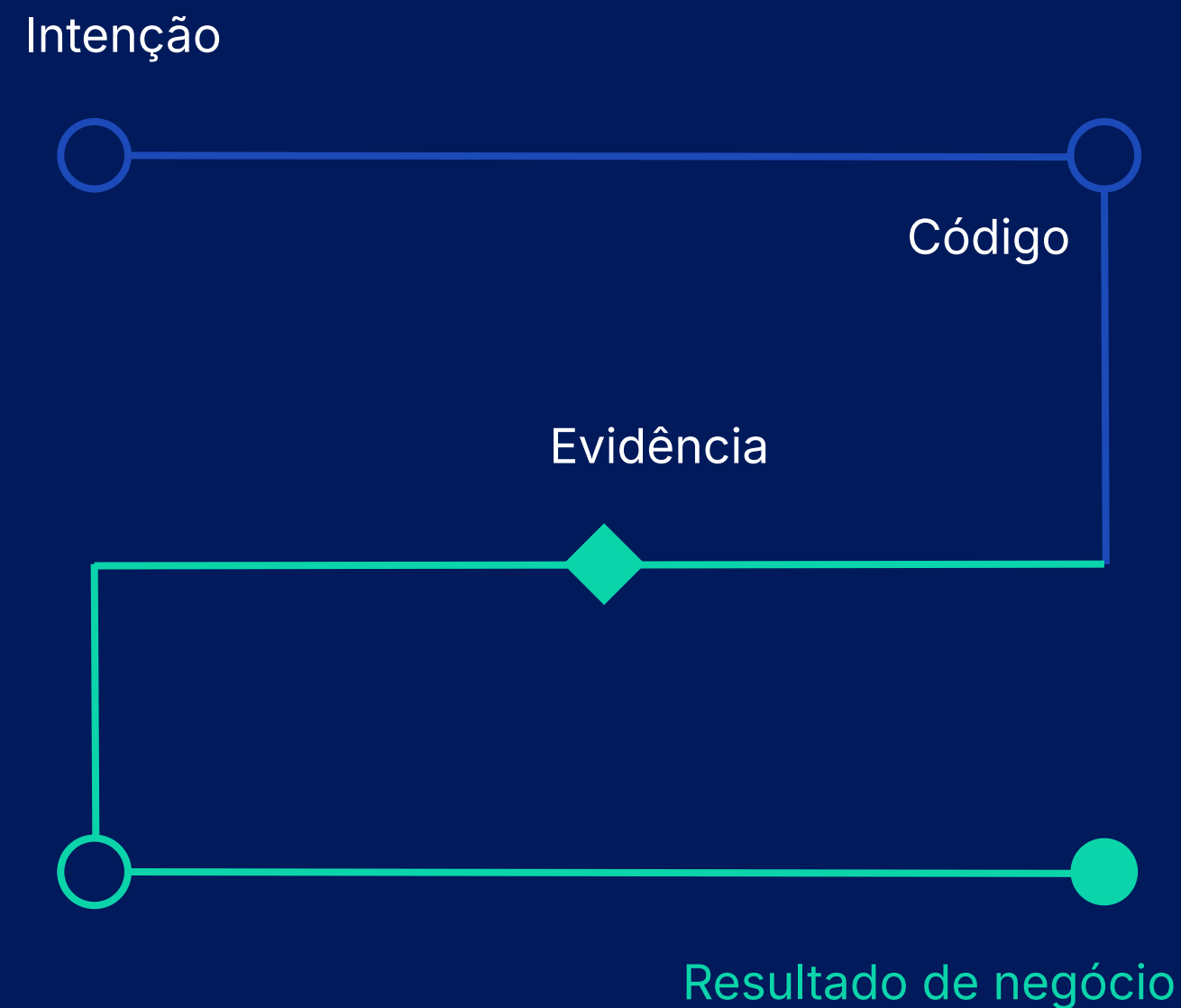


Código barato. Mudança confiável.

Três perguntas antes de escalar agentes de IA.





**João Campista
Bello Junior**

QUEM FALA

Pesquisa no CEIA. Produto na Lugui.



Mestre em IA aplicada à engenharia de software

Ciência da Computação · Universidade Federal de Goiás



Autor dos papers ICTAI 2026 e ICSME 2026

Agentes de IA guiados por especificação no ciclo de software



Cofundador e CTO da Lugui

Onde a pesquisa vira produto



joao@genaibr.ai



[linkedin.com/in/joaocampista](https://www.linkedin.com/in/joaocampista)

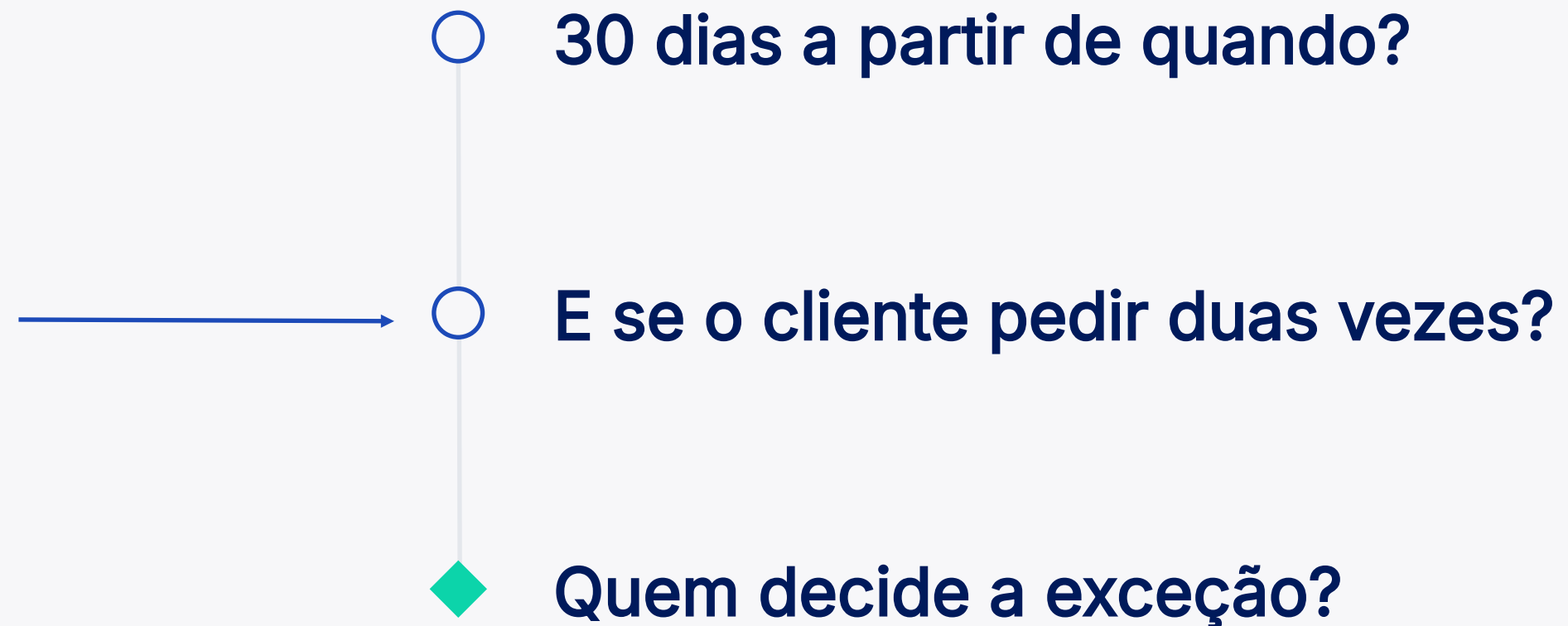
Um botão pronto não é um reembolso resolvido.

"Permitam reembolso em até 30 dias pelo app."

PEDIDO #2048

Pedido entregue

Pedir reembolso



PERGUNTA 1 DE 3

Estamos medindo mudança aceita — ou só atividade?

Onde o ganho some — e para onde vai a conta.

1

Valor e custo

2

Risco e controle

3

Conhecimento e pessoas

4

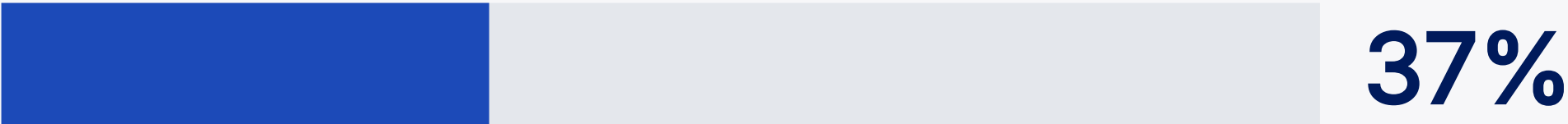
Começar

Quase todo mundo usa. Pouca gente lucra.

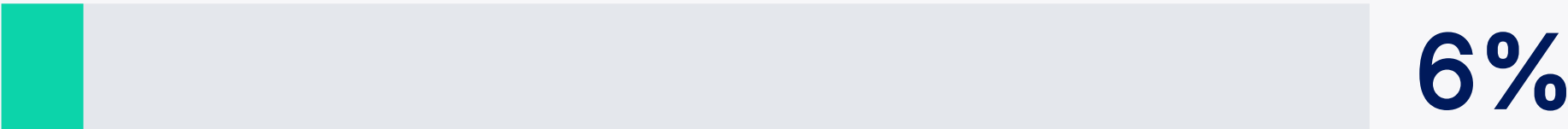
Dizem que ficaram mais produtivos



Veem algum impacto no lucro (EBIT)



Tiram 5% ou mais do EBIT da IA



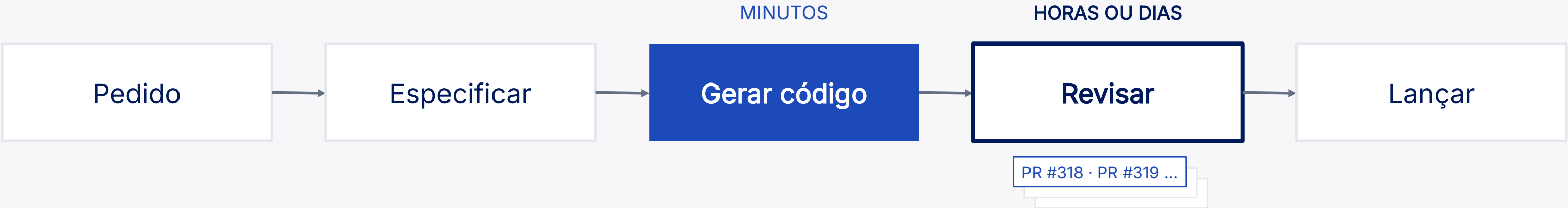
O QUE OS 6% FAZEM DIFERENTE

3 em 4

redesenharam o fluxo de trabalho.

Entre os demais: 1 em 4.

O código acelera. A fila muda de lugar.



17,6 h × 3,4 h

até o primeiro revisor: PR de agente × PR sem IA

33% × 84%

mesclados em 30 dias: com IA × manuais

 NO REEMBOLSO

O botão ficou pronto em uma tarde. A revisão levou dias.

O token ficou barato. A conta, não.

PREÇO DO TOKEN



9x a 900x

mais barato por ano

TOKENS POR PEDIDO



1,5 mil → 6 mil

em um ano

UBER · 2026

4 meses

para gastar o orçamento de IA para
código previsto para o ano inteiro.

Segundo o CTO, em reportagem do The Information

T Token = o pedaço de texto pelo qual o modelo de IA cobra. Lido ou escrito, entra na conta.

O agente paga, sobretudo, para reler.

94%

do volume é o agente relendo histórico, arquivos e instruções.



940 mil tokens relidos

1,2 mil do pedido da pessoa

Sessão de ~1 milhão de tokens: US\$ 0,55

O QUE MAIS MEXE NA CONTA



100×

do modelo barato ao caro



~7×

com vários agentes juntos



Invisível

raciocínio cobrado como texto



+26–45%

em português vs. inglês

Ponha taxímetro no agente — e meça o que ele entrega.

A ARMADILHA · META



60,2 trilhões

de tokens em 30 dias num ranking interno de consumo. **Saiu do ar.**

O QUE FUNCIONA · SHOPIFY



Disjuntor

que corta o acesso quando o gasto do dia dispara.

A MÉTRICA QUE IMPORTA

Custo por mudança aceita

$(\text{tokens} + \text{testes} + \text{revisão humana}) \div \text{mudanças aceitas}$

PERGUNTA 2 DE 3

Se o agente for enganado, o que ele alcança — e quem o para?

O tamanho do estrago — e quem responde por ele.



O estrago é do tamanho do acesso.

JUL 2025 · REPLIT	JUL 2025 · AMAZON Q	AGO 2025 · NX	OUT 2025 · CLAUDE CODE	ABR 2026 · GEMINI CLI
Banco apagado	Versão contaminada	Token roubado	Agiu antes do aviso	Falha nota 10 de 10
FALTOU	FALTOU	FALTOU	FALTOU	FALTOU
Separar teste e produção	Proteger quem publica	Acesso mínimo	Bloquear antes de perguntar	Isolar dos segredos



O agente que escreve o estorno não precisa alcançar o dinheiro.

Agente sem identidade própria é aprovação sem dono.

 NO REEMBOLSO · MERGE #4127 · ESTORNO ILUSTRATIVO

 **Aprovado por Ana Lima**

O QUE O REGISTRO NÃO MOSTRA

Quem clicou foi um agente, na conta da Ana.

71 de 263

merges aprovados só por um agente na conta do líder

ANTES DE ESCALAR

- 1** Identidade própria do agente
- 2** Um responsável com nome
- 3** Aprovação que bloqueia
- 4** Poder de parar

PERGUNTA 3 DE 3

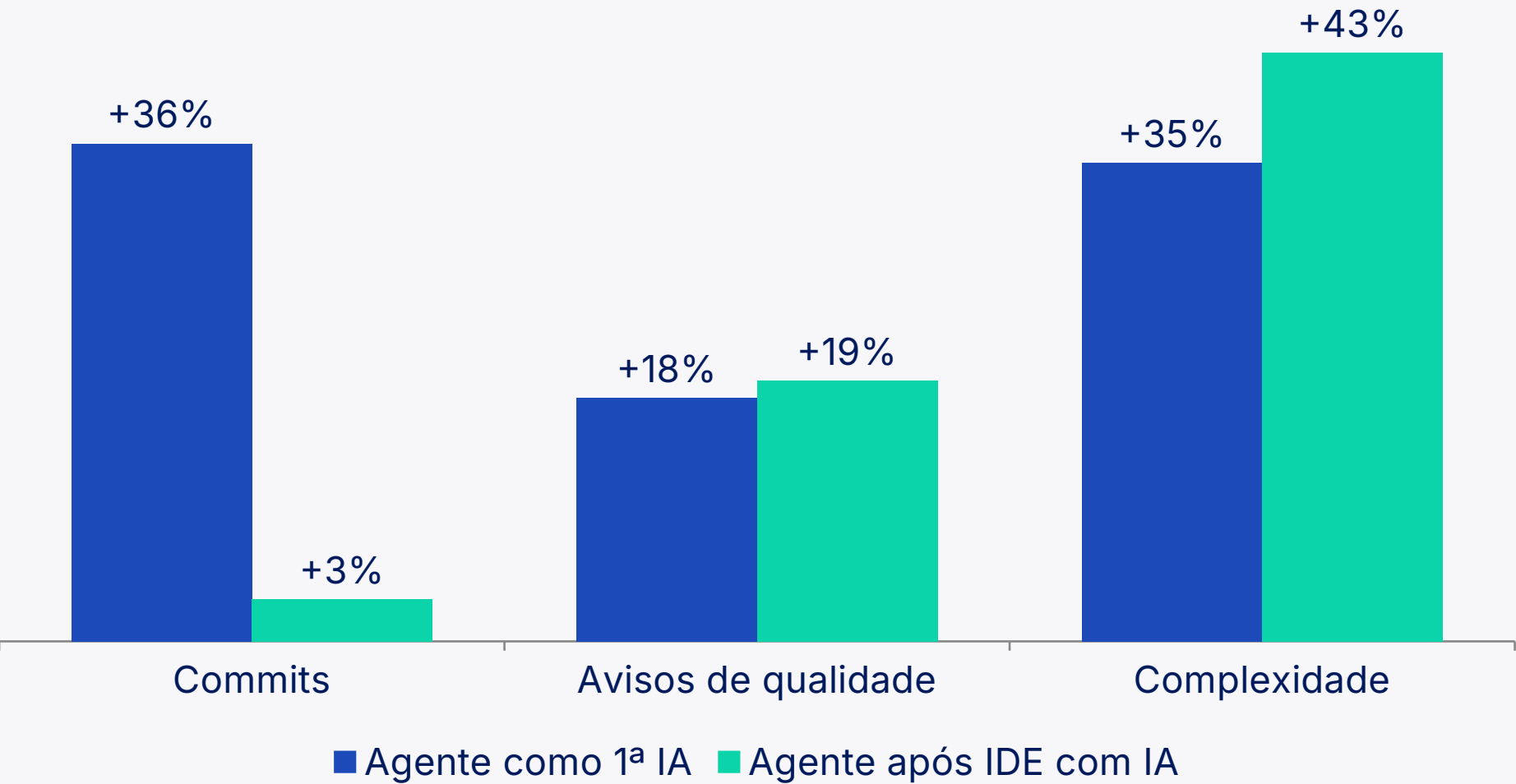
Quem vai entender este sistema daqui a três anos?

O que fica depois que a primeira versão sai.



O ganho é rápido. A complexidade fica.

Repositórios que adotaram agentes, comparados aos que não adotaram



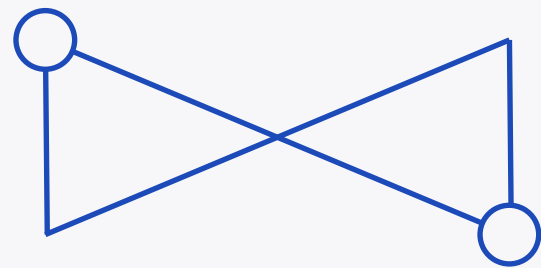
CONTRAPONTO · LOVEHOLIDAYS

80%
dos commits com agente

1-2%
das mudanças falham em produção

O que mudou: uma barreira de qualidade que bloqueia o commit.

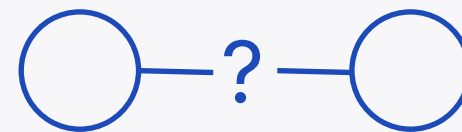
O improviso deixa três dívidas.



Técnica

A regra copiada em três lugares.

→ Barreira de qualidade antes do merge.



Cognitiva

Funciona, mas ninguém sabe explicar.

→ Explicar antes de aceitar.



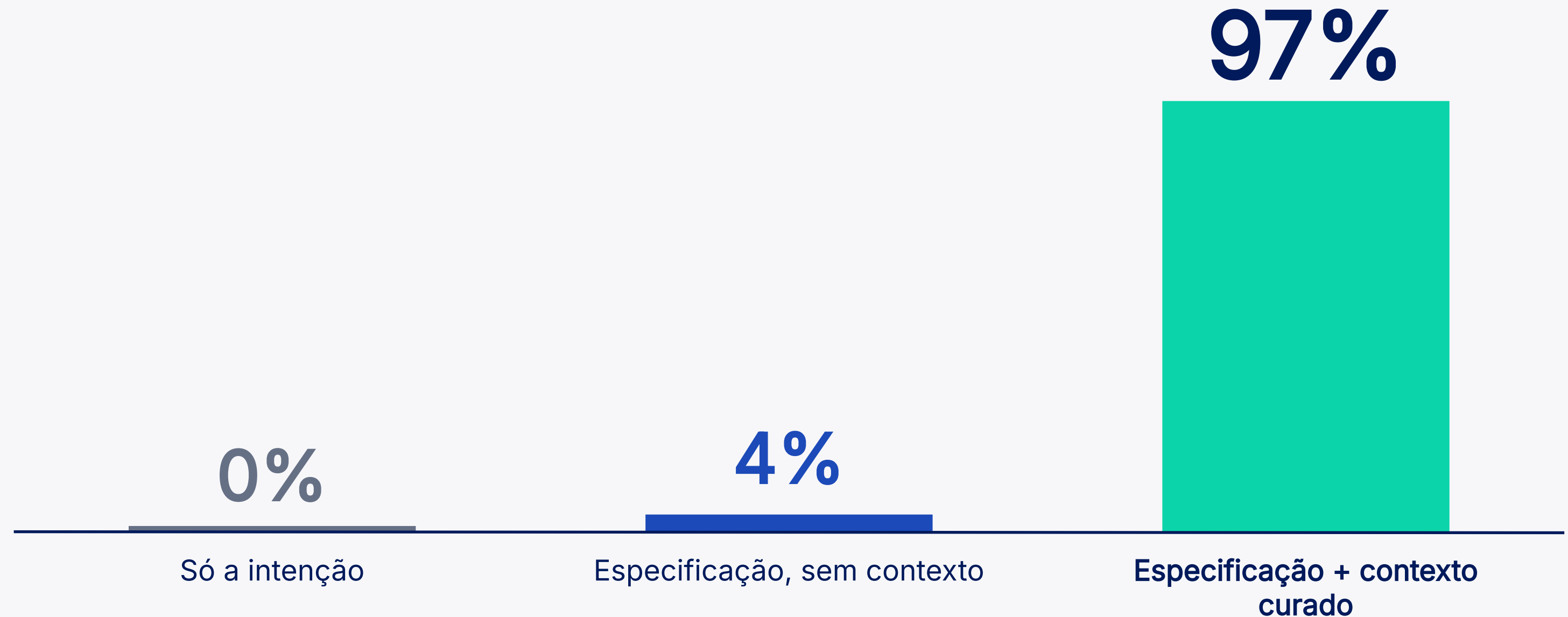
De intenção

O motivo da regra se perdeu.

→ Escrever o porquê antes de o agente agir.

Trocar de modelo não resolve o que a empresa não escreveu.

Mesmo agente. Mesma tarefa. Só mudou a informação.



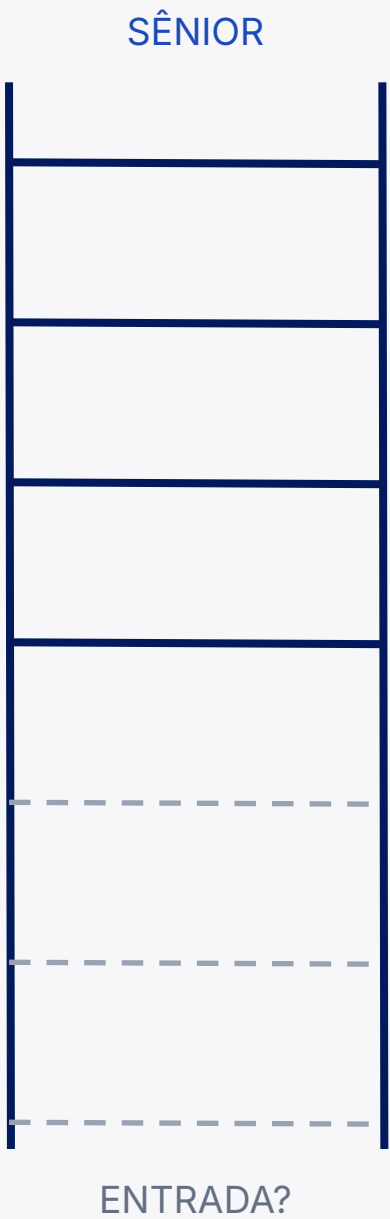
Mais contexto não é melhor contexto.

 NO REEMBOLSO · O QUE O AGENTE PRECISA RECEBER

O QUÊ  Especificação 30 dias contados da entrega.	ONDE E COMO  Contexto Use o serviço de pagamentos que já existe.	POR QUÊ  Registro de decisão Chave única por pedido: o dinheiro não sai em dobro.
---	---	--

 Despejar documentos piora: **18 modelos** perderam desempenho com entradas longas.

Sem vaga de entrada hoje, sem revisor em 2030.



4,5%

das vagas de software nos EUA são de entrada
Indeed Hiring Lab, jul. 2026

50% × 67%

de acerto: aprendeu com IA × sem IA
Anthropic, ensaio randomizado, jan. 2026

77% → 39%

de falhas: IA livre → IA que exige explicação
Explanation Gate, arXiv 2602.20206 · n=78, preprint

Use IA para formar gente, não só para entregar.

AS TRÊS PERGUNTAS, NA PRÁTICA

Por onde começar.

Onde delegar. Até onde soltar. O que perguntar.



Delegue onde dá para verificar e desfazer.



ONDE JÁ FUNCIONA

84% × 55%

PRs de agentes aceitos: rotina × desempenho

45 de 50

limpezas da DoorDash aproveitadas, com pessoa revisando

"O verificador é tudo."

Jussi Rasku, GAISE 2026

Autonomia se ganha com evidência.



No reembolso, comece pelos testes da regra de 30 dias.

SOBE QUANDO
90 dias dentro da régua

DESCE QUANDO
Incidente ou gasto acima do teto

TESTE DE PARADA

Dá para desligar às 3h, sem pedir ao agente?

Três perguntas antes de escalar agentes.

1

Medimos mudança aceita — ou só atividade?

PEÇA

Custo por mudança aceita.

2

Se for enganado, o que ele alcança — e quem o para?

PEÇA

Acesso mínimo e botão de parar.

3

Quem vai entender este sistema daqui a três anos?

PEÇA

Registro de decisão e vagas de entrada.

Segunda-feira: um fluxo (como os testes do reembolso), uma régua, 90 dias.

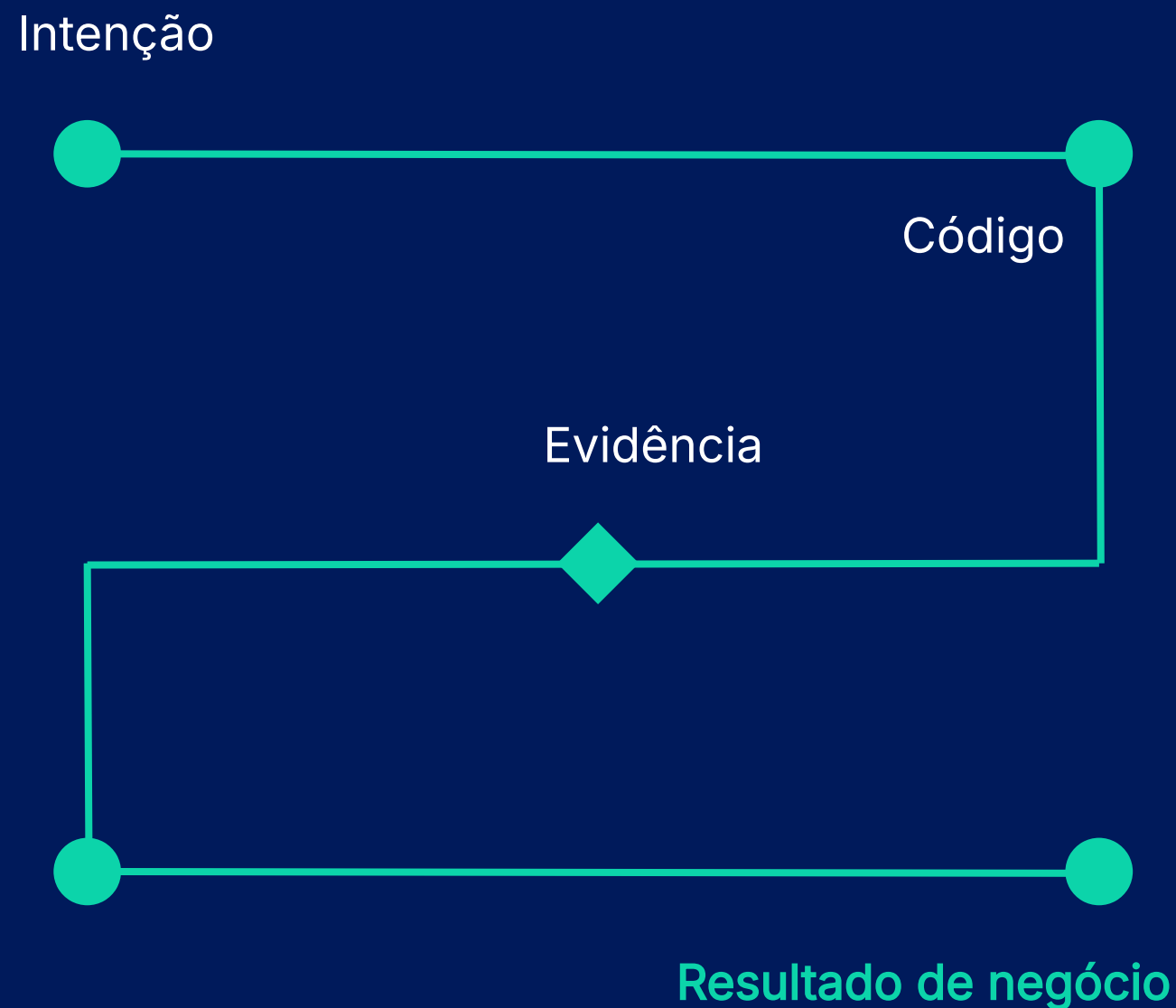
Código barato. Mudança confiável.

Obrigado.

A fonte de cada número está nas duas páginas seguintes.

✉ joao@genaibr.ai

in [linkedin.com/in/joaocampista](https://www.linkedin.com/in/joaocampista)



Fontes para continuar a conversa.

ADOÇÃO, FLUXO E VELOCIDADE

[McKinsey. The state of AI in 2026: On the road to ROI. Ago. 2026.](#)

[BCG. AI at Work: Strategy Matters More Than Tools. Jun. 2026.](#)

[Bain. From Pilots to Payoff: GenAI in Software Development. 2025.](#)

[LinearB. Engineering Benchmarks 2026. Fornecedor de métricas.](#)

[AI IDEs or Autonomous Agents? arXiv 2601.13597. MSR 2026.](#)

[Task-Stratified Analysis of Agent PR Acceptance. arXiv 2602.08915. MSR 2026.](#)

[CodeScene. loveholidays scaled to 80% agent-assisted code. Ago. 2026.](#)

[Storey. From Technical Debt to Cognitive and Intent Debt. arXiv 2603.22106.](#)

[GAISE 2026. GPT-Lab, Tampere University. Programa e keynotes.](#)

CONTEXTO E EVIDÊNCIA

[CEIA/UFG. Specification-Guided Multiagent AI for Context-Governed Software Life Cycle Automation. ICTAI 2026, manuscrito.](#)

[Benchmark Illusion. arXiv 2602.11898. 2026.](#)

[FDE-Gym: The Last Mile of AI Delivery. arXiv 2610.22350. Preprint.](#)

[Tool-Calling LLM Agents Under Realistic MCP Faults. arXiv 2610.32555.](#)

[Chroma. Context Rot. Jul. 2025.](#)

[On the Impact of AGENTS.md Files. arXiv 2601.20404. 2026.](#)

[CodeHealth Predicts the Token Cost of Agentic Code Review. arXiv 2610.21914.](#)

[DoorDash. Automating Feature Flag Cleanup at Scale. arXiv 2610.31351.](#)

Custo, risco e pessoas.

CUSTO DE TOKEN

[How Do AI Agents Spend Your Money?](#) arXiv 2604.22750. 2026.

[Epoch AI. LLM inference price trends.](#) 2025.

[OpenRouter. State of AI.](#) Estudo com 100 trilhões de tokens.

[GitHub. Copilot is moving to usage-based billing.](#) Abr. 2026.

[Forbes. Why your engineers' AI tools are wrecking your budget.](#) Mai. 2026.

[Pragmatic Engineer. Tokenmaxxing as a weird new trend.](#) Abr. 2026.

[Gartner. AI coding costs will surpass average developer salary by 2028.](#) Jun. 2026.

[Anthropic. Pricing](#) · [Claude Code: Manage costs.](#) Out. 2026.

[Hakimov et al. Language Equality has a Price.](#) arXiv 2608.01395.

RISCO, RESPONSABILIDADE E PESSOAS

[Nx. s1ngularity post-mortem.](#) Set. 2025 · [AWS-2025-015.](#)

[Claude Code, alerta de segurança](#) · [Replit.](#)

[Who Approved This? Coding Agents and Merge-Gate Evidence.](#) arXiv 2610.28683.

[Cloud Security Alliance. Agentes vs. ações humanas.](#) Mar. 2026.

[Microsoft Entra Agent ID.](#) Identidade própria para agentes.

[NCSC. Thinking carefully before adopting agentic AI.](#) Mai. 2026.

[Indeed Hiring Lab. The labor market is tilting toward seniority.](#) Jul. 2026.

[Stanford. Canaries in the Coal Mine.](#) 2025.

[Anthropic. AI assistance and coding skills.](#) Jan. 2026.

[Explanation Gate.](#) arXiv 2602.20206. Preprint, 2026.