Architektura Krytycyzmu

Strategie krytycznego myślenia w interakcji z GenAI i LLM-ami

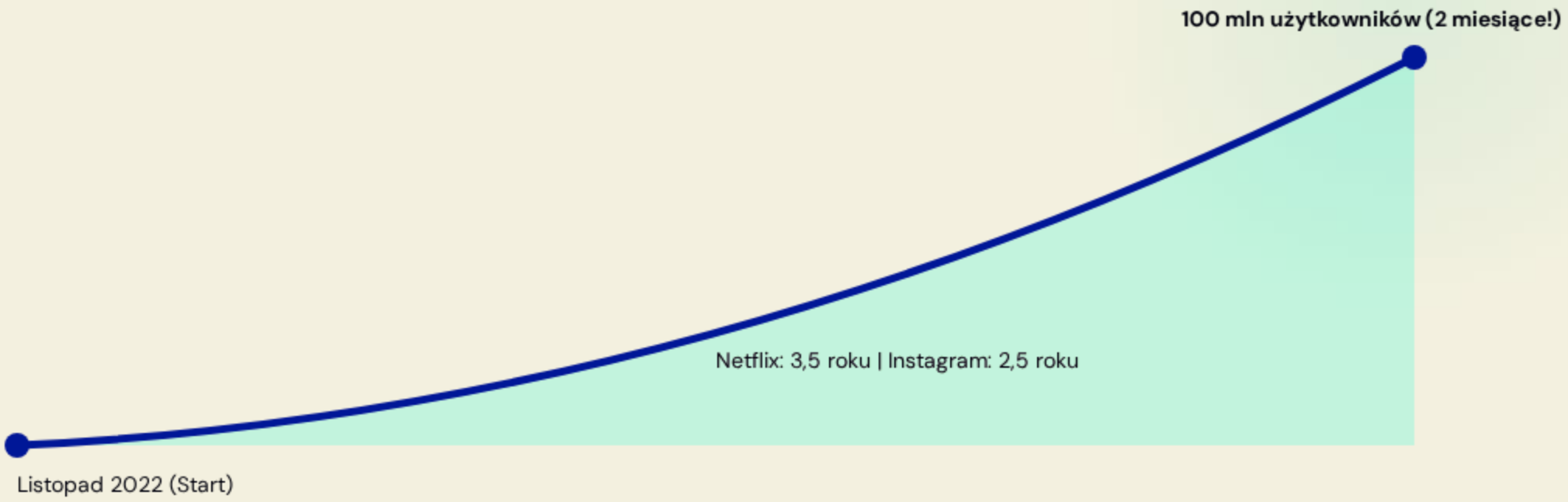
mgr Ewelina Gajewska, mgr Katarzyna Kwiatkowska

Test Turinga w Twojej kieszeni

Czy potrafisz odróżnić?

Przed Tobą dwa obrazy. Jeden jest dziełem człowieka, drugi został wygenerowany w 15 sekund za 2 grosze.

Rewolucja w tempie wykładniczym



Mamy do czynienia z najszybciej adoptowaną technologią w historii ludzkości. **Czy tempo adaptacji idzie w parze ze zrozumieniem?**

Narzędzie, Partner czy... Zagrożenie?

Super-Kalkulator

Narzędzie, które po prostu szybciej wykonuje nudną robotę. Nie "myśli", tylko liczy.

Cyfrowy Co-pilot

Partner do burzy mózgów, który rozszerza moje możliwości twórcze i poznawcze.

Obca Inteligencja

Byt, którego mechanizmów nie rozumiemy i który może przejąć nasze unikalne kompetencje.

W której grupie jesteś dzisiaj?

Zanim zajrzymy "pod maskę"...

Co tak naprawdę wiemy?

Większość z nas korzysta z GenAI codziennie, ale dla wielu pozostaje ono **"Czarną Skrzynką" (Black Box)**.

- Jak to "wie", co napisać?
- Czy to jest inteligentne, czy tylko tak wygląda?
- Czy możemy mu ufać?

Sprawdźmy, jaki jest nasz punkt startowy!

Quiz Time

Możecie się dobrać w pary/grupy i odpowiadać wspólnie lub indywidualnie

Umysł Ludzki vs. Cyfrowy Model

Zanim zrozumiemy, jak działa "sztuczna" inteligencja, musimy zdefiniować tę "naturalną"

Człowiek vs. LLM: Różnice fundamentalne

Cecha	Człowiek	Model LLM
Źródło danych	Zmysły i ciało	Tylko tekst/kod
Rozumienie	Głęboka semantyka	Czysta syntaktyka
Energia	~20W (kanapka)	Megawaty (GPU)
Kontekst	Biokulturowy	Statystyczny

Chiński pokój Johna Searle’a

Syntaktyka ≠ Semantyka

- Osoba w pokoju używa instrukcji do manipulowania symbolami chińskimi.
- Na zewnątrz wygląda to na płynną rozmowę, ale osoba wewnątrz nie rozumie ani słowa.
- **Wniosek:** Wykonywanie algorytmu (LLM) nie jest tożsame z posiadaniem umysłu czy rozumieniem znaczenia.

Czy "Sztuczna Inteligencja" to dobra nazwa?

> *"AI to nie inteligencja, to potężna statystyka przebrana w szaty języka."*

"Stochastyczne Papugi"

Termin ukuty przez Emily M. Bender. Sugeruje, że modele LLM jedynie powtarzają sekwencje słów na podstawie prawdopodobieństwa, nie rozumiejąc ich sensu.

Alternatywy:

- Statystyczne Modelowanie Języka
- Maszyny Predykcyjne
- Generatywne Systemy Probabilistyczne

Jak to działa? Podstawy generowania tekstu

Tokenizacja

Model nie czyta słów. Dzieli tekst na fragmenty (tokeny). "Kognitywistyka" może być dla modelu trzema różnymi liczbami.

Predykcja

Zadanie modelu jest jedno: przewidzieć najbardziej prawdopodobny następny token (symbol) w sekwencji. To gigantyczny autouzupełniacz.

Kontekst

Model bierze pod uwagę poprzednie tokeny (okno kontekstowe), aby zdecydować, co pasuje najlepiej w danej sytuacji.

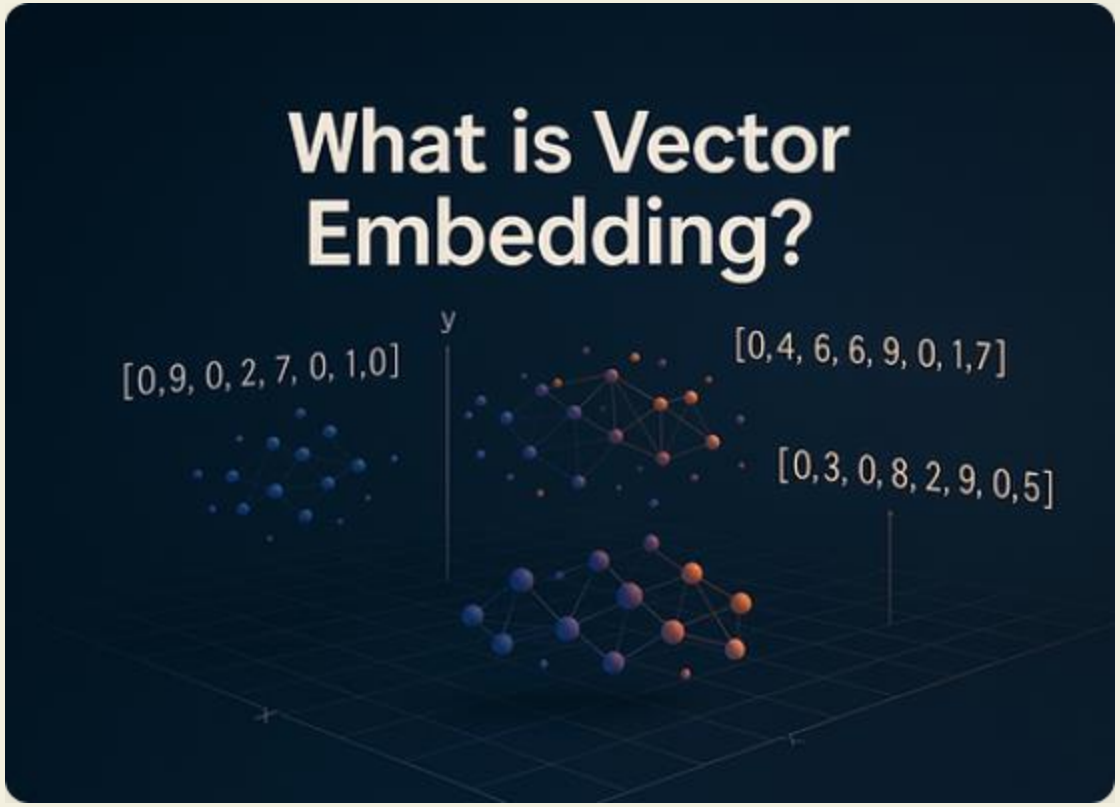
Embeddings: Matematyczna mapa pojęć

Modele LLM mapują słowa na **wielowymiarową przestrzeń** wektorową. Słowa o podobnym znaczeniu znajdują się blisko siebie.

Król – Mężczyzna + Kobieta = Królowa

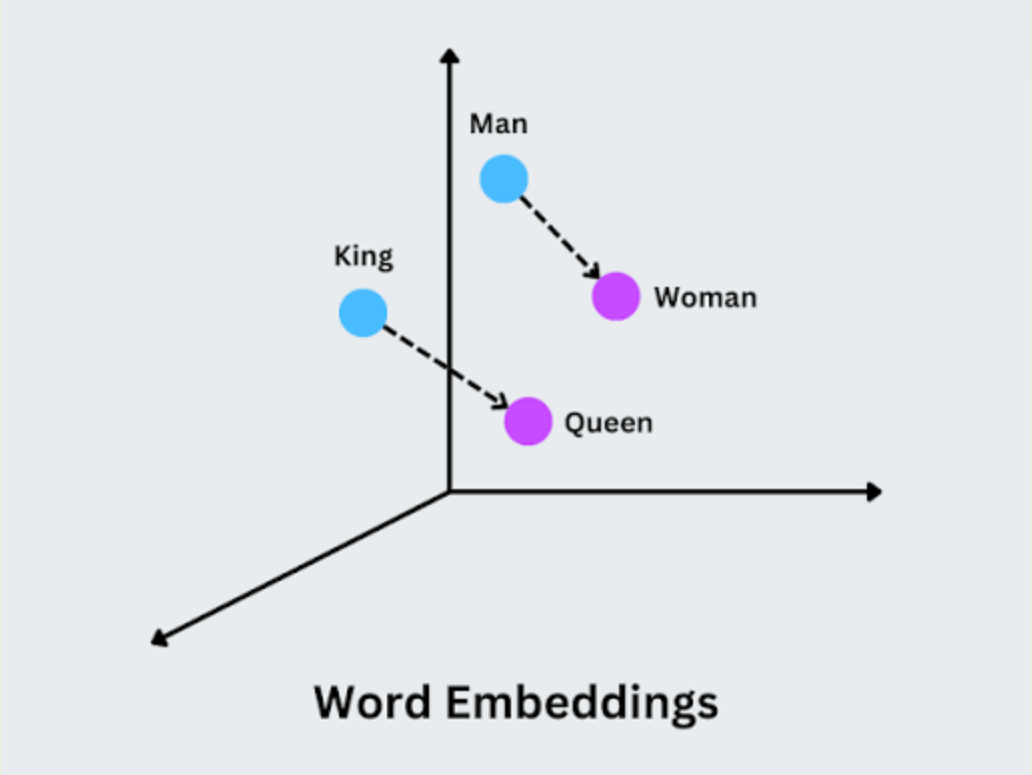
125 – 60 + 50 = 115

Relacje semantyczne są tłumaczone na dystans w przestrzeni wielowymiarowej (o tysiącach wymiarów).

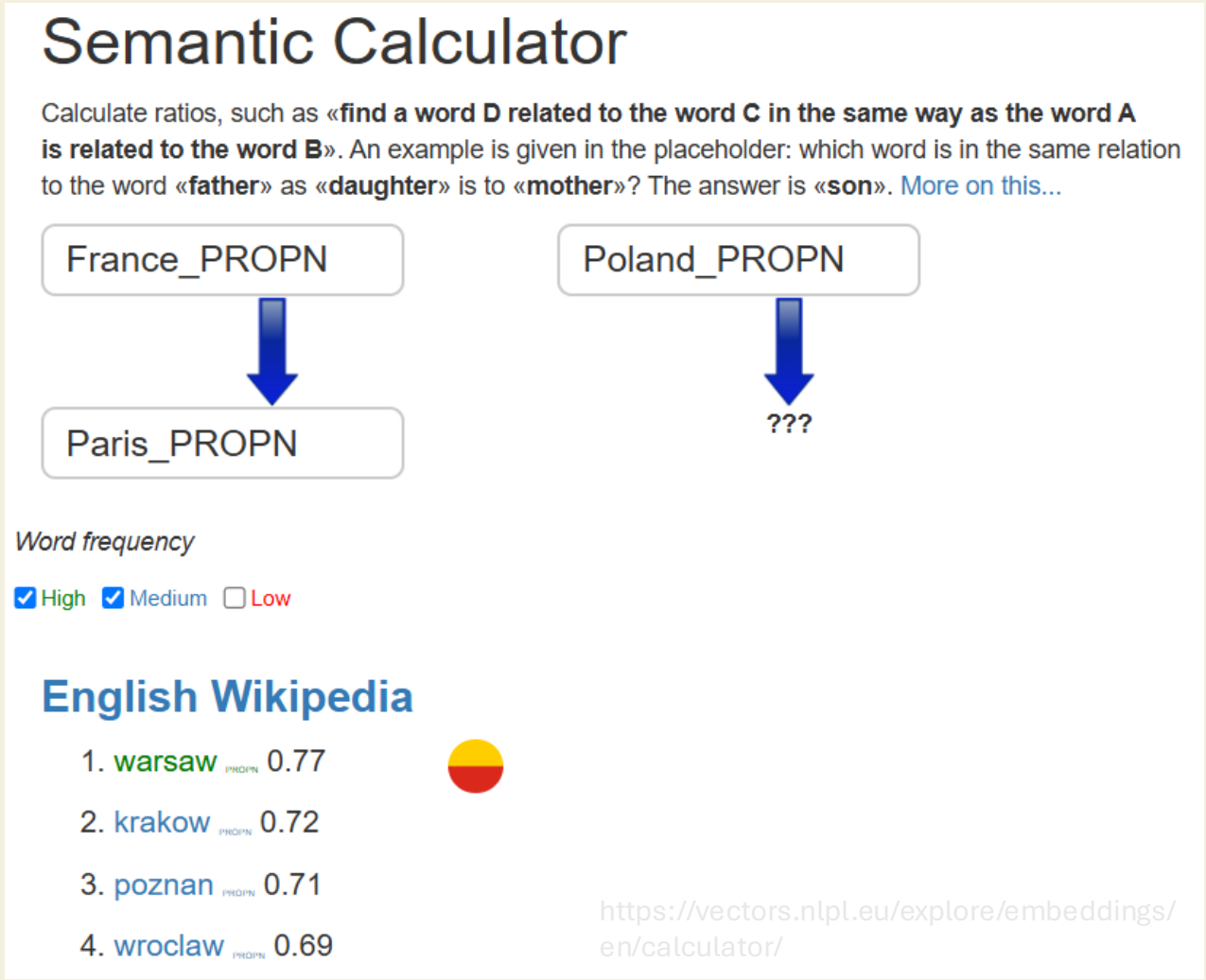


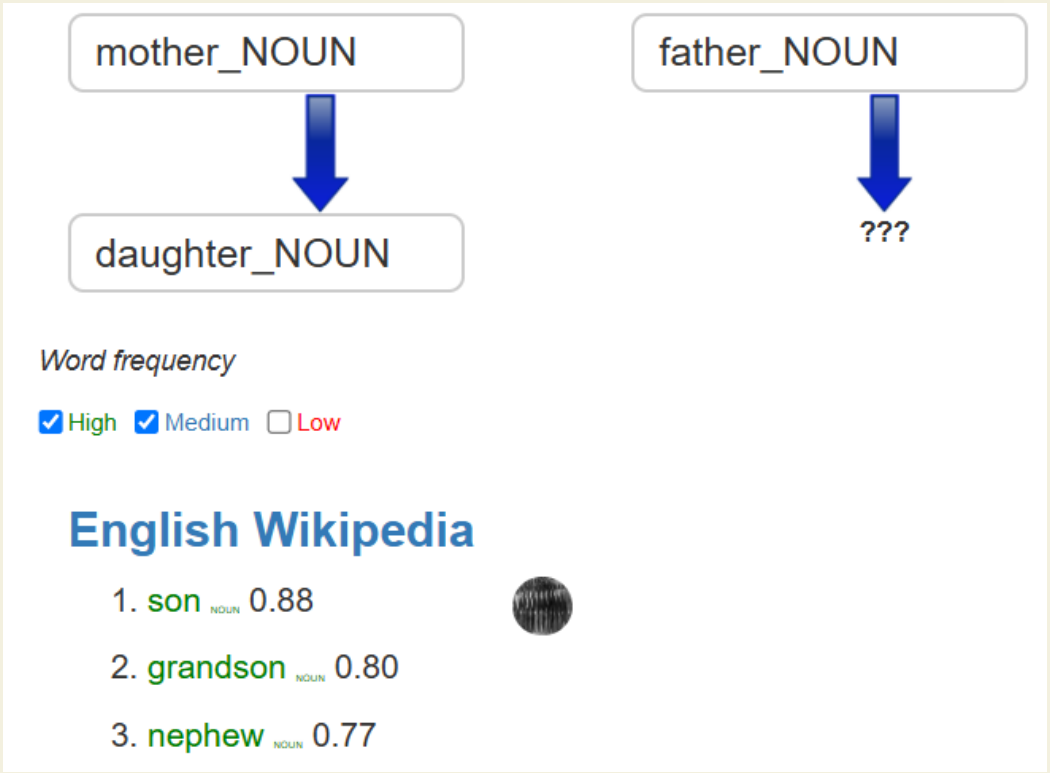
Jak powstają takie mapy?

- Każde słowo (lub token) jest zamieniane na **wektor liczb**, np. 768, 1024 lub 4096 wartości.
- Model uczy się tych reprezentacji podczas trenowania na **miliardach zdań**.
- Słowa pojawiające się w **podobnych kontekstach** otrzymują **podobne embeddingi:**
  - kot → [0.18, -0.42, 1.07, ..., 0.56]
  - pies → [0.21, -0.39, 1.01, ..., 0.61]
  - samochód → [-0.77, 0.93, -0.12, ..., -0.44]
- Embedding przechowuje informacje o **kontekście i relacjach** między pojęciami (słowami).



Kalkulator relacji semantycznych





Ewolucja: Od logiki do prawdopodobieństwa

Lata 60-80.

Systemy eksperckie. "Jeśli-to". Sztywna logika symboliczna.

Lata 90-00.

Uczenie maszynowe. Statystyka na małych zbiorach danych.

2012-2016

Deep Learning. Sieci neuronowe (CNN, RNN). Rozpoznawanie obrazu.

2017-Dziś

Era Transformerów. Generatywna AI. Skalowanie do bilionów danych.

Podsumowanie i Pytania

Zanim przejdziemy do praktyki:

Czy wiesz już, dlaczego model może kłamać (halucynować) z taką samą pewnością siebie, z jaką podaje fakty? To wynika z samej jego natury – statystyki, a nie prawdy.