Estudo de Cambridge conclui que a IA ainda não é fiável para avaliar ensaios universitários

AI in Education · Boa prática · Reino Unido

Pontuação de qualidade: 40/100

Força da evidência: Moderada

Sumário

Investigadores da Universidade de Cambridge testaram Claude e ChatGPT na avaliação de 750 ensaios reais de três universidades britânicas. Os sistemas só coincidiram com as classificações dos examinadores humanos em 35–65% dos casos, valorizando excessivamente a extensão e o vocabulário em detrimento do conteúdo.

Dimensões-chave

Dimensão	Pontuação
Evidence of learning impact	1/3
Equity & inclusion	1/3
Transparency, ethics & data protection	2/3
Teacher capability & pedagogy	1/3
Scalability & sustainability	1/3

Avaliado por C-NAPSE

Fontes de dados

C-NAPSE web research — acedido a 2026-09-01

University of Cambridge — acedido a 2026-09-01

University of Cambridge — acedido a 2026-09-01

EurekAlert — acedido a 2026-09-01

Phys.org — acedido a 2026-09-01

Citar — Evidoria. *Estudo de Cambridge conclui que a IA ainda não é fiável para avaliar ensaios universitários*. ID persistente: d938f37e-9a5c-47de-a2f4-cadb5bdbcbf3.