

▶ Mutant



O QUE É ALUCINAÇÃO NA IA?

ARRASTE PARA O LADO



ALUCINAÇÃO

É quando a IA gera uma resposta factualmente incorreta, falsa ou sem sentido, mas a apresenta com total confiança, como se fosse verdade.

O PARADOXO DA FLUIDEZ

Modelos de linguagem são treinados para prever a próxima palavra mais provável.

Quando falta conhecimento, eles preenchem lacunas com frases que soam certas, mas não são verdadeiras.

**FORMA PERFEITA
≠ CONTEÚDO CERTO.**

POR QUE ISSO ACONTECE?



DADOS INSUFICIENTES OU ENVIESADOS:

o modelo aprende pouco, aprende errado ou aprende algo desatualizado.



“CHUTE ESTATÍSTICO”:

quando falta informação, a IA preenche lacunas com padrões que já viu, inventa a próxima palavra com base em probabilidade

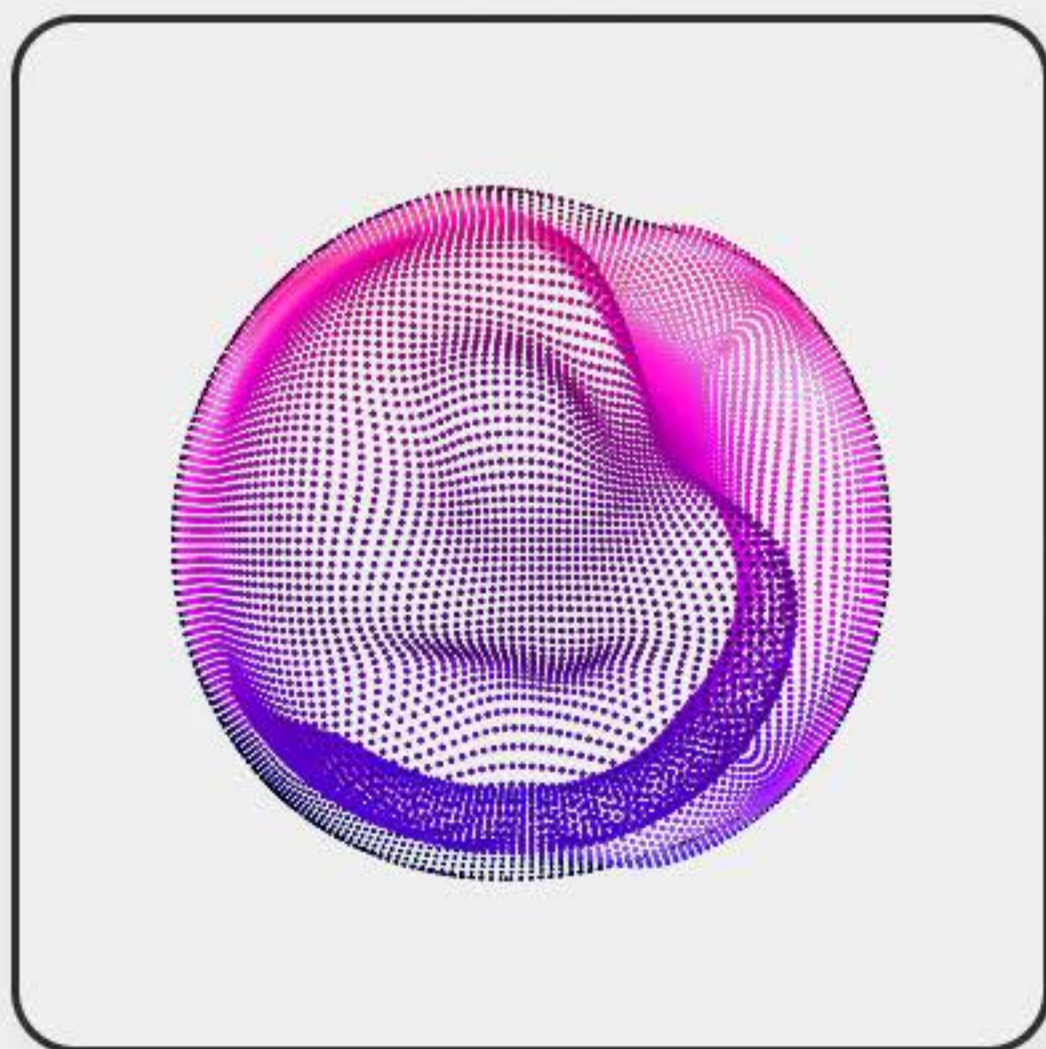
IMPACTO NO MUNDO REAL

A alucinação pode levar a decisões ruins, desinformação e problemas éticos.

Exemplos clássicos: leis inventadas, artigos falsos, fatos históricos distorcidos.

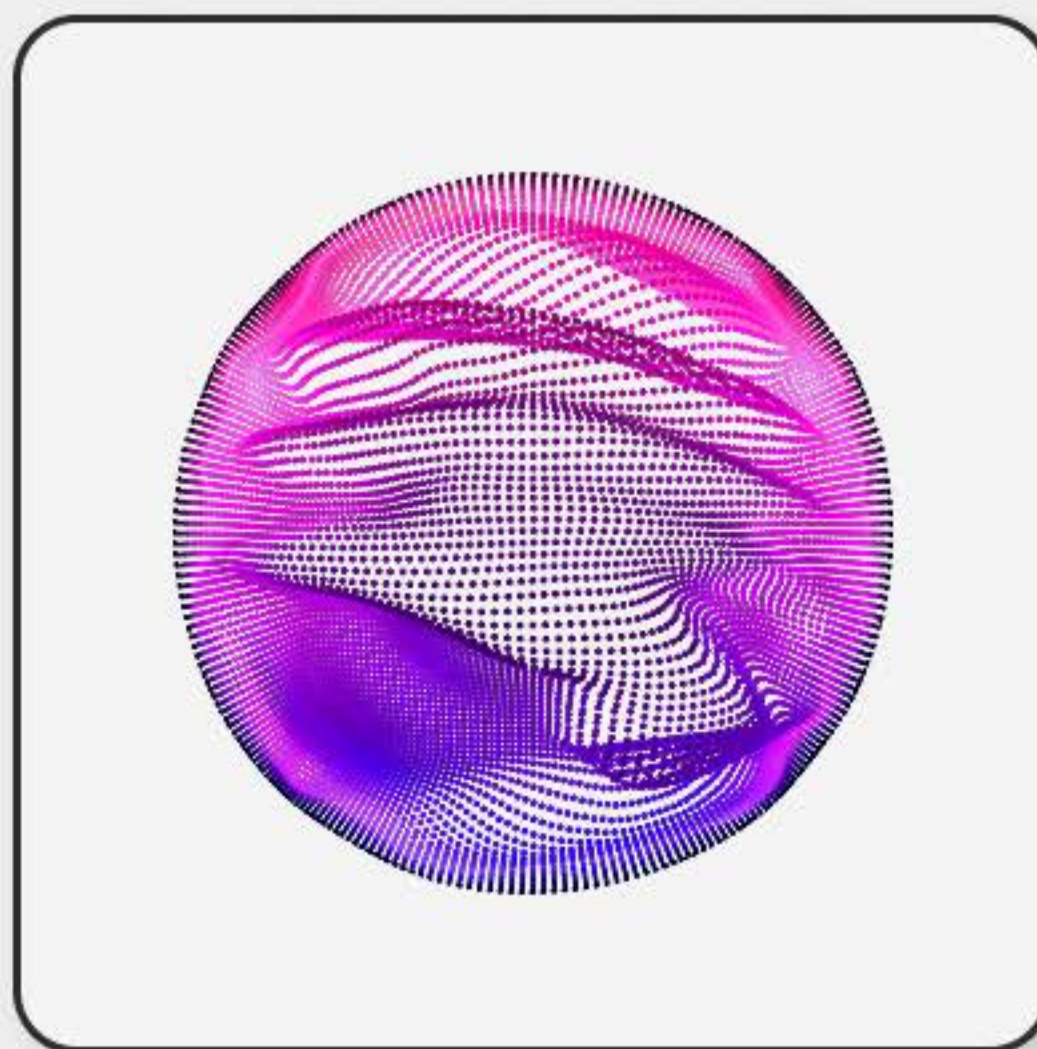
Precisão importa, e muito.

COMO MITIGAR?



RAG (Retrieval-Augmented Generation):

o modelo consulta fontes reais antes de responder.



Fine-Tuning + RLHF:

treino supervisionado que corrige comportamentos e penaliza erros.

DICAS PRA NÃO CAIR EM PEGADINHAS

1. Peça fontes.
2. Divida o problema em etapas (Chain-of-Thought).
3. Pergunte: “Tem certeza?” e peça revisão.