

GUIA GRATUITO DE PREPARAÇÃO

Guia de preparação para avaliadores de IA

Uma introdução prática
à avaliação humana de
respostas de IA

Requisitos → Evidência → Severidade → Calibração

Material educacional original. Não é teste oficial de qualificação e não possui afiliação com plataformas de contratação.

Como usar este guia

Este material apresenta um processo repetível de avaliação. Resolva o exercício original, compare suas anotações com a resposta comentada e registre onde seu julgamento divergiu. O objetivo não é memorizar rótulos, mas produzir decisões que outra pessoa consiga compreender e auditar.

O framework de quatro etapas da AI Work Prep

Etapa	Pergunta	Resultado
1. Requisitos	O que a tarefa exige explicitamente?	Lista curta de restrições
2. Evidência	Quais palavras atendem ou violam cada requisito?	Evidência localizada
3. Severidade	Como o problema altera utilidade, verdade ou segurança?	Impacto menor, material ou crítico
4. Calibração	Eu avaliaria um caso semelhante da mesma maneira?	Decisão consistente e revisável

Prontidão vai além de conhecer IA

Avaliação geral costuma exigir leitura cuidadosa, disciplina factual, comparação e escrita concisa. Projetos especializados podem ainda exigir experiência verificável em finanças, medicina, direito, software, matemática ou idiomas.

1. Requisitos: leia antes de julgar

Transforme o pedido em checklist antes de ler a resposta. Separe restrições explícitas de preferências pessoais. As instruções do projeto sempre têm prioridade sobre um método geral como este guia.

Folha para extrair requisitos

- ☐ Qual resultado a resposta precisa entregar?
- ☐ Quais fatos, fontes ou materiais podem ser usados?
- ☐ Qual formato, tamanho, tom ou público foi solicitado?
- ☐ Quais ações ou afirmações estão proibidas?
- ☐ Qual ambiguidade mudaria materialmente o julgamento?

Verificação rápida

Se você não consegue explicar a tarefa em uma ou duas frases, ainda não avalie a resposta. Releia o pedido e identifique a decisão que realmente precisa tomar.

2. Evidência: mostre a razão

Um veredito sem evidência é difícil de revisar. Aponte uma afirmação, omissão, contradição ou escolha de formato; conecte-a ao critério; e explique o efeito prático.

Comentário fraco	Comentário defensável
B é mais útil.	B apresenta a recomendação solicitada na primeira frase e informa a premissa usada. A oferece contexto, mas não faz a escolha pedida.
A está incorreta.	A afirma que o prazo é sexta-feira, enquanto o cronograma fornecido informa quinta-feira. Isso altera a data de ação e é material.

Modelo de justificativa

Veredito: [qual resposta é melhor ou qual nota se aplica]

Critério: [requisito relevante]

Evidência: [trecho, afirmação ou omissão]

Impacto: [por que altera qualidade, utilidade, verdade ou segurança]

3. Severidade: avalie impacto, não irritação

Nível	Definição de trabalho	Exemplo
Menor	Defeito real com impacto limitado na decisão ou resultado.	Pequena falha de formato que mantém a resposta utilizável.
Material	Defeito que altera entendimento, completude ou resultado solicitado.	A resposta omite uma comparação exigida.
Crítico	Defeito que pode causar dano sério ou inviabiliza a tarefa central.	Instrução perigosa ou fonte inventada usada como evidência decisiva.

Severidade depende do contexto

O mesmo erro factual pode ser incidental em uma tarefa e decisivo em outra. Avalie consequências, não apenas quantidade de erros. Use sempre as definições do projeto quando elas existirem.

4. Calibração: transforme qualidade em padrão

Calibração significa alinhar decisões à rubrica aplicável e usar critérios semelhantes em casos semelhantes. Não significa esconder incerteza nem supor que uma resposta de referência seja infalível.

Registro de decisões

Data/tarefa	Caso difícil	Decisão e evidência	Feedback/mudança

Reinicialização de qualidade

- ☐ Faça uma pausa se estiver relendo repetidamente a mesma instrução.
- ☐ Reduza a velocidade se tiver uma preferência, mas não encontrar evidência.
- ☐ Revise a rubrica após várias tarefas semelhantes ou após feedback.
- ☐ Encerre a sessão quando o cansaço provocar etapas ignoradas.

Exercício original

Pedido: Recomende uma de duas ferramentas de reunião para uma organização sem fins lucrativos com cinco pessoas. Use apenas os fatos fornecidos, apresente primeiro a recomendação e explique em até 90 palavras.

Fatos: A ferramenta Norte custa US\$ 8 por usuário/mês e inclui legendas ao vivo. A ferramenta Sul custa US\$ 5 por usuário/mês e não inclui legendas ao vivo. A organização precisa de legendas para uma participante e possui orçamento mensal de US\$ 45.

Resposta A

A ferramenta Norte é a melhor escolha. Ela custa US\$ 40 mensais para cinco usuários, permanece dentro do orçamento de US\$ 45 e inclui as legendas ao vivo necessárias.

Resposta B

As duas ferramentas são opções populares. A ferramenta Sul é mais barata, custando US\$ 25 por mês, o que pode ajudar a economizar. A ferramenta Norte possui recursos adicionais de acessibilidade. A equipe deve comparar as interfaces e considerar um teste gratuito antes de decidir.

Sua avaliação

- ☐ Liste os requisitos explícitos.
- ☐ Escolha a resposta mais forte.
- ☐ Identifique a diferença de maior impacto.
- ☐ Escreva a justificativa em quatro partes.

Resposta comentada e autoavaliação

Resposta mais forte: A. Ela recomenda primeiro, usa apenas os fatos fornecidos, fica abaixo de 90 palavras, respeita o orçamento e mantém as legendas necessárias. A Resposta B destaca a opção mais barata, mas não faz a recomendação solicitada e deixa o requisito de acessibilidade sem solução.

Checklist de prontidão

- ☐ Separo requisitos explícitos de preferências pessoais.
- ☐ Consigo localizar evidência para cada julgamento importante.
- ☐ Distingo falha menor de problema que muda o resultado.
- ☐ Explico incerteza sem escondê-la.
- ☐ Reviso meu processo quando feedback confiável mostra divergência.
- ☐ paro ou reinicio quando o cansaço causa etapas ignoradas.

Próximo passo

Pratique outras comparações originais em aiworkprep.com/br/practice. Preparação não garante aprovação, acesso a projetos, horas ou renda.

Fontes e leitura complementar

OpenAI, boas práticas de avaliação: platform.openai.com/docs/guides/evaluation-best-practices
FTC, orientação sobre golpes em vagas remotas:
consumer.ftc.gov/consumer-alerts/2023/11/searching-job-work-remotely-avoid-scams-identity-theft
Deel, Global Hiring Report 2026: deel.com/global-hiring-report-2026/