

Porcentagem de Código Produzido por IA em Big Techs

Proporção de linhas de código em produção geradas ou autocompletadas por IA.

MATRIZ DE ALOCAÇÃO ESTRATÉGICA DE RECURSOS

CUSTO VS VELOCIDADE

 **Raciocínio de Alta Precisão**

Elite

Auditorias, código crítico, decisões de crédito e planejamento estratégico.

US\$ 3.00 – US\$ 4.50 / 1M tokens

 **Soberania & On-Premise**

Zero Risk

Dados sigilosos, defesa, saúde privada e conformidade estrita com LGPD.

Custo Fixo de GPU Interna

 **Automação em Massa**

Sweet Spot

Triagem de dados, suporte ao cliente, extrações estruturadas e e-mails.

US\$ 0.10 – US\$ 0.55 / 1M tokens

 **Real-Time & Voz Interativa**

< 150ms

Agentes de voz bidirecionais, busca instantânea e automações de tela.

Sub-150ms TTFT

DESTAQUE DE MERCADO

Google reporta > 30% e Microsoft > 40% de código gerado por IA em 2026

35% a 48% do Código Global

TI

DIRETRIZ DE ENGENHARIA

Adoção de Model Routers inteligentes reduz a fatura de computação em até 70% sem sacrificar acurácia.

CANAL TELEGRAM

t.me/oquevemai

Fontes: Google Alphabet 10-K, Microsoft Build Keynote, GitHub Data (2026)

oquevem.ai · Intelligence Hub

Porcentagem de Código Produzido por IA em Big Techs

Proporção de linhas de código em produção geradas ou autocompletadas por IA.

MATRIZ DE ALOCAÇÃO ESTRATÉGICA DE RECURSOS

CUSTO VS VELOCIDADE

Raciocínio de Alta Precisão

Elite

Auditorias, código crítico, decisões de crédito e planejamento estratégico.

US\$ 3.00 – US\$ 4.50 / 1M tokens

Soberania & On-Premise

Zero Risk

Dados sigilosos, defesa, saúde privada e conformidade estrita com LGPD.

Custo Fixo de GPU Interna

Automação em Massa

Sweet Spot

Triagem de dados, suporte ao cliente, extrações estruturadas e e-mails.

US\$ 0.10 – US\$ 0.55 / 1M tokens

Real-Time & Voz Interativa

< 150ms

Agentes de voz bidirecionais, busca instantânea e automações de tela.

Sub-150ms TTFT

DESTAQUE DE MERCADO

Google reporta > 30% e Microsoft > 40% de código gerado por IA em 2026

35% a 48% do Código Global

TI

DIRETRIZ DE ENGENHARIA

Adoção de Model Routers inteligentes reduz a fatura de computação em até 70% sem sacrificar acurácia.

CANAL TELEGRAM

t.me/oquevemai

Fontes: Google Alphabet 10-K, Microsoft Build Keynote, GitHub Data (2026)

oquevem.ai - Intelligence Hub