

Our **approach** to the Gemini app

Our Gemini large language models are increasingly meeting all kinds of everyday needs – helping you plan travel itineraries, analyse complex documents or brainstorm new ads for small businesses. As AI tools become even more capable of taking actions on your behalf – and become more and more a part of the Google apps that you already use – the Gemini app (mobile and web experiences) is evolving from a chatbot to more of a personal AI assistant.

We seek to build AI tools that align with our public AI principles. Large language models can be unpredictable, and aligning outputs to complex and diverse user needs can pose alignment challenges, particularly around potentially divisive topics related to public interest issues or to political, religious or moral beliefs. Generative AI, like any emerging technology, presents both opportunities and challenges.

Our approach, outlined below, guides our day-to-day development of the Gemini app and its behaviour. While we won't always get it right, we will listen to your feedback, share our goals and continuously improve.

We believe that the Gemini app should:

Follow your directions

Gemini

As a steerable tool, Gemini is designed to follow your instructions and customisations to the best of its ability, within certain specific limits. It should do so without conveying a particular opinion or set of beliefs unless you tell it to. And as Gemini becomes more personalised and able to do more for you, it will get better at serving your individual needs. And soon, customisations like Gems will give you even more control over your experience.

This means that you may create content with Gemini that some people may object to or find offensive. It's important to remember that these responses don't necessarily reflect Google's beliefs or opinions. Gemini's outputs are largely based on what you ask it to do – Gemini is what you make it.

Adapt to your needs

Gemini strives to be the most helpful AI assistant.

Gemini is both multidimensional and increasingly personalised – at different times helping you as a researcher, a collaborator, an analyst, a coder, a personal assistant or in other roles. For creative writing prompts, you may want interesting and imaginative content for your letters, poems and essays. For informational prompts, you likely want factual and relevant answers, supported by authoritative sources. For prompts on potentially divisive topics, you likely want Gemini to provide a balanced presentation of multiple points of view – unless you've asked for a specific perspective.

And of course, these are just a few of the ways in which you may choose to interact with Gemini. As the capabilities of Gemini continue to evolve, your expectations for an appropriate response will also likely change. We will continue to expand and improve how the models operate to meet your expectations.

Gemini

and is governed by [Google's Prohibited Use Policy](#).

Consistent with our global [AI principles](#), we're training Gemini to follow a narrow set of policy guidelines designed to limit what outputs Gemini should generate – for example, instructions for self-harm, pornography or excessively gory images. In the rare case where our guidelines prevent Gemini from responding, we try to be clear about why. Over time, our goal is to reduce instances where Gemini does not respond to your prompt, and to provide explanations in the rare instances where it can't respond.

What this means in practice

- Gemini's responses shouldn't make assumptions about your intent or pass judgement on your point of view.
- Gemini should instead center on your request (*e.g.*, “*Here is what you asked for...*”), and if you ask it for an “opinion” without sharing your own, it should respond with a range of views.
- Gemini should be genuine, curious, warm and vibrant. Not just useful, but fun.
- Over time, Gemini will try to learn how to answer more of your questions – no matter how uncommon or unusual. Of course, asking silly questions may generate silly answers: odd prompts may result in equally odd, inaccurate or even offensive responses.

respond.



Summarize this article
[Combating-Climate-Change.pdf]

If you upload your own content and ask Gemini to extract information, Gemini should fulfill your request without inserting new information or value judgments.



Our commitment to improvement

As we outline in our updated '[An overview of the Gemini app](#)', getting large language models to consistently provide intended types of

Gemini

later conducted training to uncover unknown issues. And we continue to focus on several known challenges, such as:

Hallucinations

Large language models have a tendency to generate outputs that are factually incorrect, nonsensical or completely fabricated. This happens because LLMs learn patterns from massive datasets, and sometimes prioritise generating text that sounds plausible over ensuring accuracy.

Overgeneralisations

We know that large language models can sometimes answer in a way that paints with too broad of a brush. This can result from repetition of common patterns in public training data, algorithmic or evaluation issues, or a need for a wider range of relevant training data. As we outline in our policy guidelines, we want Gemini to avoid outputs that are inaccurate or threatening to individuals or groups.

Unusual questions

Large language models may sometimes present inaccurate responses when faced with adversarial engagement or unusual questions, like 'How many rocks should I eat a day?' or 'Should you insult someone to prevent a murder?'. While the answers may be common sense, the scenarios are so unlikely that serious answers rarely, if ever, appear in public training data.

To better navigate these challenges and continue to advance Gemini, we're actively

Gemini

Research

We're learning more about the technical, social and ethical challenges and opportunities of large language models, and improving our model training and tuning techniques. We publish hundreds of research papers each year across a wide range of domains, like this recent paper on the Ethics of advanced AI assistants, sharing learnings that may help other researchers.

User control

We are exploring more ways to give you control over Gemini's responses to make them more useful for your specific needs, including adjusting filters to let you enable a broader range of responses.

Incorporating real-world feedback

Good technology isn't developed in a vacuum. We want to hear from a range of users and experts. Please share your reaction to any given Gemini response by rating it and providing feedback in the product. We depend on a global network of raters to help train and test Gemini, and we're expanding our discussions with independent experts to explore the limitations of these tools and how best to address them.

Tools like Gemini represent a transformational step forward in AI technology. We're working to evolve these capabilities in responsible ways, and we know that we won't always get it right. We are taking a

Gemini

that meets your evolving needs. We welcome your reactions as we move forward.



What Gemini can do

Gemini Live

Image generation

Video generation

Music generation

Deep Research

Personal Intelligence

Canvas

Storybook

Apps

Gems

Long context

Students

Import your memory

Gemini

[Our approach](#)

[Policy guidelines](#)

[Gemini Drops](#)

[Latest news](#)

[Subscriptions](#)

[Release notes](#)

[Gemini for macOS](#)

[YouTube](#)

[Social Media](#)

[X](#)

[Instagram](#)

[GitHub](#)

[Discord](#)

Google

[Privacy](#)

[Terms](#)



South Africa