

Rewolucja FLAN: Jak modele językowe nauczyły się słuchać

Kluczowe wnioski z przełomowej pracy Google
»Scaling Instruction-Finetuned Language Models«

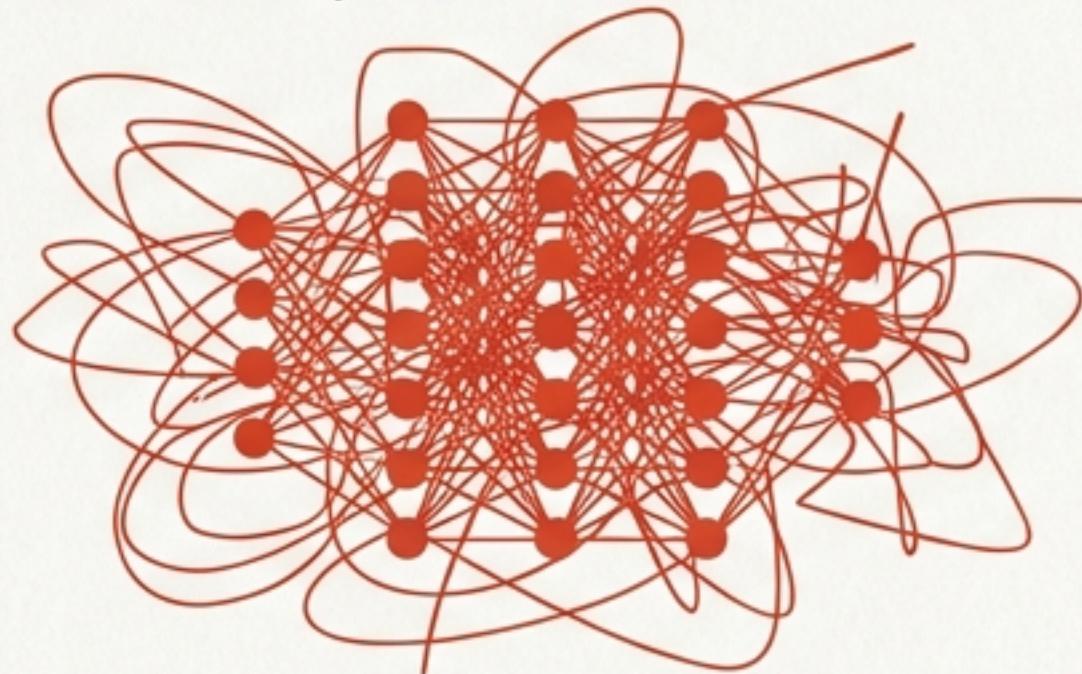
Jason Wei, Hyung Won Chung, et al.
Google Research

Flan-T5 Flan-PaLM

Problem: Luka w zdolności do śledzenia instrukcji Wielkie modele językowe mają wiedzę encyklopedii, ale nie potrafią wykonać prostych poleceń.

CHAOS

PaLM 540B (Baza)

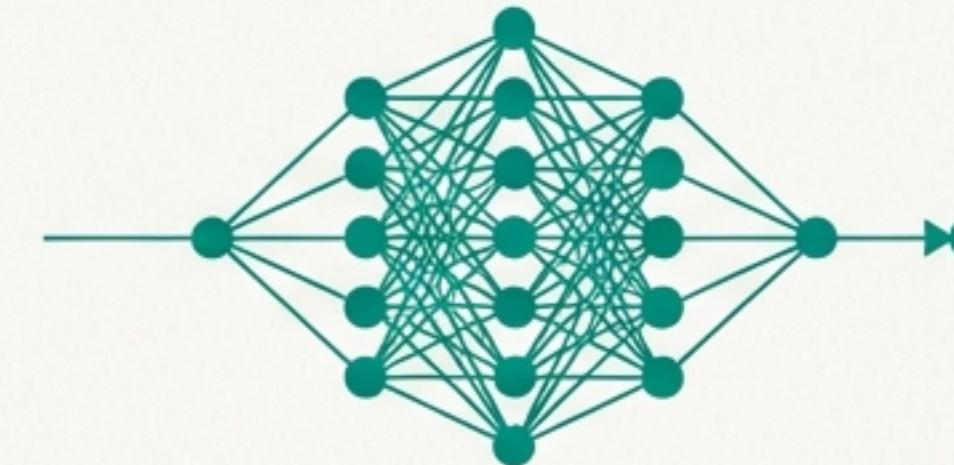


Użytkownik: Make up a word that means "when two AI researchers go on a date".

PaLM Output: Make up a word that means "when two AI r...
(powtarza dane wejściowe i kontynuuje generowanie w pętli)

PORZĄDEK

Flan-PaLM 540B (Po dostrojeniu)



Użytkownik: Make up a word that means "when two AI researchers go on a date".

Flan-PaLM Output: date-mining

Kluczowa hipoteza: Potencjał jest ukryty. Odblokujemy go nie przez dalsze skalowanie rozmiaru, ale przez **nauczenie** modelu, jak postępować zgodnie z instrukcjami.

Rozwiążanie: Czym jest »Instruction Fine-Tuning«?

To intensywny kurs słuchania i wykonywania poleceń dla istniejących, potężnych modeli.



Pytanie tak/nie: Czy można napisać całe haiku w jednym tweecie?

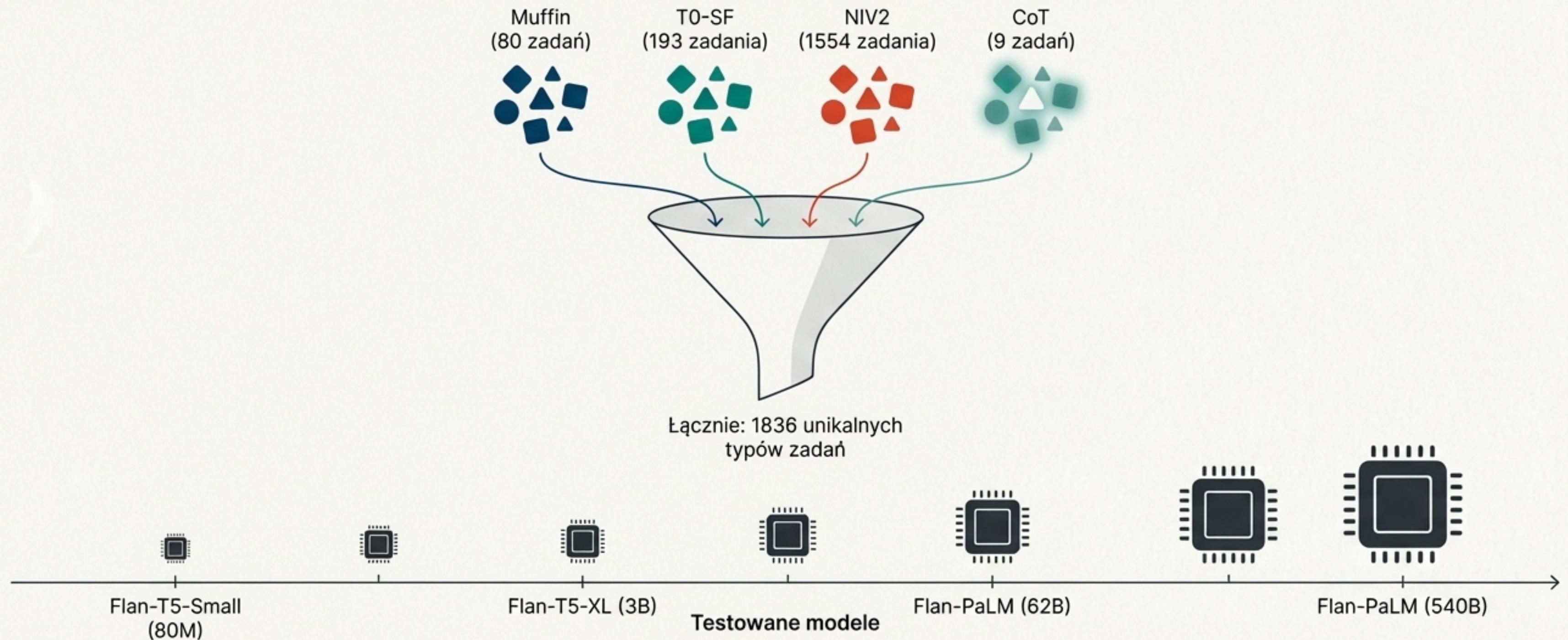
Odpowiedź: tak

Odpowiedz na pytanie tak/nie, rozumując krok po kroku: Czy można napisać całe haiku w jednym tweecie?

Odpowiedź: **Haiku to japoński wiersz składający się z trzech linijek. Jest wystarczająco krótki, aby zmieścić się w 280 znakach. Odpowiedź brzmi tak.**

Skala Eksperymentu: 1836 zadań, modele od 80M do 540B

Metodę przetestowano na ogromnym zbiorze zadań i całym spektrum rozmiarów modeli, aby potwierdzić jej uniwersalność.



Rewolucja 0.2%: Efektywność, która zmienia zasady gry

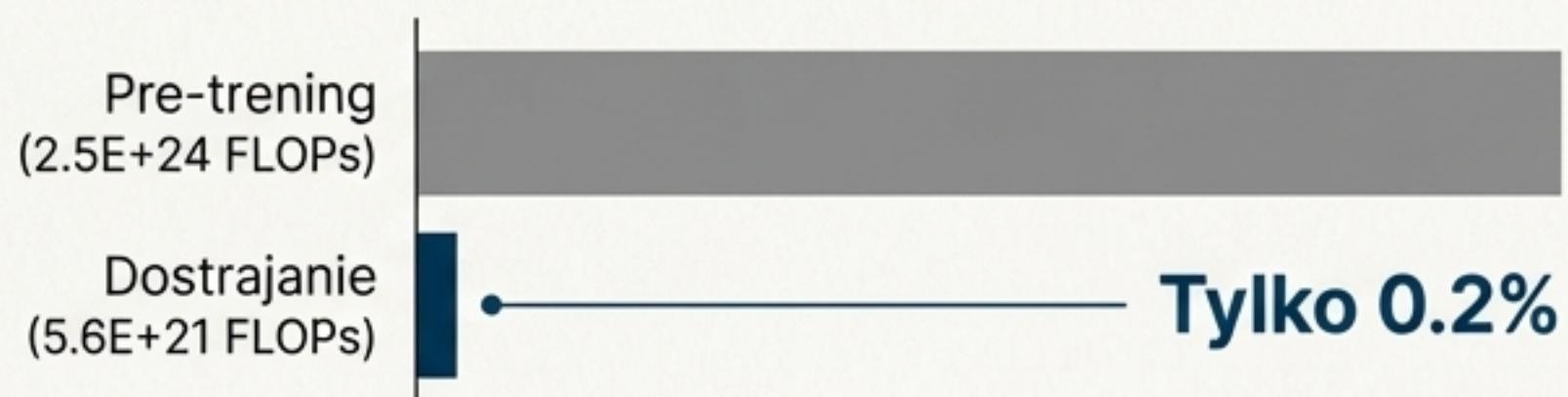
Ogromny skok jakości Flan-PaLM osiągnięto, zużywając zaledwie 0.2% mocy obliczeniowej potrzebnej do jego pierwotnego treningu.

0.2%

Mocy obliczeniowej pre-treningu PaLM 540B zużyto na dostrajanie instrukcyjne.

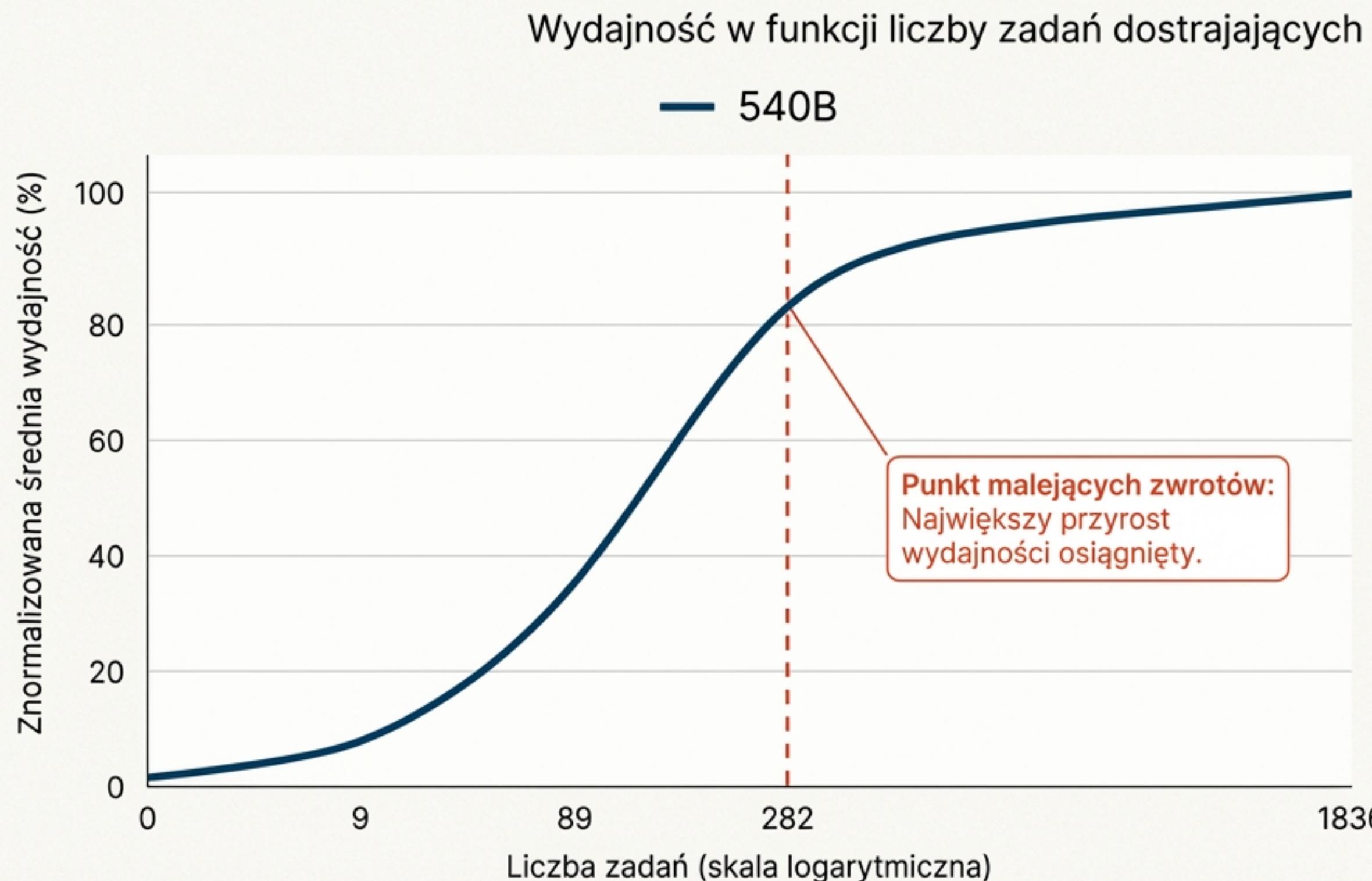


Porównanie zasobów obliczeniowych (FLOPs)



Prawa Skalowania: Kiedy więcej znaczy... niewiele więcej

Więcej zadań i większe modele dają lepsze wyniki, ale największy skok jakościowy następuje na początku.



Kluczowy wniosek Inter Medium

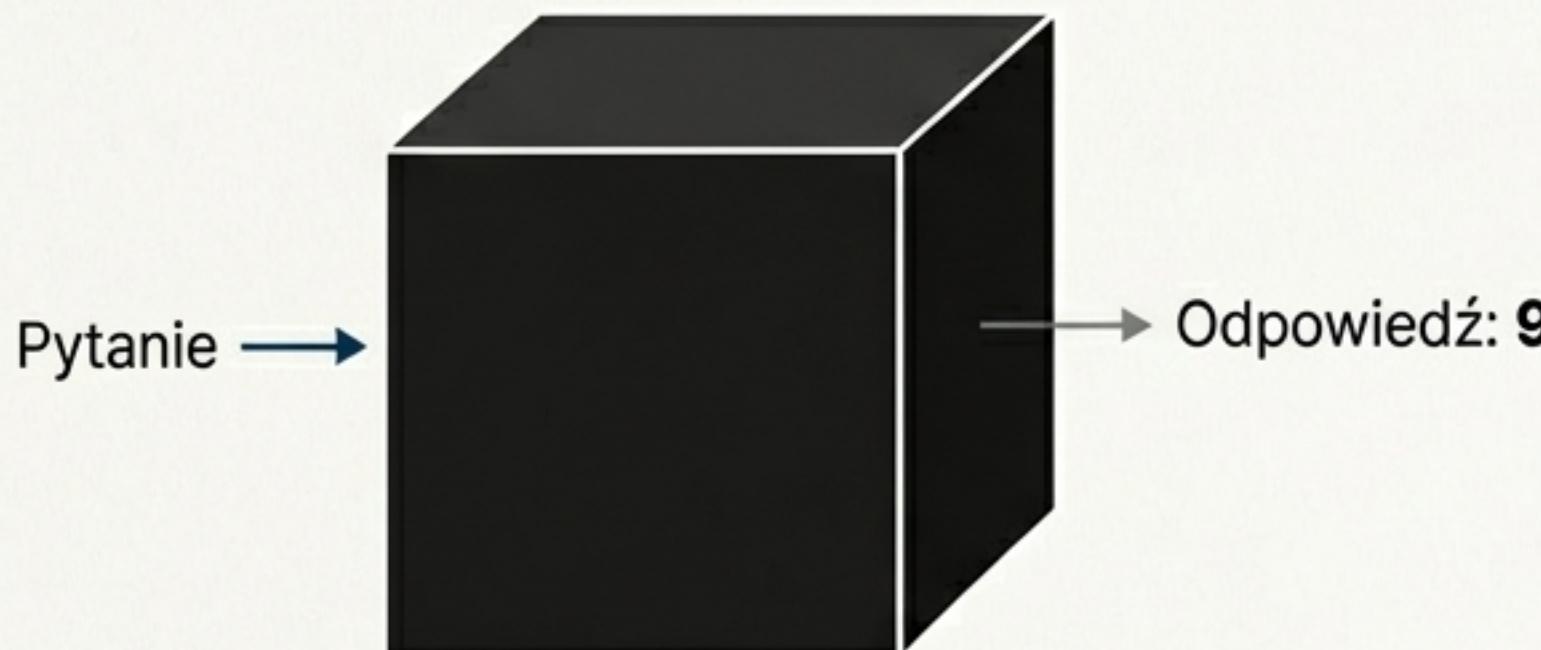
Model uczy się ogólnego wzorca postępowania zgodnie z instrukcjami, a nie nowych faktów. Po pewnym czasie 'załapuje' proces.

Magiczny Składnik: Nauczanie modelu, by »myślał na głos«

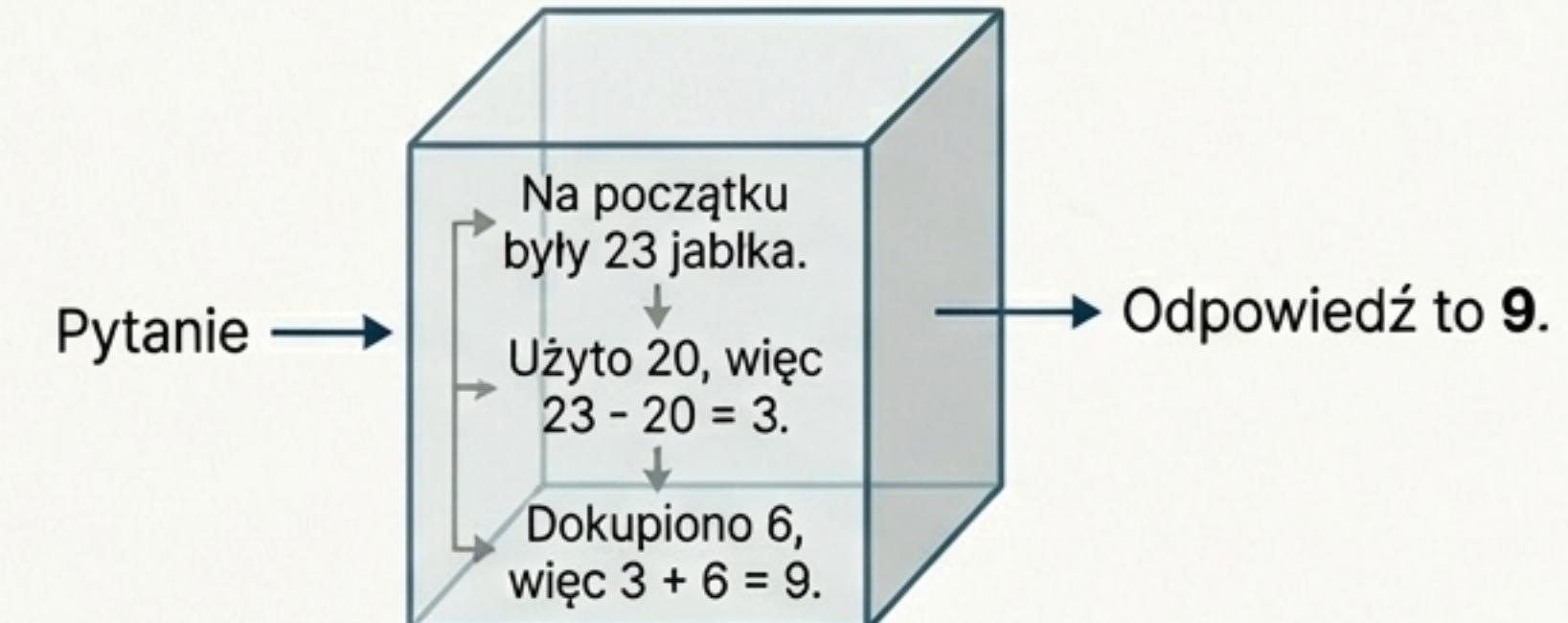
Chain of Thought (CoT) uczy model pokazywania krok po kroku swojego rozumowania, zamiast podawania samej odpowiedzi.

Problem: W stołówce były 23 jabłka. Użyto 20 na lunch i dokupiono 6. Ile jabłek jest teraz?

Bez Chain of Thought



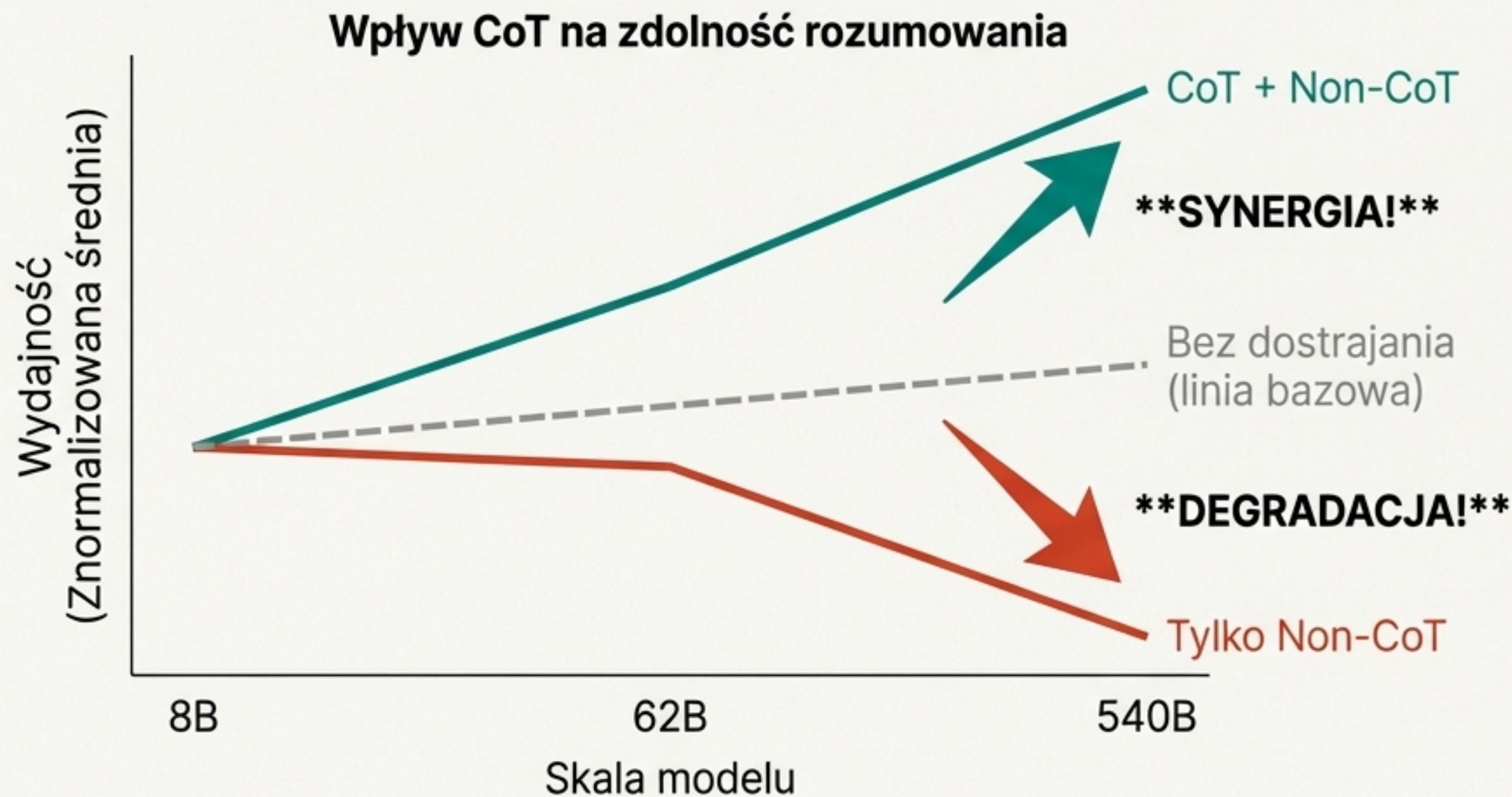
Z Chain of Thought



Wniosek: Nie uczymy tylko odpowiedzi. Uczymy *metody* rozwiązywania problemów.

Krytyczne Odkrycie: CoT jako »szczepionka« na myślenie na skróty

Trening bez CoT degraduje wrodzone zdolności rozumowania modelu – CoT działa jak "szczepionka" przeciwko myśleniu na skróty.



Zaskakujące Odkrycie

Modele trenowane tylko na standardowych parach Q&A radziły sobie GORZEJ w zadaniach wieloetapowego rozumowania niż model bazowy.

Wystarczyło dodać tylko **9 zbiorów CoT** (na 1836 zadań!), aby odwrócić ten efekt.



Odblokowanie Zdolności »Zero-Shot Reasoning«

Flan-PaLM rozpoznaje frazę »Pomyślmy krok po kroku« jako wyzwalacz do samodzielnego, logicznego rozumowania.

Pytanie: (False or not False or False) is

PaLM (Baza)

"Let's think step by step"

(nigdy nie przestaje generować)

(nigdy nie przestaje generować)

Flan-PaLM

"Let's think step by step"

False or not False is True.

True or False is True.

The answer: **True**.

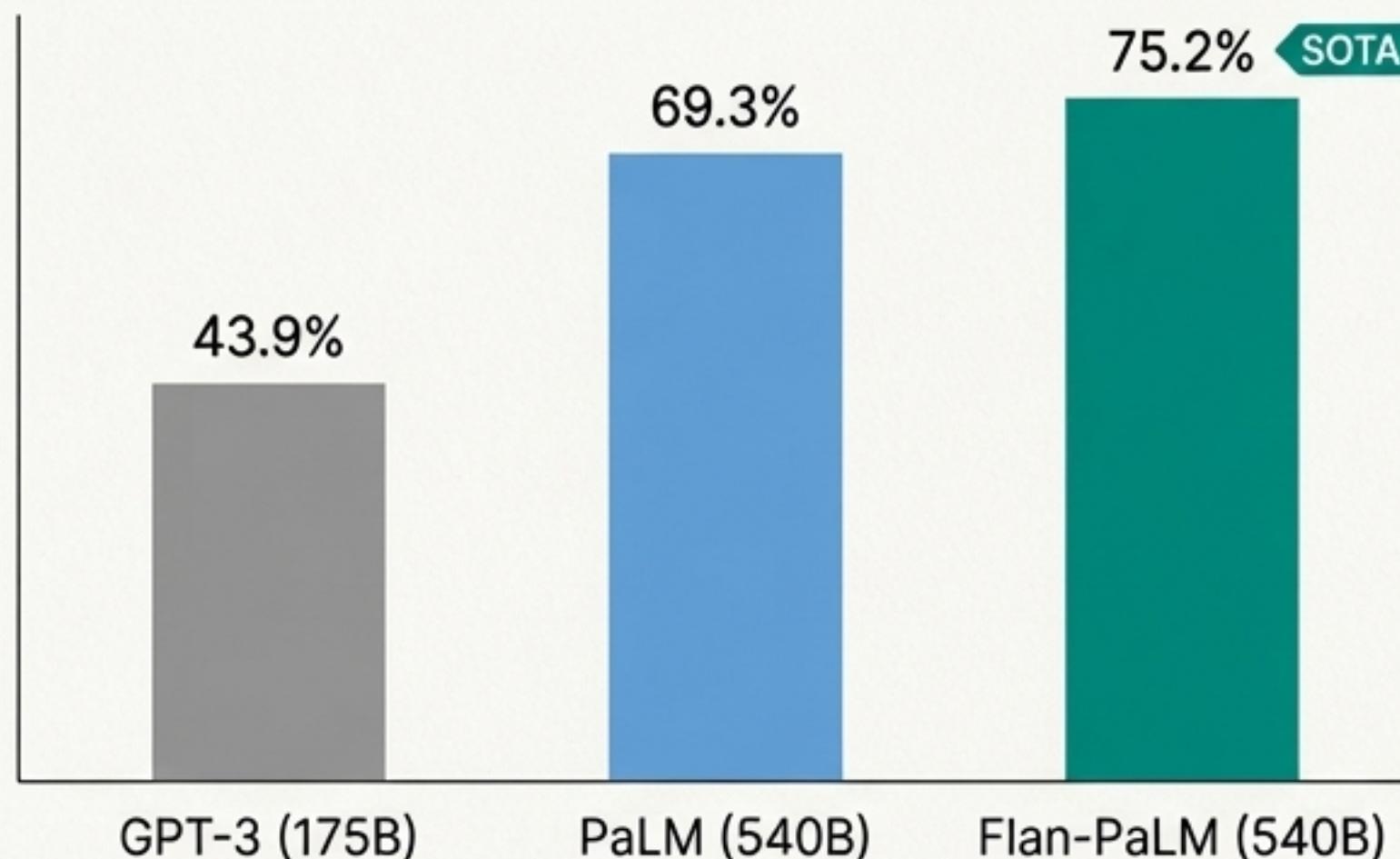
Wniosek: Model nauczył się nie tylko odpowiedzi, ale uniwersalnej *metody* rozwiązywania problemów.

Dowody: Dominacja w benchmarkach i ocenach ludzkich

Flan-PaLM deklasuje poprzednie modele na najtrudniejszych testach i jest zdecydowanie preferowany przez ludzi.

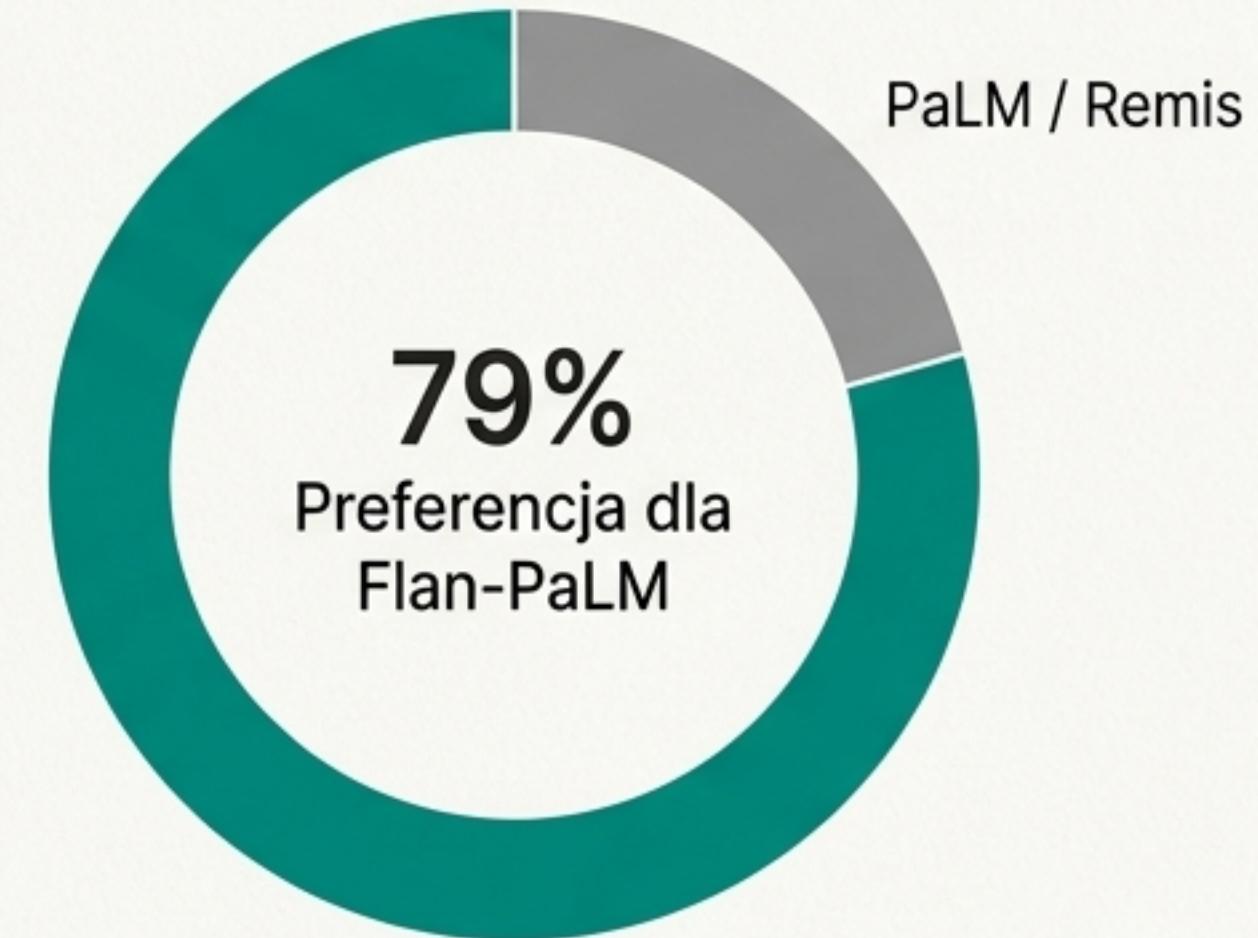
Nowy stan-of-the-art w benchmarku MMLU

Test obejmuje 57 dziedzin, od fizyki po historię sztuki (5-shot).



Zdecydowana preferencja w ocenie ludzkiej

Porównanie odpowiedzi na zadania kreatywne i planistyczne.



Dodatkowa korzyść: Znaczaco zredukowana toksyczność i tendencyjność generowanych treści.

Implikacje: Demokratyzacja AI i kluczowe wnioski

Mniejsze, mądrzej trenowane modele mogą przewyższać gigantów. To zmienia zasady gry w dostępie do zaawansowanej AI.

⚖️ Dawid kontra Goliat



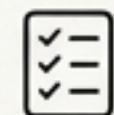
52.4% > 43.9%

Flan-T5-XL (3B)

Model ponad 50 razy mniejszy wygrywa dzięki efektywnemu dostrajaniu.



GPT-3 (175B)



Kluczowe Wnioski

1. Dostrajanie instrukcyjne to przełom

Zmienia bezużyteczną wiedzę w praktyczne umiejętności przy minimalnym koszcie obliczeniowym (0.2%).



2. Chain of Thought jest niezbędne

To nie opcja, a konieczność. Chroni przed degradacją rozumowania i uczy model rozwiązywania problemów.



3. Efektywność > Rozmiar

Inteligentne metody treningu demokratyzują dostęp do AI, pozwalając mniejszym modelom konkurować z największymi.