OFÍCIO No. 013627/2024/SUNNG/NGPG2/NGSG2

Brasília, 19 de novembro de 2024

Ao
TRIBUNAL DE CONTAS DO ESTADO DO RIO GRANDE DO NORTE
CNPJ: 12.978.037/0001-78

Assunto: Ofício Proposta - Serviços SERPRO MULTICLOUD

Prezado(a) Senhor(a),

O Serviço Federal de Processamento de Dados – SERPRO, empresa pública federal, com sede no SGAN, Quadra 601, Módulo V, Brasília/DF, CEP: 70.836-900, inscrita no CNPJ/MF sob o no 33.683.111/0001-07, apresenta as condições técnicas, financeiras e comerciais relativos ao serviço de **Serpro Multicloud**, conforme a seguir:

1. DO OBJETO PARA CONTRATAÇÃO

O objeto desta proposta é a prestação dos serviços pelo SERPRO relativos à contratação do serviço de **Serpro Multicloud,** que serão prestados nas condições estabelecidas no Termo de Adesão.

2. PREÇOS E VALOR DO SERVIÇO

O valor total dos serviços constantes para o período de 12 (doze) meses, considerando os dados informados é de R$ 18.440,14 (Dezoito mil, quatrocentos e quarenta reais e quatorze centavos) conforme tabela abaixo e demonstrada na figura a seguir:

SERPRO Multicloud

Nº do Orçamento: **86214112024**

Cliente	Projeto	Fator Câmbio	Qtd. Meses	Data do Orçamento
[TCERN] (TCE/RN) Tribunal de Contas do Estado do Rio Grande do Norte	IA Azure	R$ 5,79	12	14/11/2024

Resumo

Grupo	Atividade	Volume	Vlr. Unitário	Vlr. Mensal	Vlr. no Período
Infraestrutura	Cloud Service Brokerage	868,18	R$ 1,77	R$ 1.536,68	R$ 18.440,14
Sustentação	Cloud Service Management	0,00	R$ 0,00	R$ 0,00	R$ 0,00
Consultoria	Cloud Service Archtecture Design	0,00	R$ 1.297,00	R$ 0,00	R$ 0,00
Consultoria	Cloud Engineering and Automation	0,00	R$ 1.297,00	R$ 0,00	R$ 0,00
Consultoria	Cloud Generic Professional Service	0,00	R$ 1.297,00	R$ 0,00	R$ 0,00
Consultoria	Cloud Migration and Management	0,00	R$ 1.297,00	R$ 0,00	R$ 0,00
	TOTAL			**R$ 1.536,68**	**R$ 18.440,14**

Referências e Legendas

Cloud Service Brokerage (CSB): Serviço profissional de corretagem de serviços em nuvem e de disponibilização e operação da plataforma multinuvem, visando gerenciar o uso, o desempenho e a entrega, assim como os relacionamentos entre os provedores e consumidores desses serviços em nuvem.
Cloud Service Management (CSM): Serviço continuado de suporte a infraestrutura de nuvem providos por equipes especializadas, que atuam de forma multidisciplinar na sustentação da infraestrutura em nuvem do cliente visando disponibilidade, desempenho e segurança deste ambiente.
Cloud Migration and Management (CMM): Consultoria técnica especializada em planejamento de migrações e eventos críticos utilizada para planejamento e acompanhamento dos eventos de migração dos serviços para nuvem.
Cloud Service Archtecture Design (CAD): Consultoria técnica especializada em arquitetura e design de soluções em nuvem.
Cloud Engeneering and Automation (CEA): Consultoria técnica especializada em engenharia e automação de ambientes DevSecOps.
Cloud Generic Professional Service (CGPS): Consultoria técnica especializada para avaliação técnica, execução de procedimentos ou outras atividades correlatas ao SERPRO Multicloud não especificadas nas demais especializações.

Os valores aqui descritos já incluem a tributação necessária para execução dos serviços propostos, conforme a legislação tributária vigente até a apresentação deste.

3. DAS CONDIÇÕES DE PAGAMENTO

O pagamento do serviço prestado, será efetuado até o vigésimo dia após a emissão da nota fiscal e/ou nota fiscal eletrônica de serviços, ou de acordo com a data constante na Nota Fiscal, ou no boleto de pagamento, sendo o faturamento efetuado com base nos serviços efetivamente executados no período do dia 21 do mês anterior ao dia 20 do corrente mês de faturamento. SERPRO - SEDE SGAN Quadra 601 - Módulo V - CEP 70836900 - DF-Brasil CNPJ:33.683.111/0001-07 Telefone: (61) 2021-8000.

4. DO PROCESSO DE CONTRATAÇÃO

A contratação se dará por meio de Contrato de Adesão, celebrado por dispensa de licitação, com base no disposto no inc. IX, art. 75 da Lei nº 14.133/2021.

5. DO CONTRATO DE ADESÃO

O Contrato de Adesão foi elaborado em conformidade com a legislação vigente, de forma a atender às necessidades e peculiaridades dos órgão contratantes, tendo aprovação do Jurídico do SERPRO. Assim sendo, caso haja divergência no entendimento legal e ou conteúdo técnico, favor enviar notificação formal com parecer correlato do vosso jurídico, uma vez que será objeto de análise do

Jurídico Serpro, e sendo pertinente, emitirá parecer para aplicabilidade em todos os contratos de adesão vigentes, pertinentes aos serviços.

6. DO DETALHAMENTO DOS SERVIÇOS

Considerando que o contrato a ser firmado é de adesão ao serviço, as demais informações técnicas e legais, bem como características, tabela de preços e detalhamento da execução do serviço, estão dispostos na respectiva minuta de contrato anexo.

7. DA VALIDADE

Para execução do processo de contratação dos serviços, aguardamos manifesto com o aceite do proposto, em até **60 (sessenta) dias** a contar da data do recebimento deste.

Agradecemos e nos colocamos à disposição para sanar eventuais dúvidas, na pessoa da Analista de Negócios, Humberto Rocha de Almeida Neto, pelos telefones: (83) 99804-0116/(81) 2126-4502 (*WhatsApp*) ou pelo e-mail: humberto.almeida@serpro.gov.br.

Atenciosamente,

NeoSigner Assinatura ICP-Brasil
ALEXANDRA VITORIO DE MORAIS SILVA
CPF: 031.821.984-08
19/11/2024
SERPRO

ALEXANDRA VITORIO DE MORAIS SILVA
Gerente de Divisão
Superintendência de Novos Negócios

Preços do Serviço OpenAI Azure

Solicitar uma cotação de preços | Testar o Azure gratuitamente

Visão geral de preços do Serviço OpenAI do Azure

Unlock the power of Azure OpenAI Service's generative AI models with flexible Standard (On-Demand) and Provisioned Throughput Units (PTUs). The Standard model lets you pay only for tokens processed, while PTUs ensure consistent throughput and minimal latency variance for scalable solutions. Pricing includes costs per 1,000 tokens, and PTU rates provide a predictable cost structure. Azure OpenAI Service offers advanced capabilities like GPT-4o, fine-tuning for customization, DALL-E for image generation, and Whisper for speech-to-text. For personalized guidance on optimizing AI deployments, contact a sales specialist.

Explorar as opções de preços

Aplique filtros para personalizar as opções de preço conforme as suas necessidades.

Os preços são apenas estimativas e não pretendem ser cotações de preços reais. O preço real pode variar dependendo do tipo de contrato celebrado com a Microsoft, data de compra e taxa de câmbio. Os preços são calculados com base em dólares americanos e convertidos usando as taxas spot de fechamento de Londres capturadas nos dois dias úteis anteriores ao último dia útil do final do mês anterior. Se os dois dias úteis anteriores ao final do mês caírem em um feriado bancário nos principais mercados, o dia de definição da taxa geralmente é o dia imediatamente anterior aos dois dias úteis. Esta taxa se aplica a todas as transações durante o próximo mês. Entre na calculadora de preços do Azure para ver os preços com base em seu programa/oferta atual com a Microsoft. Entre em contato com um especialista de vendas do Azure para obter mais informações sobre preços ou para solicitar uma cotação. Veja as perguntas frequentes sobre os preços do Azure.

Região: Leste dos EUA

Moeda: Estados Unidos – Dólar – USD (US$)

Detalhes dos preços:

Modelos de linguagem

Modelos	Contexto	Entrada (por 1.000 tokens)	Saída (por 1.000 tokens)	Price per PTU per Hour	Minimum Scaling Increment	Monthly Reservation per PTU	Yearly Reservation per PTU
Implantação Global do GPT-4o	128K	$0,005	$0,015	N/A	N/A	N/A	N/A
API Regional do GPT-4o	128K	$0,005	$0,015	$2	50 PTUs	$260	$2.652
GPT-4o-mini Global Deployment	128K	$0,00015	$0,0006	N/A	N/A	N/A	N/A
GPT-4o-mini Regional API	128K	$0,000165	$0,00066	$2	25 PTUs	$260	$2.652
GPT-3.5-Turbo-0125	16K	$0,0005	$0,0015	$2	100 PTUs	$260	$2.652
GPT-3.5-Turbo-Instruct	4K	$0,0015	$0,002	N/A	N/A	N/A	N/A
GPT-4-Turbo	128K	N/D	N/D	$2	100 PTUs	$260	$2.652
GPT-4-Turbo-Visão	128K	N/D	N/D	N/A	N/A	N/A	N/A
GPT-4	8K	$0,03	$0,06	$2	50 PTUs	$260	$2.652
GPT-4	32K	$0,06	$0,12	$2	200 PTUs	$260	$2.652

This table provides a detailed comparison of Standard (On-Demand) versus Provisioned (PTU) pricing for various language models. The 'Context' column specifies the maximum number of tokens each model can handle per response. Pricing details for input and output tokens are listed, reflecting the cost per 1,000 tokens. The PTU pricing model includes an hourly rate and a minimum scaling increment, representing the minimum number of PTUs required for each model. The "Monthly Reservation per PTU" and "Yearly Reservation per PTU" columns indicate the reservation costs per PTU. This comparison helps users understand the cost implications of using each model under both Standard (On-Demand) and Provisioned (PTU) billing options, allowing for informed decisions based on their specific usage needs.

Language models are also now available in the Batch API that returns completions within 24 hours for a 50% discount.

Modelos de linguagem herdados

Modelos	Contexto	Entrada (por 1.000 tokens)	Saída (por 1.000 tokens)
GPT-3.5-Turbo-0301	4K	$0,002	$0,002

GPT-3.5-Turbo-0613	4K	$0,0015	$0,002
GPT-3.5-Turbo-0613	16K	$0,003	$0,004
GPT-3.5-Turbo-1106	16K	$0,001	$0,002

API de Assistentes

The Assistants API and its tools make it easy for developers to build AI Assistants in their applications.

The tokens used for the Assistants API are billed at the chosen language model's per token input/output rates used with each Assistant. Additionally, we charge the following fees for tool usage:

Ferramenta	Entrada
File Search*	$0,10/GB of vector-storage per day (1 GB free)
Code Interpreter**	$0,03/session

*GB refers to binary gigabytes, where 1 gb is 2^30 bytes.

**If your assistant calls Code Interpreter simultaneously in two different threads, this would create two Code Interpreter sessions (2 * $0,03). Each session is active by default for one hour, which means that you would only pay this fee once if your user keeps giving instructions to Code Interpreter in the same thread for up to one hour.

Inference cost (input and output) varies based on the GPT model used with each Assistant. If your assistant calls Code Interpreter simultaneously in two different threads, this would create two Code Interpreter sessions (2 * $0,03). Each session is active by default for one hour, which means that the price is for up to one hour of giving instructions to Code Interpreter in the same thread.

Modelos base

Modelos	Uso por 1.000 tokens
Babbage-002	$0,0004
Davinci-002	$0,002

Modelos de ajuste fino

Preço não disponível na região selecionada

Modelos de imagem

Modelos	Qualidade	Resolução	Preço (por 100 imagens)
Dall-E-3	Standard	1024 * 1024	$4
	Standard	1024 * 1792, 1792 * 1024	$8
Dall-E-3	HD	1024 * 1024	$8
	HD	1024 * 1792, 1792 * 1024	$12
Dall-E-2	Standard	1024 * 1024	$2

Inserindo modelos

Modelos	Por 1.000 tokens
Ada	$0,0001
text-embedding-3-large	$0,00013
text-embedding-3-small	$0,00002

Modelos de Fala

Preço não disponível na região selecionada

Opções de compra e preços do Azure

Conecte-se diretamente conosco

Obtenha uma explicação detalhada sobre os preços do Azure. Entenda os preços da sua solução de nuvem, aprenda sobre a otimização de custos e solicite uma proposta personalizada.

Converse com um especialista de vendas

Confira maneiras de comprar

Compre os serviços do Azure por meio do site do Azure, de um representante da Microsoft ou de um parceiro do Azure.

Explore suas opções >

Recursos adicionais

Serviço OpenAI Azure

Saiba mais sobre os recursos e as funcionalidades do Serviço OpenAI Azure.

Calculadora de preço

Estime seus custos mensais esperados para usar qualquer combinação de produtos do Azure.

SLA

Revise o Contrato de Nível de Serviço para Serviço OpenAI Azure.

Documentação

Consulte tutoriais técnicos, vídeos e outros recursos do Serviço OpenAI Azure.

Perguntas frequentes

Perguntas frequentes sobre os preços do Azure >

- Qual é o preço do Serviço OpenAI do Azure?
- Onde o Serviço OpenAI do Azure está disponível?
- Qual é o SLA do Serviço OpenAI do Azure?
- Como posso saber mais sobre PTUs (unidades de produtividade provisionadas)?

Converse com um especialista em vendas para saber mais sobre os preços do Azure. Entenda os preços da sua solução de nuvem.

Solicitar uma cotação de preços

Obtenha serviços de nuvem gratuitos e um crédito de USD200 para explorar o Azure por 30 dias.

Testar o Azure gratuitamente

Obter o aplicativo móvel do Azure

Explore o Azure
- O que é o Azure?
- Primeiros passos
- Infraestrutura global
- Regiões de datacenter
- Confie na sua nuvem
- Habilitação do cliente
- Histórias de clientes

Limites e preços
- Produtos
- Preços
- Serviços gratuitos do Azure
- Opções de compra flexíveis
- Economia da nuvem
- Otimize seus custos

Soluções e suporte
- Soluções
- Recursos para acelerar o crescimento
- Arquiteturas para soluções
- Suporte
- Demonstração do Azure e sessão de P e R ao vivo

Parceiros
- Azure Marketplace
- Encontre um parceiro
- Ingresse no Sucesso do ISV

Referências
- Treinamento e certificações
- Documentação
- Blog
- Recursos para desenvolvedores
- Alunos
- Webinars e eventos
- Relatórios de analistas, white papers e livros eletrônicos
- Vídeos

Computação em nuvem
- O que é computação em nuvem?
- O que é migração na nuvem?
- O que é uma nuvem híbrida?
- O que é IA?
- O que é IaaS?
- O que é SaaS?
- O que é PaaS?
- O que é o DevOps?

Alterar idioma

Português (Brasil)

Suas opções de privacidade

Chat com Vendas

Pricing

Simple and flexible. Only pay for what you use.

Contact sales

Latest models

Multiple models, each with different capabilities and price points. Prices can be viewed in units of either per 1M or 1K tokens. You can think of tokens as pieces of words, where 1,000 tokens is about 750 words.

Language models are also available in the Batch API that returns completions within 24 hours for a 50% discount.

GPT-4o

GPT-4o is our most advanced multimodal model that's faster and cheaper than GPT-4 Turbo with stronger vision capabilities. The model has 128K context and an October 2023 knowledge cutoff.

Learn about GPT-4o ↗

Model	Pricing	Pricing with Batch API*
gpt-4o	$5.00 / 1M input tokens	$2.50 / 1M input tokens
	$15.00 / 1M output tokens	$7.50 / 1M output tokens
gpt-4o-2024-08-06	$2.50 / 1M input tokens	$1.25 / 1M input tokens
	$10.00 / 1M output tokens	$5.00 / 1M output tokens
gpt-4o-2024-05-13	$5.00 / 1M input tokens	$2.50 / 1M input tokens
	$15.00 / 1M output tokens	$7.50 / 1M output tokens

Vision pricing calculator

Set model

chatgpt-4o-latest

Set width: 150 px by Set height: 150 px = $0.001275

Low resolution

*Batch API pricing requires requests to be submitted as a batch. Responses will be returned within 24 hours for a 50% discount. Learn more about Batch API ↗

GPT-4o mini

GPT-4o mini is our most cost-efficient small model that's smarter and cheaper than GPT-3.5 Turbo, and has vision capabilities. The model has 128K context and an October 2023 knowledge cutoff.

Learn about GPT-4o mini ↗

Model	Pricing	Pricing with Batch API*
gpt-4o-mini	$0.150 / 1M input tokens	$0.075 / 1M input tokens
	$0.600 / 1M output tokens	$0.300 / 1M output tokens
gpt-4o-mini-2024-07-18	$0.150 / 1M input tokens	$0.075 / 1M input tokens
	$0.600 / 1M output tokens	$0.300 / 1M output tokens

Vision pricing calculator

Set width: 150 px by Set height: 150 px = $0.001275

☐ Low resolution

*Batch API pricing requires requests to be submitted as a batch. Responses will be returned within 24 hours for a 50% discount. Learn more about Batch API ↗

Embedding models

Build advanced search, clustering, topic modeling, and classification functionality with our embeddings offering.

Learn about embeddings ↗

Model	Pricing	Pricing with Batch API*
text-embedding-3-small	$0.020 / 1M tokens	$0.010 / 1M tokens
text-embedding-3-large	$0.130 / 1M tokens	$0.065 / 1M tokens
ada v2	$0.100 / 1M tokens	$0.050 / 1M tokens

*Batch API pricing requires requests to be submitted as a batch. Responses will be returned within 24 hours for a 50% discount. Learn more about Batch API ↗

Fine-tuning models

Create your own custom models by fine-tuning our base models with your training data. Once you fine-tune a model, you'll be billed only for the tokens you use in requests to that model.

Learn about fine-tuning ↗

Model	Pricing	Pricing with Batch API*
gpt-4o-mini-2024-07-18**	$0.30 / 1M input tokens	$0.15 / 1M input tokens
	$1.20 / 1M output tokens	$0.60 / 1M output tokens
	$3.00 / 1M training tokens	
gpt-3.5-turbo	$3.00 / 1M input tokens	$1.50 / 1M input tokens
	$6.00 / 1M output tokens	$3.00 / 1M output tokens
	$8.00 / 1M training tokens	
davinci-002	$12.00 / 1M input tokens	$6.00 / 1M input tokens
	$12.00 / 1M output tokens	$6.00 / 1M output tokens
	$6.00 / 1M training tokens	
babbage-002	$1.60 / 1M input tokens	$0.80 / 1M input tokens
	$1.60 / 1M output tokens	$0.80 / 1M output tokens
	$0.40 / 1M training tokens	

*Batch API pricing requires requests to be submitted as a batch. Responses will be returned within 24 hours for a 50% discount. Learn more about Batch API ↗

**Fine-tuning for GPT-4o mini is free up to a daily token limit through September 23, 2024. Each qualifying org gets up to 2M complimentary training tokens daily and any overage will be charged at the normal rate of $3.00/1M tokens.

Assistants API

The Assistants API and its tools make it easy for developers to build AI assistants in their applications. The tokens used for the Assistant API are billed at the chosen language model's per-token input / output rates.

Learn about Assistants API ↗

Additionally, we charge the following fees for tool usage:

Tool	Input
Code Interpreter	$0.03 / session
File Search	$0.10 / GB of vector-storage per day (1 GB free)

GB refers to binary gigabytes (also known as gibibyte), where 1 GB is 2^30 bytes.

Image models

Build DALL·E directly into your apps to generate and edit novel images and art. DALL·E 3 is the highest quality model and DALL·E 2 is optimized for lower cost.

Learn about image generation ↗

Model	Quality	Resolution	Price
DALL·E 3	Standard	1024×1024	$0.040 / image
	Standard	1024×1792, 1792×1024	$0.080 / image
DALL·E 3	HD	1024×1024	$0.080 / image
	HD	1024×1792, 1792×1024	$0.120 / image
DALL·E 2		1024×1024	$0.020 / image
		512×512	$0.018 / image
		256×256	$0.016 / image

Audio models

Whisper can transcribe speech into text and translate many languages into English. Text-to-speech (TTS) can convert text into spoken audio.

Learn about Whisper ↗

Learn about Text-to-speech (TTS) ↗

Model	Usage
Whisper	$0.006 / minute (rounded to the nearest second)
TTS	$15.000 / 1M characters
TTS HD	$30.000 / 1M characters

Other Models

While we continuously improve our latest models, here is a list of other models that we support.

Model	Input	Output
chatgpt-4o-latest	$5.00 / 1M tokens	$15.00 / 1M tokens
gpt-4-turbo	$10.00 / 1M tokens	$30.00 / 1M tokens
gpt-4-turbo-2024-04-09	$10.00 / 1M tokens	$30.00 / 1M tokens
gpt-4	$30.00 / 1M tokens	$60.00 / 1M tokens
gpt-4-32k	$60.00 / 1M tokens	$120.00 / 1M tokens
gpt-4-0125-preview	$10.00 / 1M tokens	$30.00 / 1M tokens
gpt-4-1106-preview	$10.00 / 1M tokens	$30.00 / 1M tokens
gpt-4-vision-preview	$10.00 / 1M tokens	$30.00 / 1M tokens
gpt-3.5-turbo-0125	$0.50 / 1M tokens	$1.50 / 1M tokens
gpt-3.5-turbo-instruct	$1.50 / 1M tokens	$2.00 / 1M tokens
gpt-3.5-turbo-1106	$1.00 / 1M tokens	$2.00 / 1M tokens
gpt-3.5-turbo-0613	$1.50 / 1M tokens	$2.00 / 1M tokens
gpt-3.5-turbo-16k-0613	$3.00 / 1M tokens	$4.00 / 1M tokens
gpt-3.5-turbo-0301	$1.50 / 1M tokens	$2.00 / 1M tokens
davinci-002	$2.00 / 1M tokens	$2.00 / 1M tokens
babbage-002	$0.40 / 1M tokens	$0.40 / 1M tokens

☐ Save 50% with Batch API

FAQ

What's a token? −

You can think of tokens as pieces of words used for natural language processing. For English text, 1 token is approximately 4 characters or 0.75 words. As a point of reference, the collected works of Shakespeare are about 900,000 words or 1.2M tokens.

To learn more about how tokens work and estimate your usage...

- Experiment with our interactive Tokenizer tool.
- Log in to your account and enter text into the Playground. The counter in the footer will display how many tokens are in your text.

Which model should I use? +

How will I know how many tokens I've used each month? +

How can I manage my spending? +

Is the ChatGPT API included in the ChatGPT Plus, Teams, or Enterprise subscription? +

Does Playground usage count against my quota? +

How is pricing calculated for Completions? +

How is pricing calculated for Fine-tuning? +

Which models support vision capabilities and how is pricing calculated? +

Is there an SLA on the various models? +

Start creating with OpenAI's powerful models.

Get started | Contact sales ›

Preço para ajudar você a levar seu app para o mundo

Já disponível

Gemini 1.5 Pro

Nosso modelo de última geração com uma janela de contexto inovadora de 2 milhões. Agora em disponibilidade geral para uso na produção.

Sem custo financeiro*	Pagamento por uso (preços em USD)***
Limites de taxa** 2 RPM (solicitações por minuto) 32.000 TPM (tokens por minuto) 50 RPD (solicitações por dia)	Limites de taxa** 360 RPM (solicitações por minuto) 4 milhões de TPM (tokens por minuto)
Preço de entrada (por 1 milhão de tokens) Sem custo financeiro	Preço de entrada (por 1 milhão de tokens) US$ 3,50 para 128 mil tokens <= US$ 7,00 para > 128 mil tokens
Armazenamento em cache de contexto (por 1 milhão de tokens) Não relevante	Armazenamento em cache de contexto (por 1 milhão de tokens) US$ 0,875 para 128 mil tokens <= US$ 1,75 para > 128 mil tokens US$ 4,50 / 1 milhão de tokens por hora (armazenamento) Saiba mais
Preço de saída (por 1 milhão de tokens) Sem custo financeiro	Preço de saída (por 1 milhão de tokens) US$ 10,50 para 128 mil tokens <= US$ 21,00 para > 128 mil tokens
Ajuste de preço Os preços de entrada/saída são os mesmos para modelos ajustados. O serviço de ajuste é sem custo financeiro.	Ajuste de preço Os preços de entrada/saída são os mesmos para modelos ajustados. O serviço de ajuste é sem custo financeiro.
Comandos/respostas usados para melhorar nossos produtos Sim Saiba mais	Comandos/respostas usados para melhorar nossos produtos Não Saiba mais
Teste agora no Google AI Studio	Configurar o faturamento no Google AI Studio

*As restrições de uso do nível sem custo financeiro da API Gemini se aplicam ao EEE (incluindo a UE), ao Reino Unido e à Suíça. Consulte Perguntas frequentes sobre faturamento para mais detalhes.

**Os limites de taxa especificados não são garantidos, e a capacidade real pode variar. Solicite um limite máximo de taxa maior (somente para o nível pago).

***Os custos de inferência de modelos ajustados são cobrados com o mesmo preço dos modelos base. Para receber ajuda com faturamento, consulte Suporte do Cloud Billing.

****Os preços podem ser diferentes dos listados aqui e dos oferecidos na Vertex AI. Para preços da Vertex, consulte a documentação da Vertex.

Teste a IA empresarial do Google Cloud