.
- **Stay on top of current and evolving areas of concern and risk** (e.g. [safety](#), [bias](#), [toxicity](#) factuality), and address, research, and innovate to respond to challenges/risks as they emerge. Share learnings and innovations (e.g. [open-sourcing the Monk scale](#) and tools for detecting synthetic speech). Develop methods to monitor deployed systems, ensuring that we can quickly mitigate dynamically-occurring risks in production and in-use services.
- **Lead on and help shape responsible governance**, accountability, and regulation that encourages innovation and maximizes the benefits of AI while mitigating risks (e.g. our role in setting up Partnership on AI, our support for Global Partnership on Artificial Intelligence and [our contributions](#) to flagship AI governance efforts, including the EU AI Act, NIST AI Risk Management Framework, and OECD AI Principles).
- **Help users and society understand what AI is (and is not) and how to benefit from its potential**—how it might be helpful in their daily lives (e.g. education), what the risks are, and how to mitigate those risks.

We are leaders in driving change in many areas of Responsible AI, but at the same time we continue to learn from users, other researchers, affected communities, and our experiences. As a result, we are continually refining our approaches to ensure that the above considerations are incorporated in all we do and address issues as they arise. We aim to work in meaningful ways that help shape but don't slow down innovation that can benefit people and society.

5 Why a collective approach to Responsible AI is needed

We believe that getting AI right requires a collective effort. We don't have all the answers, but our experience so far suggests that everyone involved in AI (researchers, developers, deployers, academics, civil society, governments, and users, including individuals, businesses, and other organizations) must work together to get AI right including in the following areas:

- Responsible approaches to AI development and deployment of AI systems
- Data and privacy practices that protect privacy and enable benefits for people and society (e.g. sharing traffic and public safety data)
- Robust AI infrastructure and cybersecurity to mitigate security risks
- Regulations that encourage innovation and safe and beneficial uses of AI and avoid misapplications, misuse, or harmful uses of AI
- Cross-community collaboration to develop standards and best practices
- Sharing and learning together with leaders in government and civil society
- Practical accountability mechanisms to build trust in areas of societal concern
Investment in AI safety, ethics, and sociotechnical research
- Growing a larger and more diverse community of AI practitioners to fully reflect the diversity of the world and to better address its challenges and opportunities

If pursued boldly and responsibly, we believe that AI can be a foundational technology that transforms the lives of people everywhere—this is what excites us!

Explore more at ai.google.