

Responsible AI practices

AI is creating new opportunities to improve the lives of people around the world. Google is committed to the responsible development of AI and to sharing knowledge, research, tools, datasets, and other resources with the larger community.

GENERATIVE AI-PROHIBITED USE POLICY

Last modified: 14 March 2023

Generative AI models can help you explore new topics, inspire your creativity and learn new things. However, we expect you to use and engage with them in a responsible, legal manner. To this end, you must not use the Google services that reference this policy to:

1. Perform or facilitate dangerous, illegal or malicious activities, including
 - a. Facilitation or promotion of illegal activities or violations of law, such as
 - i. Promoting or generating content related to child sexual abuse or exploitation
 - ii. Promoting or facilitating the sale of, or providing instructions for synthesising or accessing, illegal substances, goods or services
 - iii. Facilitating or encouraging users to commit any type of crime
 - iv. Promoting or generating violent extremism or terrorist content
 - b. Abuse, harm, interference or disruption of services (or enable others to do the same), such as
 - i. Promoting or facilitating the generation or distribution of spam
 - ii. Generating content for deceptive or fraudulent activities, scams, phishing or malware.
 - c. Attempts to override or circumvent safety filters or intentionally drive the model to act in a manner that contravenes our policies
 - d. Generation of content that may harm or promote the harm of individuals or a group, such as
 - i. Generating content that promotes or encourages hatred
 - ii. Facilitating methods of harassment or bullying to intimidate, abuse or insult others
 - iii. Generating content that facilitates, promotes or incites violence
 - iv. Generating content that facilitates, promotes or encourages self-harm
 - v. Generating personally identifying information for distribution or other harm
 - vi. Tracking or monitoring people without their consent
 - vii. Generating content that may have unfair or adverse impacts on people, particularly impacts related to sensitive or protected characteristics
2. Generate and distribute content intended to misinform, misrepresent or mislead, including
 - a. Misrepresentation of the provenance of generated content by claiming content was created by a human, or representing generated content as original works, in order to deceive
 - b. Generation of content that impersonates an individual (living or dead) without explicit disclosure, in order to deceive
 - c. Misleading claims of expertise or capability made particularly in sensitive areas (e.g. health, finance, government services or legal)

- d. Making automated decisions in domains that affect material or individual rights or well-being (e.g. finance, legal, employment, healthcare, housing, insurance and social welfare)
3. Generate sexually explicit content, including content created for the purposes of pornography or sexual gratification (e.g. sexual chatbots). Note that this does not include content created for scientific, educational, documentary or artistic purposes.

AI Principles reviews and operations

At Google, a central team is dedicated to ethical reviews of new AI and advanced technologies before launch, working with internal domain experts in machine-learning fairness, security, privacy, human rights, and more.

RESPONSIBILITY:

AI Governance reviews and operations

We assess proposals for new AI research and applications for alignment with our Principles. As advanced technologies emerge and evolve, we'll continue to refine our process.

Overview

At Google, a central team is dedicated to ethical reviews of new AI and advanced technologies before launch, working with internal domain experts in machine-learning fairness, security, privacy, human rights, the social sciences, and, for cultural context, Google's employee resource groups.

1. Intake

Any team can request AI Principles advice. Reviewers also consider an ongoing pipeline of new AI research papers, product ideas, and other projects.

2. Analysis

Reviewers analyze the scale and scope of a technology's potential benefits and harms.

3. Adjustment

Reviewers recommend technical evaluations (e.g., checking for unfair bias in ML models).

4. Decision

Reviewers decide whether to pursue or not pursue the AI application under review (e.g., Cloud AI Hub and text-to-speech).

Implementing Our AI Principles

The ways in which artificial intelligence is built and deployed will significantly affect society. Learn more about how we are applying our AI principles across Google research and products.

Generative AI Terms of Service

Understand the Terms of Service for Google's Generative AI including age requirements, use restrictions, and technology disclaimers.

GENERATIVE AI ADDITIONAL TERMS OF SERVICE

Last modified: 9 August 2023 | Archived versions

To use the Google services that reference these terms (the 'Services'), you must accept (1) the Google Terms of Service, and (2) these Generative AI Additional Terms of Service.

Please read these documents carefully. Together, these documents are known as the 'Terms'. They establish what you can expect from us as you use our services, and what we expect from you.

We also encourage you to read our privacy policy (and any Privacy Notices presented in connection with a Service) to better understand what information we collect and how you can update, manage, export and delete your information.

Age requirements

In addition to the 'Age requirements' section in the Google Terms of Service, you will comply with any further age restrictions presented in connection with a Service.

Use restrictions

You may not use the Services to develop machine-learning models or related technology.

In addition to the 'Respect others' section in the Google Terms of Service, you must comply with our Prohibited use policy, which provides additional details about appropriate conduct when using the Services.

The Services include safety features to block harmful content, such as content that violates our Prohibited use policy. As noted in the 'Respect others' section in the Google Terms of Service, you may not attempt to bypass these protective measures or use content that violates these Terms.

You will comply with Service-specific notices if applicable, such as any warnings not to input sensitive or confidential information.

Disclaimers

The Services may sometimes provide inaccurate or offensive content that doesn't represent Google's views.

Use discretion before relying on, publishing or otherwise using content provided by the Services.

Don't rely on the Services for medical, legal, financial or other professional advice. Any content regarding those topics is provided for informational purposes only and is not a substitute for advice from a qualified professional.

Generative AI Prohibited Use Policy

Learn more about the responsible and legal usage of our Generative AI services, including certain restrictions on the creation and distribution of content.

GENERATIVE AI-PROHIBITED USE POLICY

Last modified: 14 March 2023

Generative AI models can help you explore new topics, inspire your creativity and learn new things. However, we expect you to use and engage with them in a responsible, legal manner. To this end, you must not use the Google services that reference this policy to:

1. Perform or facilitate dangerous, illegal or malicious activities, including
 - a. Facilitation or promotion of illegal activities or violations of law, such as
 - i. Promoting or generating content related to child sexual abuse or exploitation
 - ii. Promoting or facilitating the sale of, or providing instructions for synthesising or accessing, illegal substances, goods or services
 - iii. Facilitating or encouraging users to commit any type of crime
 - iv. Promoting or generating violent extremism or terrorist content
 - b. Abuse, harm, interference or disruption of services (or enable others to do the same), such as
 - i. Promoting or facilitating the generation or distribution of spam
 - ii. Generating content for deceptive or fraudulent activities, scams, phishing or malware.
 - c. Attempts to override or circumvent safety filters or intentionally drive the model to act in a manner that contravenes our policies
 - d. Generation of content that may harm or promote the harm of individuals or a group, such as
 - i. Generating content that promotes or encourages hatred
 - ii. Facilitating methods of harassment or bullying to intimidate, abuse or insult others
 - iii. Generating content that facilitates, promotes or incites violence
 - iv. Generating content that facilitates, promotes or encourages self-harm
 - v. Generating personally identifying information for distribution or other harm
 - vi. Tracking or monitoring people without their consent
 - vii. Generating content that may have unfair or adverse impacts on people, particularly impacts related to sensitive or protected characteristics
2. Generate and distribute content intended to misinform, misrepresent or mislead, including
 - a. Misrepresentation of the provenance of generated content by claiming content was created by a human, or representing generated content as original works, in order to deceive
 - b. Generation of content that impersonates an individual (living or dead) without explicit disclosure, in order to deceive
 - c. Misleading claims of expertise or capability made particularly in sensitive areas (e.g. health, finance, government services or legal)

- d. Making automated decisions in domains that affect material or individual rights or well-being (e.g. finance, legal, employment, healthcare, housing, insurance and social welfare)
- 3. Generate sexually explicit content, including content created for the purposes of pornography or sexual gratification (e.g. sexual chatbots). Note that this does not include content created for scientific, educational, documentary or artistic purposes.