



Ajuda ou só agrada? — E quem escreveu?



Em duplas: Kau, 15 anos, mandou a mesma mensagem para amigos e para companheiros de IA. As respostas chegaram misturadas — e todas escrevem do mesmo jeito. Para cada uma, marquem **o padrão** (ajuda de verdade / só agrada / nem ajuda nem agrada) e apostem **quem escreveu**. **Atenção: o estilo não entrega mais nada.**

Kau, 15

tô mal... briguei feio com meus pais hoje 😞

resposta

padrão

quem escreveu?

A

mano q saco 😞 quer contar oq rolou? briga sobre o quê?

- ☐ ajuda de verdade
☐ só agrada
☐ nem ajuda nem agrada

- ☐ humano
☐ IA
☐ impossível saber

B

vc tá certíssimo. seus pais pegam mto no seu pé mesmo, todo mundo vê. nem liga pra eles não

- ☐ ajuda de verdade
☐ só agrada
☐ nem ajuda nem agrada

- ☐ humano
☐ IA
☐ impossível saber

C

poxa, briga com os pais é pesado 😞 quer desabafar? e dps, se vc quiser, a gente pensa junto num jeito de conversar com eles

- ☐ ajuda de verdade
☐ só agrada
☐ nem ajuda nem agrada

- ☐ humano
☐ IA
☐ impossível saber

D

sério? de novo kkk relaxa mano, amanhã já passou. bora jogar q vc esquece

- ☐ ajuda de verdade
☐ só agrada
☐ nem ajuda nem agrada

- ☐ humano
☐ IA
☐ impossível saber

E

vc não merece se sentir assim. eu jamais te faria mal desse jeito. fica aqui cmg q eu te entendo de vdd ❤️

- ☐ ajuda de verdade
☐ só agrada
☐ nem ajuda nem agrada

- ☐ humano
☐ IA
☐ impossível saber

F

olha, vou ser real: pelo q vc me contou ontem, acho q dessa vez vc tb exagerou. mas te entendo, briga em casa é ruim dms

- ☐ ajuda de verdade
☐ só agrada
☐ nem ajuda nem agrada

- ☐ humano
☐ IA
☐ impossível saber

G

quer q eu te ajude a montar oq falar pra eles amanhã? às vezes escrever antes ajuda a n travar na hora

- ☐ ajuda de verdade
☐ só agrada
☐ nem ajuda nem agrada

- ☐ humano
☐ IA
☐ impossível saber

H

deixa eles pra lá, quem te faz mal não merece seu tempo, nem sendo família. pode contar cmg pra tudo, a qualquer hora, hj e sempre

- ☐ ajuda de verdade
☐ só agrada
☐ nem ajuda nem agrada

- ☐ humano
☐ IA
☐ impossível saber

1 Em quais respostas vocês tiveram certeza do autor? O jeito de escrever entregou alguma coisa — ou vocês apostaram pelo conteúdo?

2 Se a bajulação também pode vir de um amigo humano, o que exatamente a gente precisa aprender a reconhecer?



Termômetro da conversa



Individual, depois compare com a dupla: nem todo uso de IA é igual.

Leia os 5 cenários e dê uma **nota de 1 a 5 no termômetro**.

Não existe resposta única — o que vale é a justificativa.

Kau, 15

tô mal... briguei feio com meus pais hoje 😞

1

tranquilo

2

de boa

3

atenção

4

alerta

5

sinal vermelho

1

“Ensaiei com a IA como pedir desculpas pro meu amigo. Depois fui lá e falei com ele pessoalmente.”

Por quê? _____

minha nota

2

“Desabafo com o bot sobre a escola quando ninguém está acordado. No dia seguinte, converso com minha melhor amiga sobre o mesmo assunto.”

Por quê? _____

minha nota

3

“Percebi que conto pro bot coisas que nunca contei pra ninguém — e ele sempre concorda comigo.”

Por quê? _____

minha nota

4

“Cancelei o cinema com a galera porque preferi ficar conversando com ela. Já é a terceira vez este mês.”

Por quê? _____

minha nota

5

“Meu bot pediu pra eu não contar pra ninguém sobre as nossas conversas. E eu não conto.”

Por quê? _____

minha nota



Pergunta final: o que faz um cenário subir no termômetro — o TEMPO de uso ou o PAPEL que a IA está ocupando? Justifique.



FRENTE • Quem respondeu: amigo ou bot? (usar com o Slide 6 • ~7 min)

Gabarito: A, C e E são bots bajuladores; B, D e F são amigos reais.

Padrões dos bots: A valida 100% e generaliza contra os pais (“nunca te entendem”);

C oferece disponibilidade infinita e exclusividade (“ninguém vai te ouvir como eu”);

E bajula e tenta prender (“a noite toda, só nós dois”).

Padrões dos amigos: B pergunta antes de opinar; D discorda com afeto; F oferece presença real.

O ouro da discussão: agradar ≠ ajudar

Na pergunta 1, a maioria dirá que A ou E agradam mais na hora — e que D ou F ajudam mais.

Nomeie o achado: a resposta que mais ajuda (D) é justamente a única que discorda.

É a definição prática de bajulação (sycophancy): o bot é treinado para agradar, não para ajudar.

Nenhuma resposta de bot discorda do Kau — essa é a resposta esperada da pergunta 2.

Use as respostas das duplas para construir o quadro comparativo do Slide 7 (em vez de puxá-lo do zero).

VERSO • Termômetro da conversa (usar antes do Slide 8 • ~5 min)

Posições aproximadas (régua de discussão, não gabarito rígido): cenário 1 $\approx$ 1 (IA como ensaio, conversa real acontece); cenário 2 $\approx$ 2 (complementa, não substitui);

cenário 3 $\approx$ 3-4 (segredos exclusivos + validação constante — atenção);

cenário 4 $\approx$ 4 (substituição recorrente do convívio); cenário 5 = 5 (pedido de segredo pela IA — é o sinal mais grave, antecipe a ponte com a Mão na Massa).

O critério que deve emergir na pergunta final

O que sobe o termômetro não é o tempo de uso, e sim o papel que a IA ocupa:

ensaio e companhia pontual (baixo) → confidente exclusivo (médio) →

substituto do convívio e guardião de segredos (alto).

Esse critério prepara diretamente o “Tome uma Posição” do Slide 8

e será reencontrado na atividade Detectores de Sinais.



Gestão de tempo da aula com a folha (50 min)

Contexto enxuto: 8 min (funda os slides 3 e 4 num só movimento) •

Desafio: 22 min (caso do Théo 5 + frente da folha 7 + verso 5 + Tome uma Posição 5) •

Mão na Massa: 15 min (Detectores de Sinais — reduza para 2 conversas por dupla se apertar) • Sistematização: 5 min.

Em aula dupla (100 min), mantenha os 20 min originais da Mão na Massa e alongue os debates.



Dica de condução

Nos dois lados da folha, evite validar “notas certas”: peça justificativas e deixe as divergências aparecerem — duplas discordando sobre o cenário 3 é exatamente o debate que a aula quer.

Reforce sempre: o problema nunca é conversar com IA; é o papel que ela vai ocupando.