

Astra Effort — guia completo

Mark Kashef | Data de referência da pesquisa: 8 de setembro de 2026.

Esta edição em português traduz o material original. Os fatos sobre produtos e preços abaixo refletem a pesquisa datada do autor; não foram revalidados nesta tradução. O original está em [original-en/SUPER-GUIDE.md](#).

Este material documenta sete execuções de pesquisa e construção, uma demonstração aritmética separada com três configurações, verificações aprofundadas do produtor, orientações com fontes e testes reutilizáveis.

Comece com Medium. Exija um motivo para aumentar.

Mark Kashef | Sete primeiras tentativas, verificações práticas e prompts prontos para usar.

Medium é minha escolha inicial para esta tarefa de pesquisa e construção. Entregou um fluxo verificado e um plano útil em 30:41. High é a próxima configuração que eu testaria em um trabalho com restrições que interagem entre si.

Este guia apresenta as evidências por trás dessa escolha e um caminho para encontrar a configuração adequada ao seu trabalho. Inclui a tarefa original, sete quadros editáveis, dez prompts reutilizáveis e uma planilha de registro.

Pesquisa encerrada em 8 de setembro de 2026. Os experimentos ocorreram em 7 de setembro.

Escolha a próxima configuração por um motivo.

1. Defina o que significa concluir antes de escolher o esforço. Nomeie o artefato necessário e três verificações que o tornariam útil. "Construa um aplicativo" dá muito menos informação que "altere um preço, aprove, recarregue e preserve o valor".
2. Comece com Medium em um projeto delimitado que misture pesquisa e construção como este. Comece mais baixo em uma tarefa fácil e bem especificada se puder verificar o resultado rapidamente. São hipóteses práticas de partida, não garantias de desempenho.
3. Aumente o esforço para tratar uma falha específica. Experimente High quando restrições que interagem, um defeito difícil ou um raciocínio fraco persistirem após um esclarecimento direcionado. A ausência de uma captura de tela, sozinha, não justifica aumentar todas as configurações.

- 4. Compare o trabalho inteiro. Inclua acompanhamentos, verificação, espera e trabalho dos agentes filhos. Pare quando o resultado passar nas verificações; não persiga um quadro maior ou uma explicação mais elaborada.
- 5. Mantenha velocidade como decisão separada. O experimento usou Standard. A preferência pessoal de Mark por Medium + Fast é um julgamento separado, não um resultado testado neste conjunto.

Quadro comparativo das sete execuções.

A tabela registra estas primeiras tentativas preservadas, não um ranking universal de velocidade. As sete rodaram simultaneamente em uma máquina. Ultra podia delegar; as demais condições não.

Tokens processados = entrada acumulada + saída. A entrada em cache já está incluída na entrada. A maior parte da entrada nestas execuções veio de cache. Não transforme esses totais em comparação de preços.

Nenhuma das sete precisou de um lembrete do produtor para devolver os artefatos. Esse resultado vale para estas instruções e tarefas. Não estabelece se esforço maior resolve interrupções precoces.

Configuração	Tempo	Tokens processados	Parcela da entrada em cache	Fontes	Elementos do quadro
Astra Low	37:37	14.814.481	98,3%	8 Reddit + 0 X	161
Astra Medium	30:41	10.233.386	98,0%	6 Reddit + 0 X	133
Astra High	41:21	14.328.598	98,1%	7 Reddit + 1 X	186
Astra Extra High	45:14	12.286.194	97,8%	6 Reddit + 1 X	155
Astra Max	46:10	10.547.404	97,5%	7 Reddit + 0 X	155
Sol High	32:49	6.698.569	97,5%	6 Reddit + 2 X	215
Astra Ultra	42:10	21.730.368	97,3%	7 Reddit + 2 X	240

Astra Low — Clearshift

Acompanha tarefas de limpeza não realizadas até a correção e a nova verificação.

Para: donos de empresas de limpeza comercial que atendem escritórios regularmente.

Vantagem competitiva proposta: padrões específicos de cada cliente e histórico das correções que funcionaram.

Verificação observada: a correção precisava ser verificada novamente antes de encerrar o serviço. O estado encerrado permaneceu após recarregar.

Lacuna: nenhuma evidência do X. Mais tempo decorrido e tokens processados que Medium.

O que eu aproveitaria: lembre-se de que Low também pode produzir um fluxo completo. Menor esforço não garante um trabalho total menor.

Astra Medium — Scopewell

Calcula preços de serviços extras de limpeza e registra a aceitação do cliente.

Para: donos de empresas de limpeza comercial com contratos de prestação de serviços por escrito.

Vantagem competitiva proposta: registros precisos de escopo e estimativas aprimoradas por serviços reais.

Verificação observada: alterar a mão de obra de 45 para 60 minutos mudou a proposta mensal de \$164 para \$212. Uma aceitação simulada e o novo valor permaneceram após recarregar.

Lacuna: nenhuma evidência do X. O quadro com menos elementos.

O que eu aproveitaria: minha escolha inicial para esta tarefa. Entregou um fluxo verificado e um plano útil no menor tempo observado.

Astra High — Fieldwork

Redistribui visitas quando uma equipe de limpeza não pode trabalhar.

Para: donos de empresas de limpeza residencial que coordenam equipes e agendamentos.

Vantagem competitiva proposta: restrições e resultados específicos da empresa que melhoram o agendamento.

Verificação observada: um serviço impossível de duas horas dentro de uma janela de uma hora foi rejeitado. A redistribuição válida foi salva corretamente e manteve outras quatro visitas intactas.

Lacuna: o deslocamento usou uma margem fixa, sem otimização de rotas em tempo real.

O que eu aproveitaria: uma próxima configuração útil para testar quando há restrições que interagem. O valor veio dos comportamentos que o produto conseguiu tratar.

Astra Extra High — Scopekeep

Leva serviços extras de reforma da aprovação até o faturamento.

Para: donos de pequenas empresas de reforma que perdem o controle de adicionais cobráveis.

Vantagem competitiva proposta: verificações de escopo adaptadas ao histórico do que foi aprovado, faturado e pago.

Verificação observada: uma nova tarifa de mão de obra foi propagada para um total de \$530, para a aprovação e para uma linha pronta para faturar. A revisão 1 permaneceu após recarregar.

Lacuna: mais tempo que High. Seu contraexemplo do X foi lido pelo participante; o produtor não conseguiu reabri-lo de forma independente.

O que eu aproveitaria: leia as evidências contrárias antes de comprar a ideia de negócio. Uma solução já existente com Grok e contabilidade de código aberto colocou o produto proposto em questão.

Astra Max — Fieldnote

Mantém escopo, preço e aprovação de um serviço extra no mesmo registro.

Para: donos de empresas de reforma residencial com equipes de 2–10 pessoas.

Vantagem competitiva proposta: uma rotina adotada no canteiro de obras, configuração assistida e indicações de profissionais de escrituração contábil.

Verificação observada: mão de obra de \$275,25 mais materiais de \$84,75 resultaram exatamente em \$360. A aprovação simulada persistiu.

Lacuna: a execução mais longa. O download de uma exportação nativa não foi verificado de forma independente.

O que eu aproveitaria: a profundidade útil apareceu nos centavos exatos, no estado de aprovação e nas regras de revisão. São verificações melhores que contar detalhes de acabamento visual.

Sol High — RelayOps

Planeja a recuperação quando atrasos afetam uma rota de serviços.

Para: donos que também coordenam operações de empresas de serviços recorrentes com 2–20 funcionários.

Vantagem competitiva proposta: resultados de recuperação, regras mantidas e suporte responsável.

Verificação observada: a progressão entre etapas e o estado resolvido persistiram.

Lacuna: selecionar “Split with Crew 1” não alterou o resumo seguinte: ele continuou mostrando outro plano e números de resultado fixos.

O que eu aproveitaria: mantenha Sol na comparação e teste se as escolhas realmente mudam os resultados. Um consumo processado menor não compensa um caminho de decisão quebrado.

Astra Ultra — Accord

Registra e combina adicionais de reforma antes do início do trabalho.

Para: donos de pequenas empresas de reforma residencial que lidam com adicionais frequentes.

Vantagem competitiva proposta: configuração específica por especialidade e uma rotina comprovada, melhoradas por resultados reais.

Verificação observada: uma nova alteração personalizada de \$200 exigiu uma nota de aprovação. Título, escopo, total e aprovação persistiram.

Lacuna: três agentes filhos acrescentaram 7,12 milhões de tokens processados. É uma condição separada, com delegação habilitada.

O que eu aproveitaria: use delegação quando frentes independentes de pesquisa ou verificação tiverem valor real. Aqui ela produziu pesquisa mais profunda sobre os fluxos dos concorrentes, não um produto drasticamente diferente.

Um resultado bonito pode esconder uma escolha quebrada.

O protótipo do Sol avançou pelas etapas e manteve um estado resolvido. Em uma demonstração rápida, isso poderia parecer aprovação.

O produtor escolheu “Split with Crew 1”, mas o resumo seguinte continuou descrevendo outro plano. O resultado também ficou fixo. A entrada escolhida não determinou a saída posterior.

Teste causa e efeito: escolha outra opção, altere um número, use um novo registro e confira a consequência exata. Recarregue depois. Verifique também se registros não relacionados permaneceram iguais.

Essa verificação isolada é mais útil que contar componentes, telas ou formas do quadro. Aplique-a ao trabalho de qualquer modelo.

Revise o protótipo concluído sem editar seu código-fonte primeiro. Use dados de exemplo novos ou restaurados e registre o estado inicial.

1. Altere uma entrada relevante ou selecione outra opção.
2. Preveja qual total, registro ou resumo posterior deve mudar.
3. Conclua o fluxo principal.
4. Compare a saída final com a escolha realmente feita.
5. Recarregue a página e verifique a persistência.
6. Verifique uma entrada inválida e um registro não relacionado.

Capture evidências de cada resultado. Separe comportamento testado, inspeção de código e comportamento não verificado. Preserve o artefato da primeira tentativa antes de propor correções. Um botão funcionando ou uma captura de tela bonita não bastam para provar que a decisão foi propagada pelo fluxo.

Separe esforço de continuidade até a conclusão.

Esforço muda quanto raciocínio o modelo pode dedicar a uma tarefa. Um escopo incompleto, uma autorização pouco clara ou um bloqueio real de ferramenta ainda podem interromper o trabalho. Aumentar a configuração não resolve todo tipo de pausa.

Dê uma linha de chegada observável e indique quais decisões reversíveis o agente pode tomar. Se quer pesquisa, implementação e verificação, diga isso. Especifique quais ações ainda dependem de sua decisão.

Neste conjunto, as sete execuções devolveram artefatos sem um lembrete do produtor. Não podemos classificar “preguiça” em um conjunto cuja quantidade medida de lembretes não variou.

Teste autonomia separadamente: preserve a primeira tentativa, conte requisitos ausentes e diferencie bloqueios reais de paradas precoces evitáveis. Aplique a mesma política neutra de acompanhamento a cada execução.

Conclua esta tarefa: [TAREFA].

Concluído significa: [ENTREGAS OBSERVÁVEIS E TESTES DE ACEITAÇÃO].

Você pode decidir [ESCOLHAS REVERSÍVEIS] sem me perguntar. Use suposições razoáveis quando elas não mudarem o objetivo; informe brevemente qualquer suposição relevante.

Continue até implementar e verificar o que está neste escopo. Um plano, uma confirmação de entendimento ou uma oferta para continuar não são a entrega final. Se encontrar um bloqueio real, identifique-o, preserve o trabalho e conclua as partes independentes que ainda forem possíveis. Pergunte antes de [AÇÕES ESPECÍFICAS QUE EXIGEM MINHA DECISÃO].

Ao final, mostre o que mudou, as verificações realmente executadas e qualquer requisito ainda não atendido. Não diga que um teste passou se você não o executou.

Faça sua comparação com isolamento adequado.

Use o mesmo texto fixo da tarefa, arquivos iniciais, ferramentas disponíveis, contas conectadas, velocidade e critérios de conclusão. Mantenha cada execução em uma pasta nova. Não deixe a saída de um modelo entrar no contexto de outro.

Confira o modelo e o esforço reais. Uma tarefa chamada "High" é apenas um rótulo. Se seu aplicativo não consegue criar ou configurar tarefas, abra e configure manualmente.

Para um teste de tempo mais justo, rode uma execução por vez ou isole os recursos de execução. O controle compartilhado do computador pode causar interferência em testes simultâneos. Repita em outra ordem antes de fazer uma afirmação forte.

Arquive as primeiras tentativas antes de corrigir. Se possível, avalie o resultado sem ver o rótulo de esforço. Preserve falhas como evidências; não ajude silenciosamente sua configuração preferida.

Crie três tarefas separadas para comparação: "LOW | Meu teste", "MED | Meu teste" e "HIGH | Meu teste". Use GPT-6 Astra com esforço de raciocínio low, medium e high, respectivamente, se houver suporte neste ambiente. Mantenha a mesma configuração de velocidade. Verifique e informe o modelo e o esforço efetivos; os rótulos, sozinhos, não bastam.

Dê às três a mesma tarefa abaixo, em pastas novas e isoladas. Não inclua outras execuções nem suas saídas. Mantenha desativada a delegação opcional. Dê a cada uma até [LIMITE DE TEMPO] para a primeira tentativa. Preserve a saída antes de qualquer acompanhamento. Registre trabalho concluído, requisitos faltantes, bloqueios, tempo decorrido e uso atribuível à tarefa, quando disponível.

Execute em paralelo somente se ferramentas e espaços de trabalho estiverem isolados; execute em sequência qualquer fluxo que compartilhe o controle do computador. Acompanhe o progresso sem orientar os participantes. Se este ambiente não puder criar tarefas ou configurar o esforço, diga exatamente qual etapa devo realizar manualmente. Nunca substitua configurações silenciosamente.

TAREFA:
[COLE UMA ÚNICA VERSÃO FIXA DA TAREFA]

Milhões de tokens não são uma conta a pagar.

O modelo pode ver o mesmo contexto em cache em muitas chamadas. A contagem processada acumulada soma essas leituras. Um total de 20 milhões não significa uma única janela de contexto de 20 milhões de tokens nem 20 milhões de tokens recém-gerados.

Neste conjunto, total = entrada + saída. Entrada em cache é um subconjunto da entrada; saída de raciocínio é um subconjunto da saída. Somar esses subconjuntos de novo duplicaria a contagem.

Os 21.730.368 tokens de Ultra incluem o agente principal e três agentes filhos. Os filhos contribuíram com 7.121.868. Comparar apenas o principal esconderia parte do trabalho.

Um cálculo real de API exige tarifas de entrada sem cache, entrada em cache e saída do modelo, velocidade e faixa de contexto específicos. Uma assinatura do Codex usa suas próprias regras de créditos. Uma alteração na franquia da conta inteira não isola uma tarefa quando outras estão rodando.

O registro completo está em `evidence/results.csv` e `evidence/results.json`. Deixe campos de custo indisponíveis em branco em vez de adivinhar.

Faça a pesquisa argumentar contra a ideia.

“Quais são os melhores insights?” costuma produzir um resumo bem acabado. Uma pergunta mais específica pode encontrar uma contradição útil: quem já resolve isso com uma ferramenta existente, se recusa a pagar ou diz que o problema é outro?

Na pesquisa de Extra High, o fluxo existente de um dono de negócio com Grok e contabilidade de código aberto questionou a necessidade de um novo produto. Foi útil porque mudou o argumento de negócio. O produtor não conseguiu reabrir de forma independente a publicação no X, então a fonte continua identificada como lida pelo participante.

Use Grok ou outro caminho de busca para encontrar publicações diretas e depois examine a evidência original e o contexto ausente. Pergunte qual tarefa falhou, qual configuração foi usada e se alguém testou a correção sugerida.

Um relato pode revelar um modo de falha que merece teste. Não estabelece a frequência da falha nem qual nível de esforço a causou.

Ajude a investigar esta afirmação específica: [AFIRMAÇÃO]. Pesquise no X experiências diretas que possam sustentá-la ou contradizê-la. Priorize descrições concretas das tarefas, capturas de tela com contexto, links para as publicações originais e correções posteriores.

Para cada descoberta, informe URL da publicação original, data, tarefa, configuração quando mencionada, resultado observado e contexto ausente. Separe a opinião do autor da publicação do comportamento verificado do produto. Não transforme curtidas, republicações ou linguagem confiante em prova.

Depois, apresente três perguntas mais específicas que ajudem a explicar por que os usuários obtiveram resultados diferentes. Se não conseguir acessar a evidência original, diga isso.

Uma vantagem competitiva precisa sobreviver à cópia.

O prompt perguntou explicitamente o que cliente ou concorrente poderiam copiar em um fim de semana, o que ainda faltaria e por que alguém pagaria. Essa pergunta é mais difícil que pedir uma lista de funcionalidades.

Os sete produtos ficaram em serviços de campo. O cliente compartilhado era um pequeno empresário de serviços com 2–20 funcionários, o que já restringia a busca. A convergência não prova que esses setores sejam o melhor mercado para todos.

Registros armazenados, uma interface melhor e modelos gerados por IA são fáceis de apresentar como vantagens defensáveis. A pergunta útil é se um hábito operacional real, um histórico de resultados usado com permissão, a distribuição ou um serviço confiável de configuração geram valor suficiente para reter clientes.

Todas as vantagens competitivas destas saídas são propostas. Nenhum protótipo demonstrou adoção paga ou vantagem duradoura. Defina um pequeno teste de refutação antes de construir o sistema maior.

Suponha que um concorrente competente e meu cliente consigam reproduzir a interface e a lógica básica do software em um fim de semana. Coloque este produto à prova: [PRODUTO].

Separe o que é facilmente copiável de qualquer vantagem que precise ser conquistada. Explique o valor para o cliente no primeiro dia, antes de existir uma vantagem defensável (moat). Identifique como conquistar a primeira vantagem a partir do zero, por que um concorrente estabelecido ainda poderia nos superar e o que o cliente poderia usar no lugar.

Planeje um pequeno experimento com o primeiro cliente, com um limiar explícito de aprovação ou reprovação. Identifique a capacidade de defesa proposta como hipótese. Não chame um banco de dados genérico, uma interface sobre IA ou uma lista de funcionalidades de vantagem competitiva comprovada.

272 mil é um limite por requisição.

Segundo as notas de preços consultadas pelo autor, em prompts da API Astra acima de 272.000 tokens de entrada, a requisição inteira usa tarifas de entrada e cache de 2× e tarifas de saída de 1,5×. A regra não se limita aos tokens excedentes, e nem toda tarifa dobra.

Uma tarefa pode processar milhões de tokens em chamadas repetidas sem que nenhuma chamada ultrapasse esse limite. Observe a maior contagem individual de entrada, não o total exibido no fim do projeto.

Esse limite da API não estabelece um desconto equivalente na assinatura do Codex. Antes de alterar uma configuração TOML, confira o cliente instalado, a forma de autenticação, a configuração suportada e o valor atual. A compactação também pode reduzir detalhes retidos.

Fonte do original: notas de preços do modelo Astra, OpenAI, consultadas em 8 de setembro de 2026.

Audite as configurações de contexto e cobrança deste fluxo sem alterá-las. Identifique se ele usa login do ChatGPT ou cobrança de API. Para chamadas de API, informe a maior contagem individual de tokens de entrada e se alguma requisição excedeu 272.000 tokens de entrada. Separe isso dos tokens acumulados da tarefa. Se uma alteração de configuração ajudar, mostre primeiro a opção suportada, o valor atual e o diff proposto.

Fast custa mais. Esforço é outro controle.

Conforme a documentação consultada pelo autor, com login do ChatGPT, Astra Fast consome créditos à taxa de 2,5× Standard onde está disponível. É um multiplicador de créditos. Não promete que uma tarefa terminará 2,5 vezes mais rápido.

A API tem estrutura de preços separada: Astra Fast usa 2× as tarifas aplicáveis por token de API. Não aplique o multiplicador de assinatura à cobrança da API.

Mantenha o esforço fixo ao testar velocidade. Experimente uma tarefa representativa em Standard e Fast, meça a experiência e use dados reais de cobrança atribuíveis, se disponíveis. As sete execuções usaram Standard e não mediram essa troca.

No Codex CLI, `/fast status` consulta a configuração; `/fast off` e `/fast on` a alteram. Confirme a opção correspondente no seu aplicativo.

Fontes do original: documentação de velocidade do Codex e notas do modelo Astra na API, consultadas em 8 de setembro de 2026.

Inspecione meu modelo, esforço de raciocínio, modo de velocidade e forma de autenticação. Explique a relação de uso aplicável com base na documentação oficial atual. Mantenha o esforço de raciocínio inalterado e mostre como mudar somente a velocidade para comparar a mesma tarefa em Standard e Fast.

Aumente o esforço na fase difícil.

O guia original descreve suporte a um item de entrada `configuration_update` na API Astra para mudar o esforço de raciocínio preservando a configuração original da requisição e o prefixo em cache. Isso permite rascunhar em low e depois aumentar o esforço para uma revisão difícil na mesma conversa.

Coloque o item antes da próxima mensagem do usuário e mantenha o `reasoning_effort` da requisição no valor original. A atualização persiste até ser substituída. Após compactar, adicione uma nova atualização com o valor desejado.

É um recurso de desenvolvimento para Astra em modo padrão de agente único, não uma frase mágica que muda a configuração de uma conversa do Codex. O `reasoning.effort` da resposta continua informando a configuração da requisição; esse campo sozinho não comprova o esforço efetivo.

Histórico compartilhado é útil em produção. Não é uma comparação independente e isolada entre níveis de esforço.

Fonte do original: OpenAI, alteração de raciocínio durante a conversa, consultada em 8 de setembro de 2026.

```
{
  "type": "configuration_update",
  "reasoning": { "effort": "high" }
}
```

Depois, peça uma revisão difícil específica, por exemplo:

Revise esta migração em busca de perda de dados, falhas de concorrência e lacunas de reversão. Para cada risco relevante, aponte a parte correspondente do plano. Proponha uma verificação ou alteração concreta dentro do escopo combinado.

Pergunte o que interrompeu o trabalho. Delimite a delegação.

O guia da OpenAI citado pelo autor descreve sensibilidade de Astra a skills e instruções em `AGENTS.md`. Quando ocorrer uma pausa desnecessária, pergunte qual instrução exata ou fato ausente a causou. Isso oferece algo concreto para corrigir.

Se a tarefa se beneficiar de delegação, dê a cada agente uma pergunta diferente e delimitada. Evidências de clientes, alternativas existentes e objeções mais fortes são frentes separadas úteis. O agente principal mantém a decisão e verifica o resultado integrado.

Os três especialistas de Ultra acrescentaram 7,12 milhões de tokens processados nesta execução. Registre seus modelos e consumo. Mais agentes não são automaticamente mais baratos, e pesquisa duplicada pode consumir o benefício esperado.

Fontes: orientações da OpenAI citadas no original sobre seguir instruções e delegar a subagentes; revisão preservada de Ultra pelo produtor.

Se estiver bloqueado, identifique o fato exato ausente ou a instrução exata e o arquivo que causam a parada. Conclua primeiro o trabalho independente já autorizado.

Se o trabalho em paralelo ajudar, use até três agentes com perguntas distintas e entregas apoiadas em fontes diretas. Evite pesquisas duplicadas. Mantenha a síntese e a verificação final com o agente principal. Se não houver trabalho independente útil, continue sem delegar.

Uma tarefa simples. Três respostas idênticas.

Também preservamos as três tarefas pequenas usadas para demonstrar a abertura de conversas de comparação durante a gravação. Low, Medium e High receberam a mesma pergunta aritmética sem ferramentas. As três devolveram a resposta correta de \$8 no formato solicitado de duas linhas.

O turno registrado de Low levou 4.641 ms; Medium, 5.461 ms; High, 5.130 ms. Rodaram simultaneamente, com inícios separados por menos de um segundo. As pequenas diferenças são sensíveis ao agendamento e não estabelecem um ranking confiável de velocidade.

Neste problema fácil, esforço maior não trouxe benefício visível de resposta ou formatação. É um motivo útil para testar esforço menor quando a tarefa é simples e a resposta é fácil de conferir.

Essas demonstrações curtas são separadas das sete execuções de pesquisa e construção. O uso de tokens não estava disponível nos resumos recuperados. Prompt, resposta e dados completos estão em `evidence/BONUS-ARITHMETIC-DEMO.md` e `.json`.

Conclua esta tarefa simples sem ferramentas: um caderno custa \$4 e uma caneta custa \$2. Maya compra 3 cadernos e 5 canetas e paga com \$30. Quanto deve receber de troco? Responda com exatamente duas linhas: o cálculo e depois a resposta.

As três respostas, traduzidas:

$\$30 - (3 \times \$4 + 5 \times \$2) = \8

Maya deve receber \$8 de troco.

A tarefa exata do vídeo.

O enunciado completo foi traduzido em [experiment/ASSIGNMENT.txt](#) e reproduzido na seção “Enunciado completo” abaixo. A versão inglesa inalterada, efetivamente apresentada aos modelos no vídeo, está em [original-en/experiment/ASSIGNMENT.txt](#). As condições de execução e o quadro de contexto anotado são arquivos separados em `experiment/`.

Guarde a evidência. Depois melhore.

Escolha uma tarefa real que você já precise terminar. Escreva três verificações de aprovação ou reprovação. Execute Medium uma vez, guarde o resultado e teste. Se falhar em alguma verificação, tente um acompanhamento direcionado antes de concluir que faltou esforço.

Para comparar, use o CSV vazio em `experiment/YOUR-RUN-LOG.csv`. Registre configurações efetivas, velocidade, delegação, resultados da primeira tentativa e acompanhamentos separadamente. Guarde capturas de tela ou arquivos de saída junto de cada registro.

Abra os quadros originais no Excalidraw para examinar sua estrutura e suas fontes. Use as sete imagens comparativas como referência visual rápida. Todo o material essencial deste kit funciona offline.

Material complementar do vídeo: <https://astra-field-guide.markkashef.chatgpt.site/>

Você está escolhendo uma configuração que justifique seu lugar no fluxo de trabalho. Comece pelo resultado útil. Deixe a falha indicar o que mudar.

Fontes e medição.

Os fatos oficiais sobre os produtos foram consultados pelo autor em 8 de setembro de 2026. `SOURCE-NOTES.md` contém pesquisa mais completa, casos de uso e links. Comportamentos, disponibilidade e preços podem mudar.

Dados do experimento: sete primeiras tentativas preservadas de 7 de setembro. A entrada e a saída acumuladas de cada execução foram conciliadas de forma independente. A duração veio do primeiro evento de conclusão. A máquina compartilhada e as escolhas diferentes de produto limitam conclusões causais sobre tempo e esforço.

Verificações do produtor: os caminhos reais de interação descritos em `evidence/PRODUCER-CHECKS.md`. São inspeções mais profundas das sete saídas originais, não novas execuções independentes dos modelos.

As sugestões de configuração e os prompts reutilizáveis são recomendações práticas. O experimento não estabeleceu adoção paga, vantagens competitivas duradouras nem superioridade universal.

- [Orientações para Astra: instruções e delegação](#)
- [Modelo Astra e notas de preços da API](#)
- [Velocidade e uso de créditos no Codex](#)
- [Responses API: mudar o raciocínio durante a conversa](#)
- [Cache de prompts: medição e condições](#)
- [Uso do computador: continuar e verificar o resultado](#)

- [Publicação comparativa de Tibo](#)
- [Complemento interativo do vídeo](#)

Enunciado completo

Encontre uma oportunidade de SaaS que valha a pena para donos de pequenas empresas de serviços com 2–20 funcionários. Ainda não escolhi a categoria de negócio, o problema ou o produto. Comece pelos problemas deles, considere três oportunidades e escolha uma.

Pesquise conversas reais no Reddit e no X e examine três alternativas existentes. Busque sustentar a lista de oportunidades com seis conversas distintas e relevantes nessas plataformas. Procure frustrações recorrentes, soluções improvisadas atuais, sinais de gasto e evidências contrárias. Escolha seus métodos de pesquisa com os recursos disponíveis. Inclua links para as fontes originais. Explique limitações de acesso e lacunas de evidência; não invente fontes para atingir uma contagem. Popularidade, sozinha, não prova que alguém pagará.

Suponha que clientes e concorrentes possam usar Astra para reproduzir rapidamente um software competente. Explique o que poderia ser copiado em um fim de semana, qual parte valiosa ainda faltaria e por que o cliente pagaria em vez de construir ou trocar de solução. Proponha um caminho plausível para conquistar essa vantagem a partir do zero, incluindo uma estratégia para o primeiro cliente e um experimento que possa refutá-la. O produto deve entregar valor antes de essa vantagem existir. Trate a capacidade de defesa como hipótese, não como uma vantagem competitiva já estabelecida.

Crie um quadro editável no Excalidraw com zonas distintas e conectadas para evidências dos clientes, alternativas e oportunidade, jornada principal do usuário, telas do produto e direção visual, arquitetura e fluxo de dados, e escopo do MVP com sequência de construção. Use uma combinação útil de imagens, diagramas e anotações concisas, com links para as fontes.

Construa um site bem acabado que reúna um plano de produto explorável e um protótipo funcional em uma única experiência. Atenda a um perfil principal de usuário e um problema central. Implemente um fluxo completo com 3–5 etapas relevantes e no máximo três telas principais do produto. Use dados de exemplo realistas e preserve alterações após recarregar a página. Mantenha pesquisa e planejamento em painéis de apoio. Exclua autenticação de produção, pagamentos, mensagens ao vivo e integrações externas; identifique claramente os comportamentos simulados.

Entregue o protótipo funcional, o quadro editável do Excalidraw, o site de planejamento e as fontes de pesquisa. Verifique o fluxo principal e explique o que ainda não foi comprovado. Tome as decisões de produto e implementação dentro deste escopo. Não contate pessoas, publique mensagens, compre serviços ou faça implantação pública. Não alegue pesquisa ou verificação que não realizou.

Próximos arquivos

Veja [PROMPTS.md](#) para dez modelos prontos para usar; [SOURCE-NOTES.md](#) para oito complementos pesquisados; e [evidence/results.csv](#) para os dados exatos.