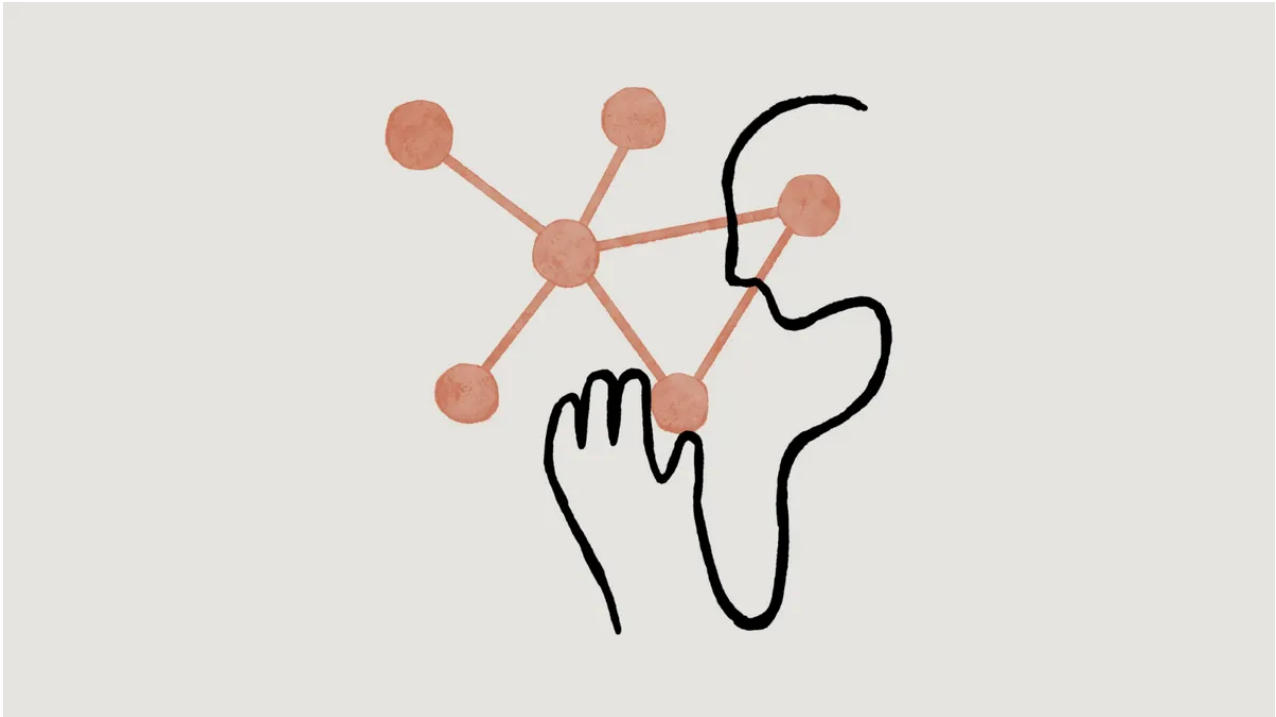


Claude, nový chatbot s umělou inteligencí je vytvořen bývalými zaměstnanci OpenAI

interestingengineering.com/innovation/anthropic-launches-ai-chatbot

15. března 2023



Společnost Anthropic, založená bývalými zaměstnanci OpenAI Danielou a Dariem Amodeiovými, spustila vlastního chatbota s umělou inteligencí (AI), přezdívaného Claude. Anthropic testoval chatbota s dalšími technologickými společnostmi, jako jsou DuckDuckGo, Quora a Notion, a je nyní otevřený spolupráci s dalšími klienty, kteří chtějí do svých produktů začlenit „užitečného“ chatbota.

Viz také

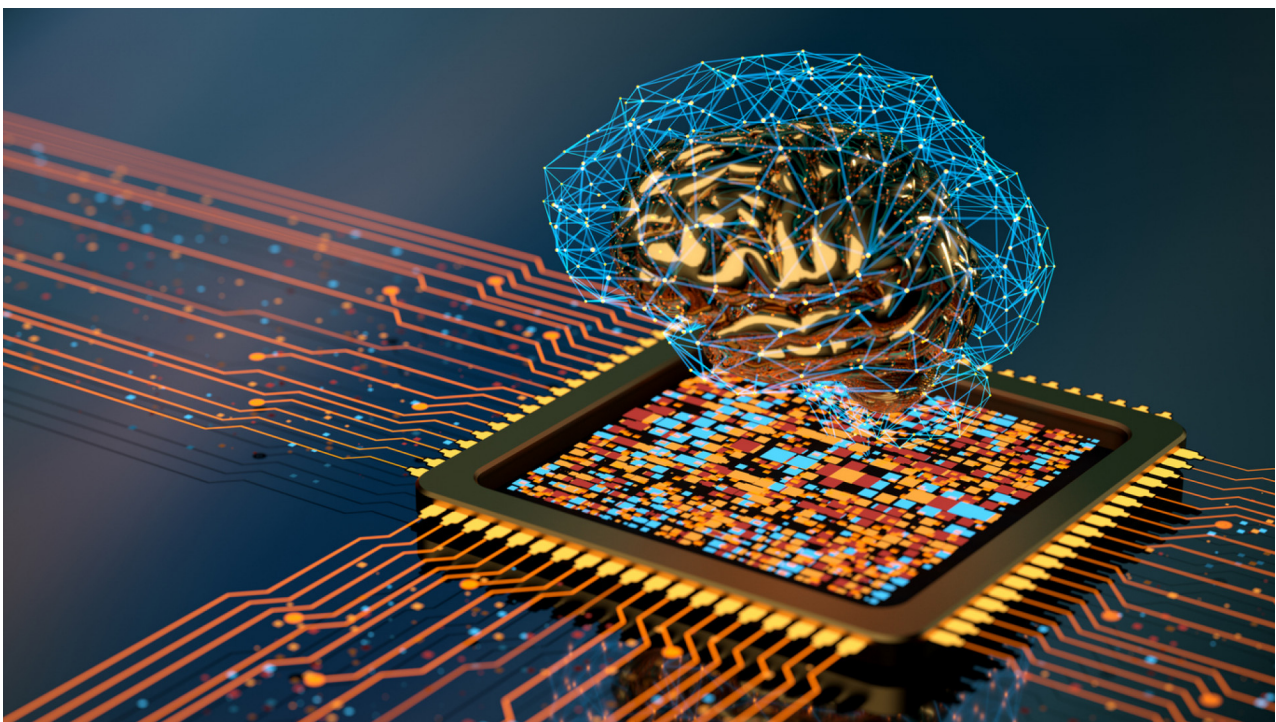
Po úspěchu OpenAI s ChatGPT se může zdát, že jakýkoli jiný chatbot je nyní ztracený případ. Pokus Googlu předvést svou umělou inteligenci Bard selhal poté, co chatbot během veřejného dema odpověděl nesprávně. Zatímco Microsoft byl agresivní při prosazování AI poháněné GPT ve svých vlastních produktech, pečlivě oznámení Anthropic o příchodu svého chatbota vyznačuje důvěru v produkt.

Jak Anthropic postavil Claude

Společnost Anthropic byla založena v roce 2021 lidmi, kteří opustili OpenAI, včetně hlavního inženýra pro GPT-3, Toma Browna.

Existuje tedy mnoho podobností v technologii, která byla použita k jejímu výcviku. Zakladatelé však tvrdí, že použili seznam pravidel a principů k trénování modelu AI, který označují jako Ústavní AI.

Zajímavé inženýrství již dříve informovalo, jak se uživatelům podařilo proniknout do původních pravidel chatbota Bing společnosti Microsoft a přiměl je je odhalit. Takové chování bylo něco, co dokonce způsobilo, že chatbot byl později trochu defenzivní a podrážděný.



Stroj se učí s jakýmkoli zásahem člověka

MF3D/iStock

Claude, jak tvrdí Anthropic, byl vycvičen, aby byl nápomocný a zdvořile odmítal odpovídat na otázky, které by mohly být škodlivé. Podle webové stránky na webových stránkách společnosti toho dosáhla dalším školením chatbota během fáze posílení učení (RL) s jiným preferenčním modelem, aby byl chatbot neškodný, ale také neúhybný.

Nejoblíbenější

Aby Anthropic otestoval Claudovu efektivitu, nabídl společností přístup k chatbotovi. Mezi nimi DuckDuckGo využívá technologii ke spuštění vyhledávacího nástroje pro své uživatele. Google také shledal tuto technologii docela užitečnou a koncem loňského roku investoval do Anthropic 300 milionů dolarů.

Nyní Anthropic uvedl Claude ve dvou variantách, Standard a Instant, přičemž druhá varianta nabízí rychlejší, lehčí a levnější způsob integrace AI do produktů. Anthropic navrhuje, že by chatbot mohl být na začátku pověřen hledáním, zákaznickým servisem, prodejem a také back-office prací.

Claude je navržen tak, aby byl samostatný a nemá přístup k internetu. Pokud jde o vylepšení, výzvy jsou nejlepším způsobem, jak Clauda naučit, co se od úkolu vyžaduje. Kromě toho mu lze také přiřadit osobnost a tón, což dokonce Microsoft pracuje na svém chatbotu .

Bude zajímavé vidět, jak do boje vstoupí nový chatbot a jak se s tím Microsoft hodlá vypořádat.

1. Domov
2. Inovace

 ZOBRAZIT KOMENTÁŘ (0) 