

Sztuczna inteligencja i rozwój nowych kompetencji

dr hab. inż. Piotr Szczuko

Konferencja (Nie)ludzkie oblicze sztucznej inteligencji.
Szanse i wyzwania związane z rozwojem sztucznej inteligencji.

Olsztyn, 16 października 2023



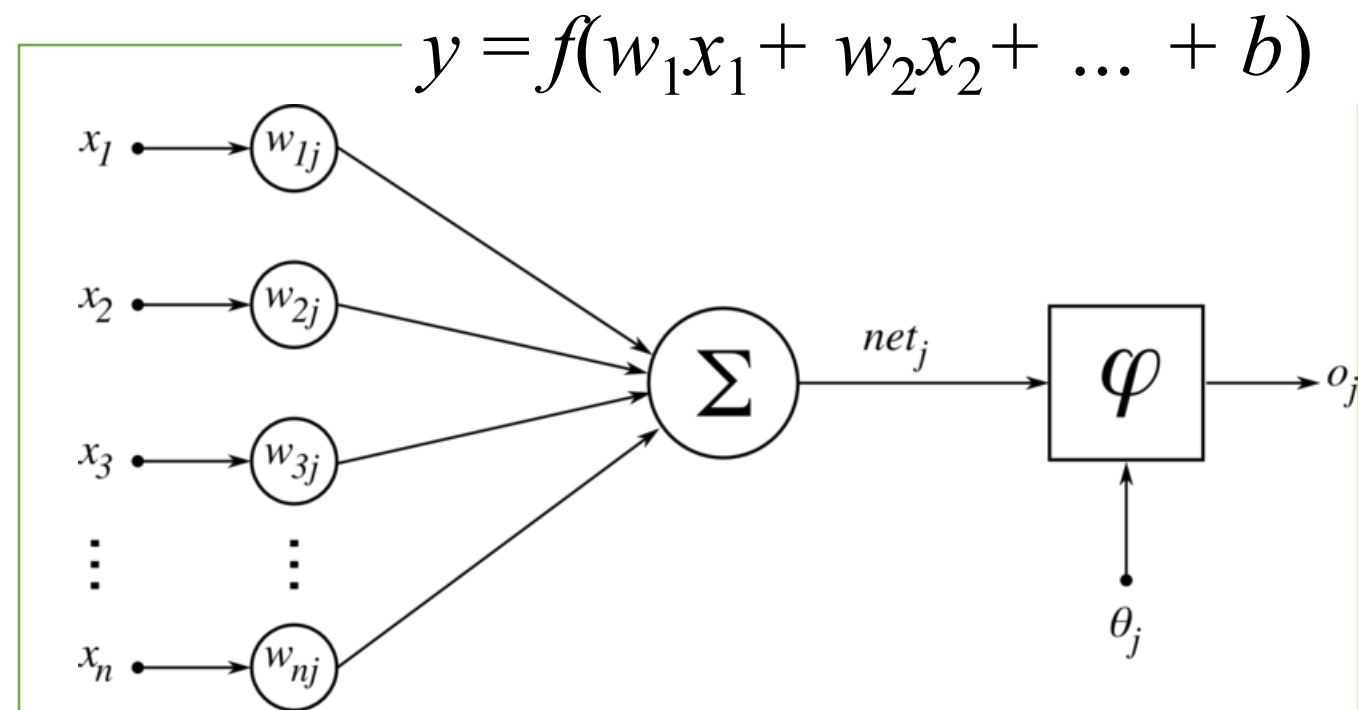
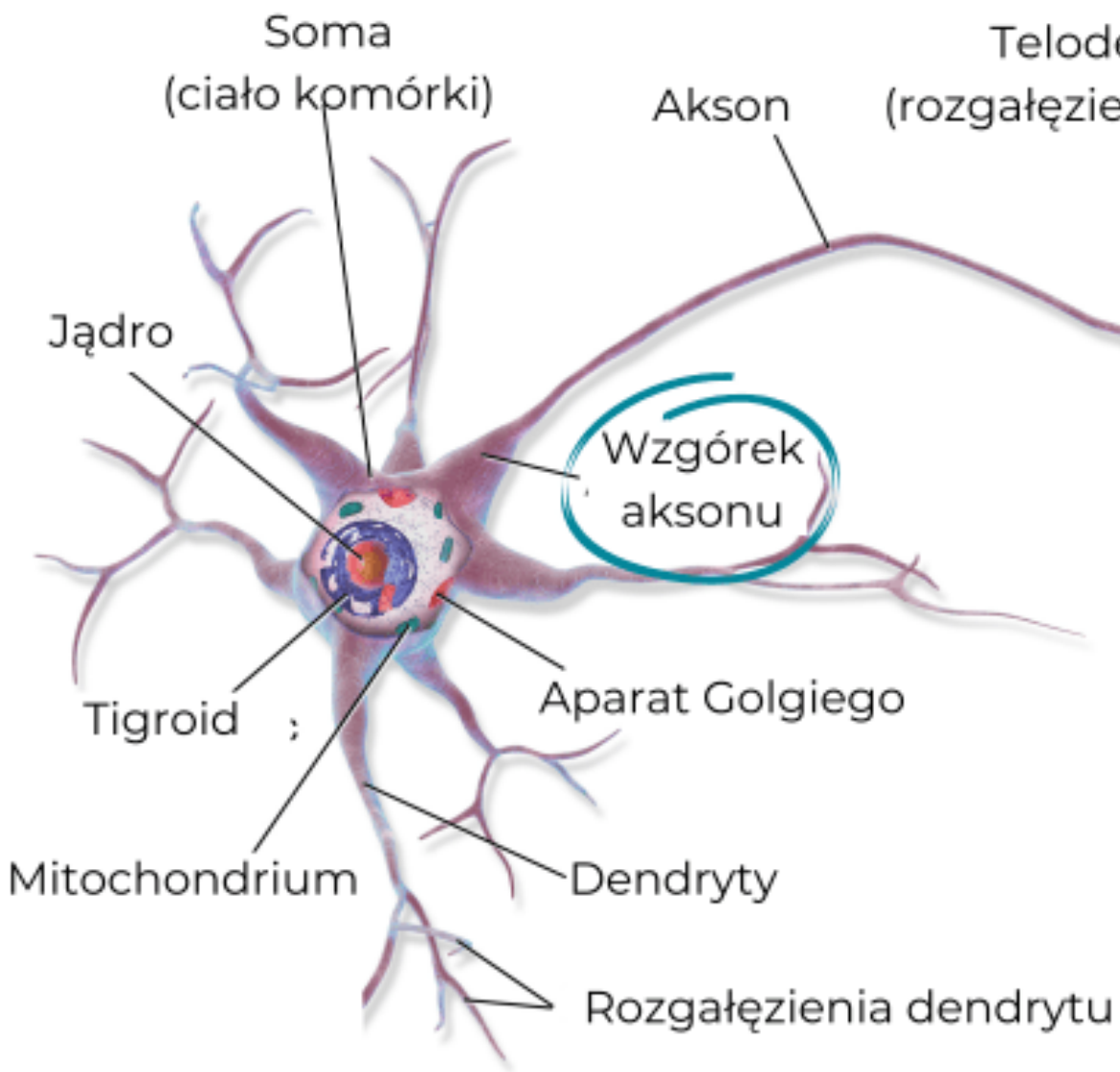
3 tysiące lat rozwoju

od prostych automatów (1000 p.n.e.) do
“inteligentnych” asystentów (2023 n.e.)

wybór subiektywny

1943-1965

- 1943 Warren S. McCulloch, Walter Pitts "A Logical Calculus of the Ideas Immanent in Nervous Activity"
- Działanie biologicznego neuronu opisane prostym modelem matematycznym
- Co potrafi jeden sztuczny neuron?





<https://www.gregadunn.com/self-reflected/>

▶ Ludzki mózg

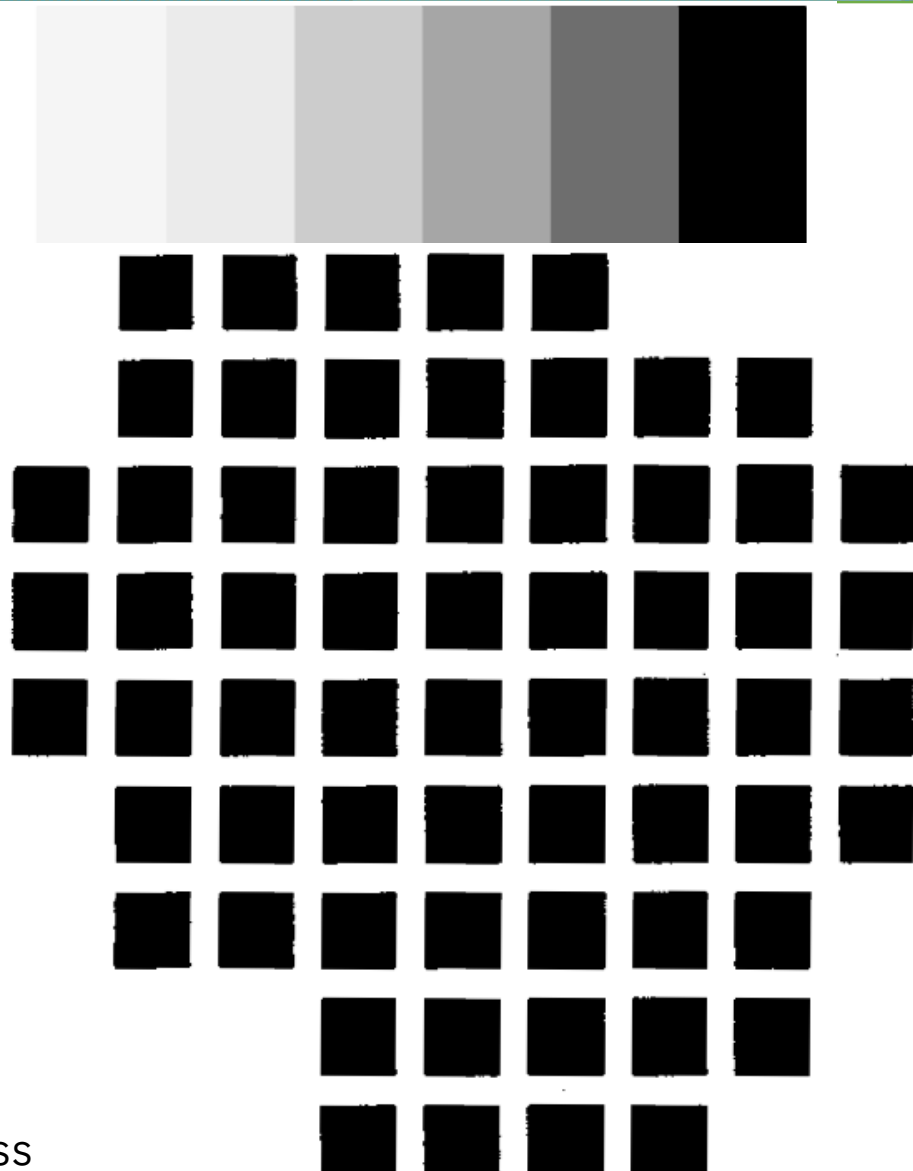
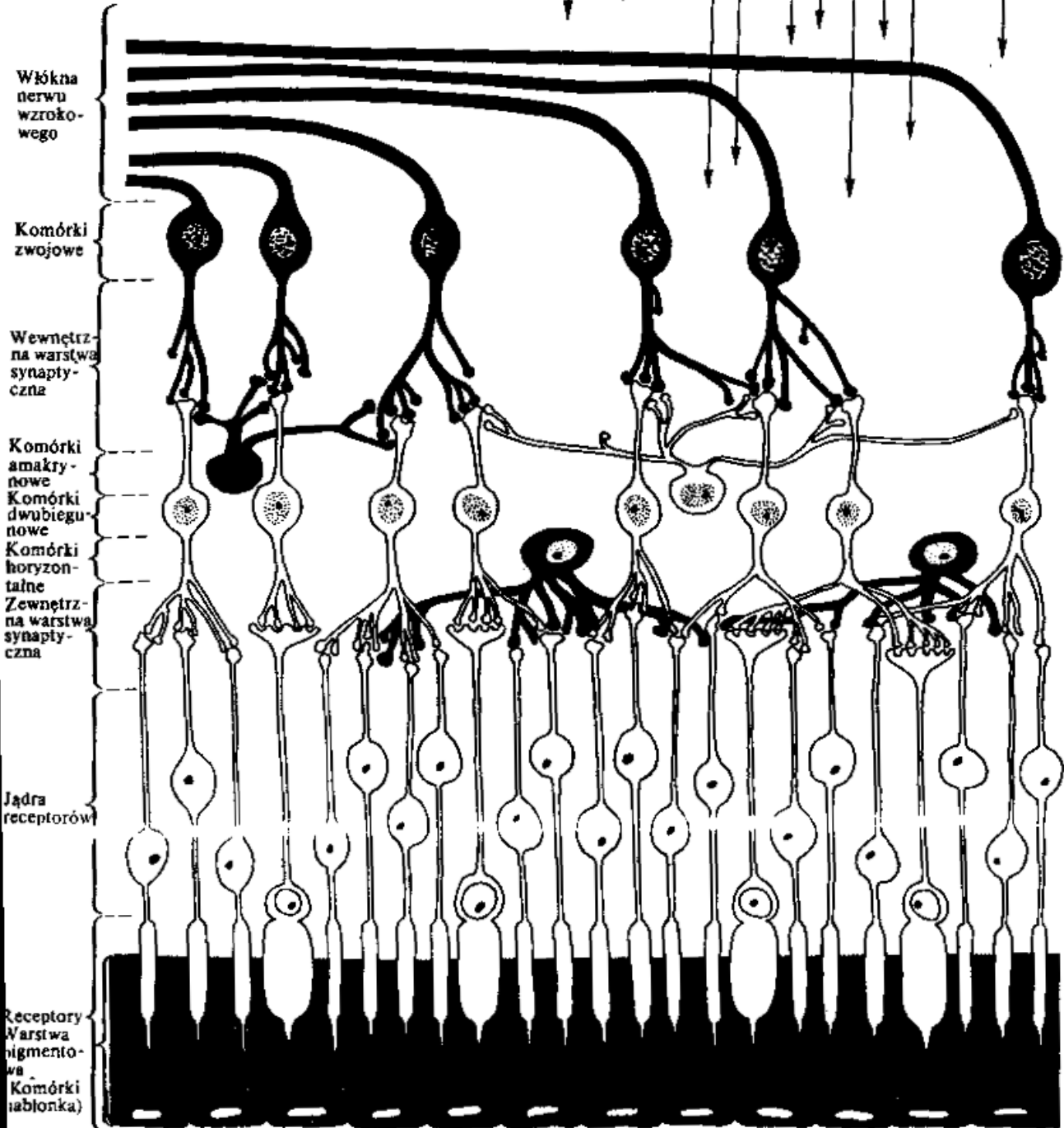
- 86 miliardów neuronów
- 100 bilionów połączeń
- Specjalizowane ośrodki
- Plastyczność
- Do zamodelowania jednego neuronu potrzebne 1000 sztucznych



1950-1989

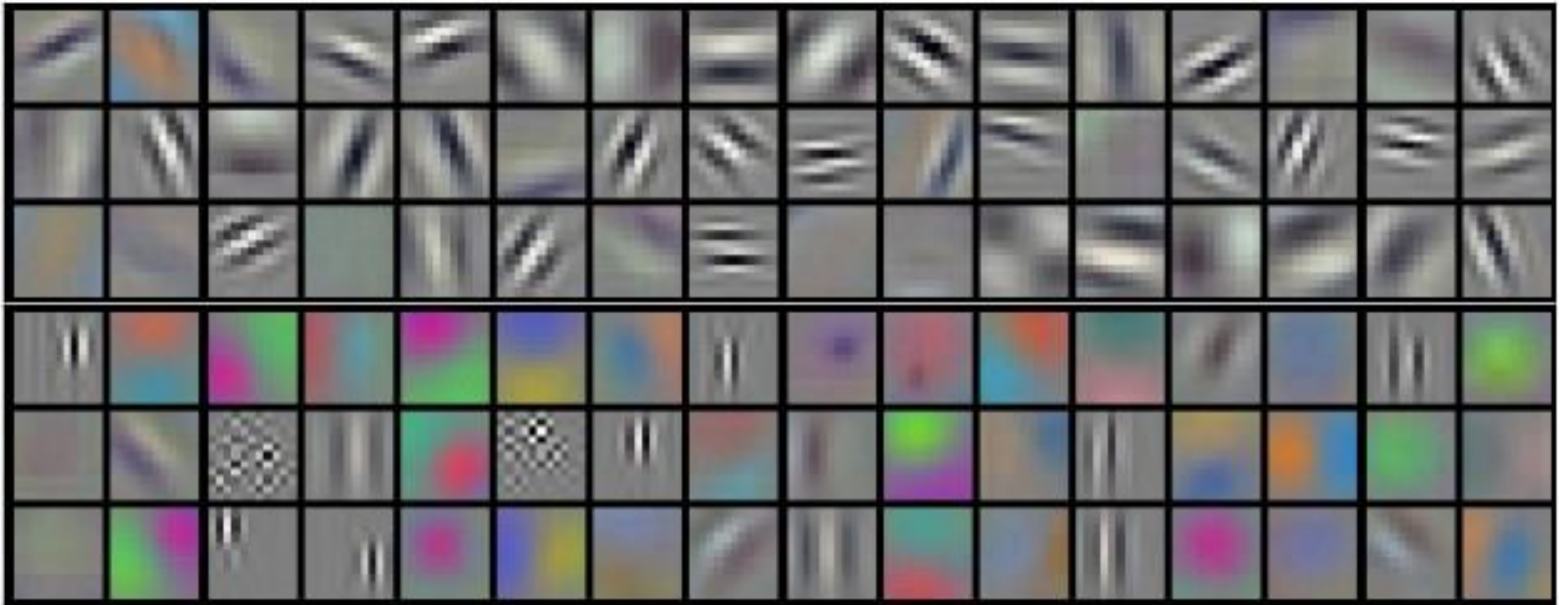
- 1950 **Test Alana Turinga**
- 1951 Programy szachowe
- 1956 Termin „Sztuczna inteligencja” (AI artificial intelligence) Johna McCarthy’ego
- ... Nowe algorytmy uczenia
- 1965 **Uczenie głębokie** Alexey G. Ivakhnenko, Valentin Lapa
- 1965 Joseph Weizenbaum, ELIZA, interaktywny **program dialogowy**
- 1980 Kunihiro Fukushima **sieć splotowa**
- 1989 Yann LeCun **uczy sieć splotową**

Siatkówka oka



Ewolucja sieci do rozpoznawania obrazów

Na co „reaguje” sieć neuronowa nauczona analizowania obrazów?

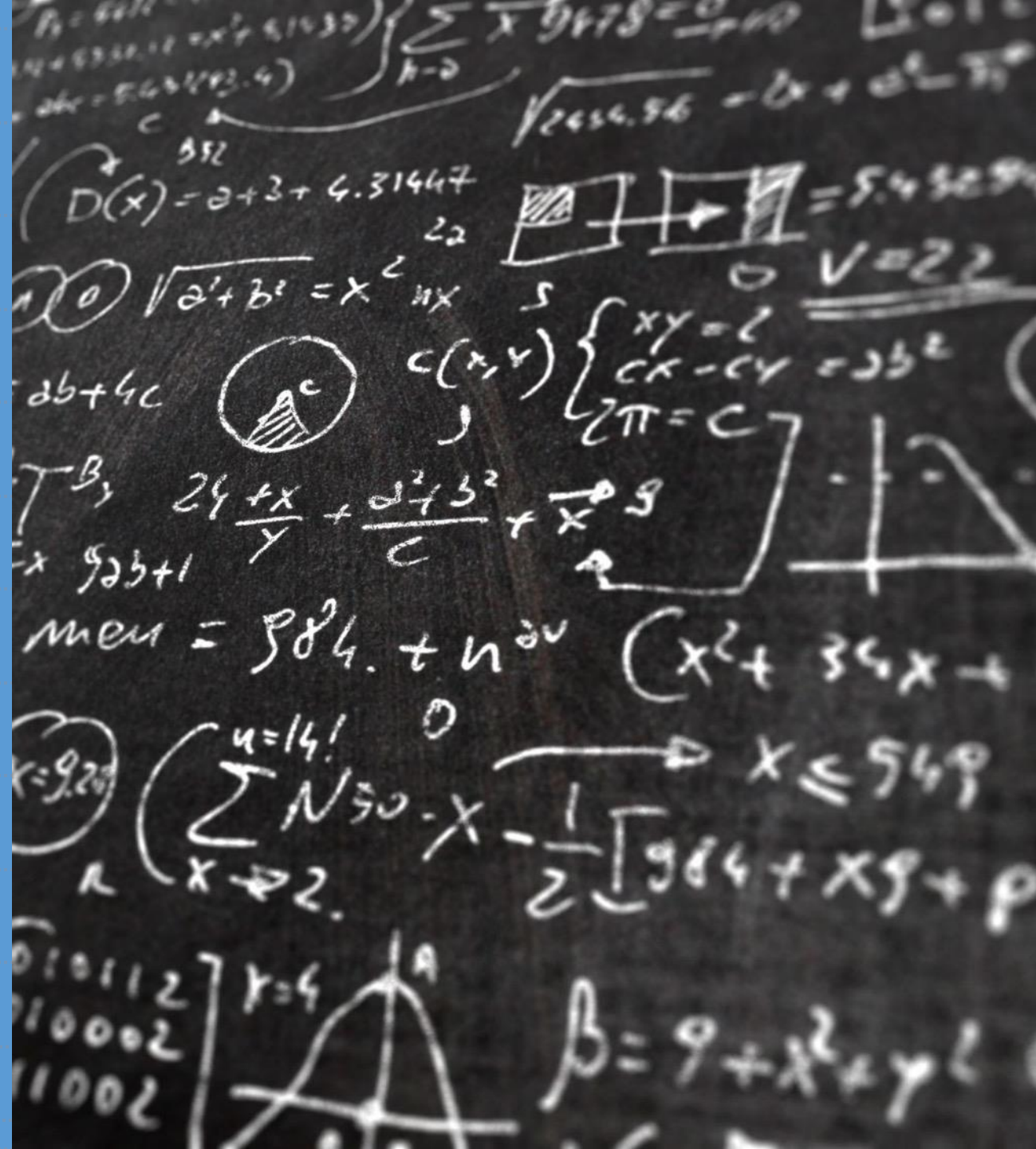


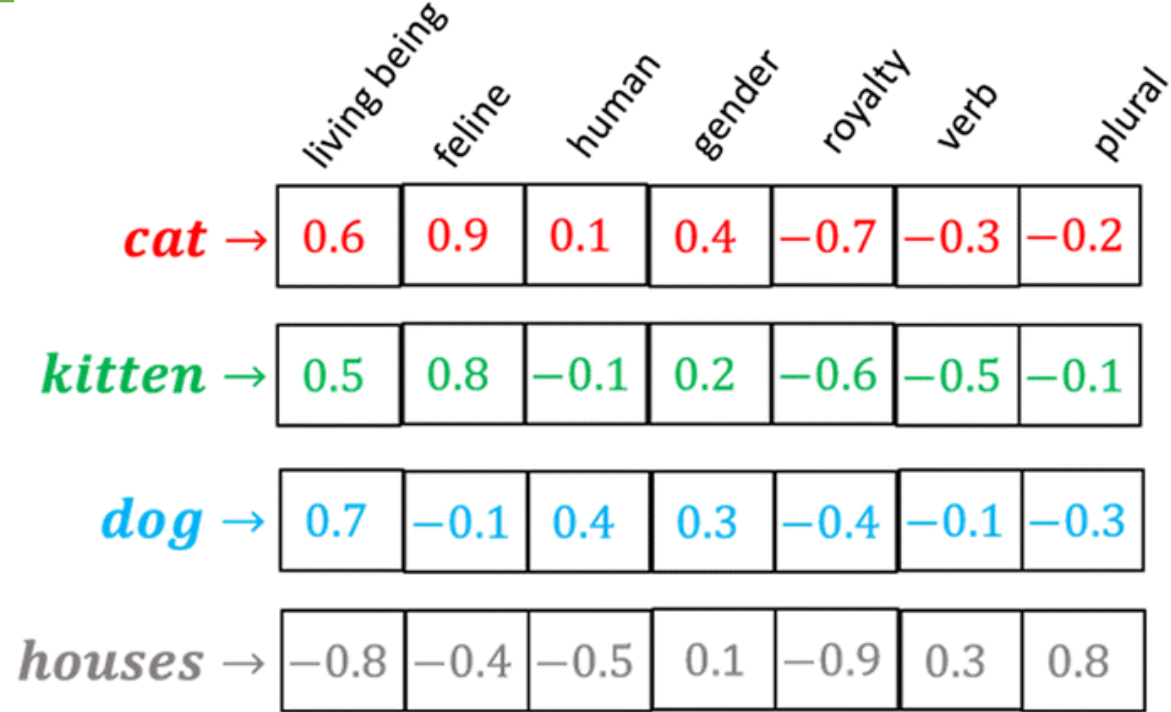
1997-2017

- Nowe algorytmy uczenia
- Wzrost **wydajności** maszyn
- 1997 wygrana DeepBlue nad Garrym Kasparovem
- 1997 Hochreiter, Jürgen . "Long Short-Term Memory". Neural Computation
- Lata 1990-2010 roboty, zabawki, algorytmy wyszukiwania stron internetowych, pierwsze samochody autonomiczne
- 2011 Siri, Cortana, Asystent Google
- 2015-16 Zwycięstwa AlphaGo
- 2017 Architektura **transformer**

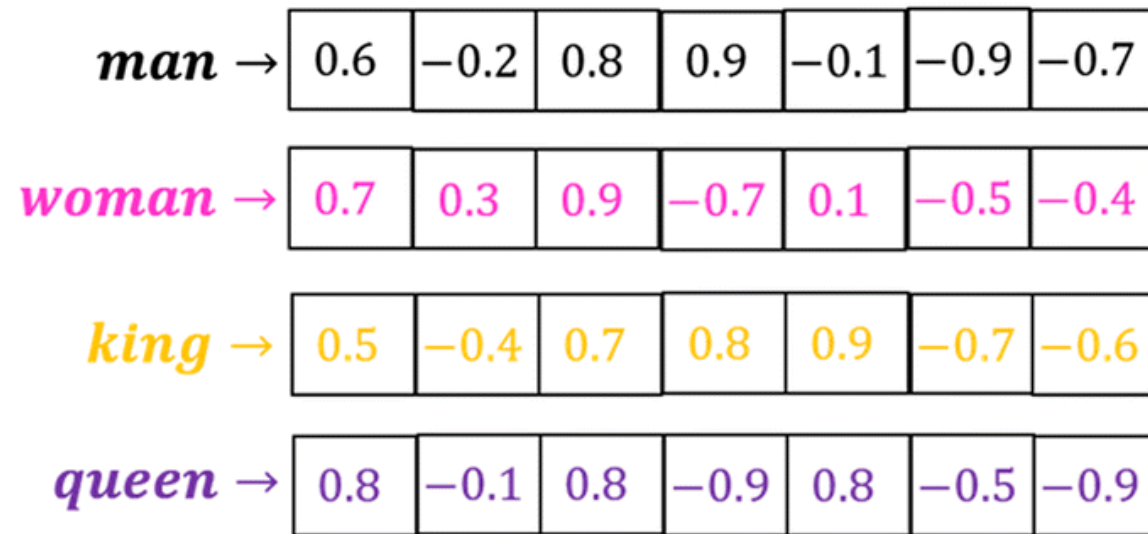
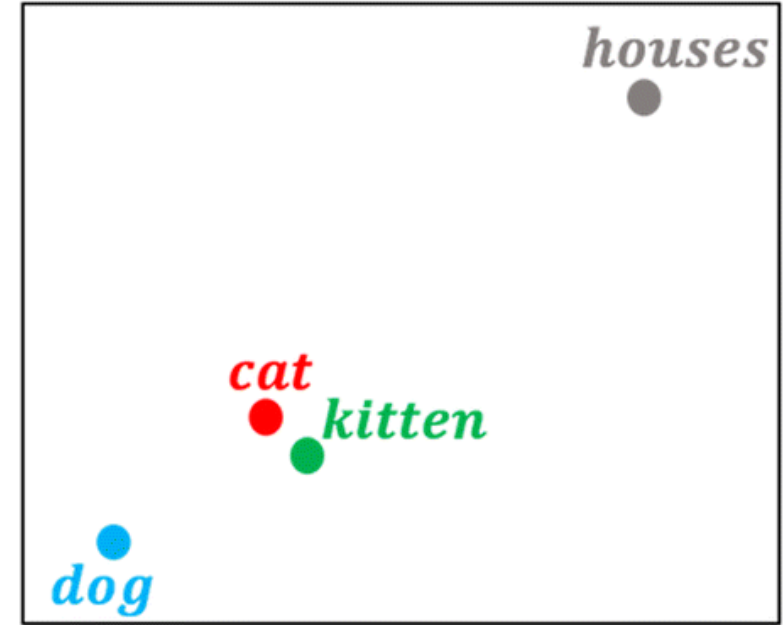
▶ Dlaczego algorytmy SI działają?

Dowolne koncepcje
wyrażane są jako liczby,
obrazy, tekst, mowa,
muzyka

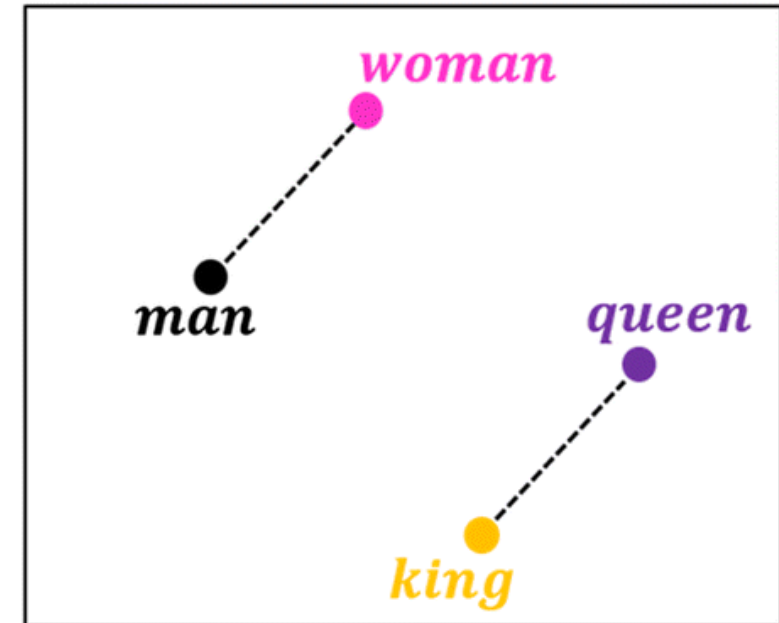




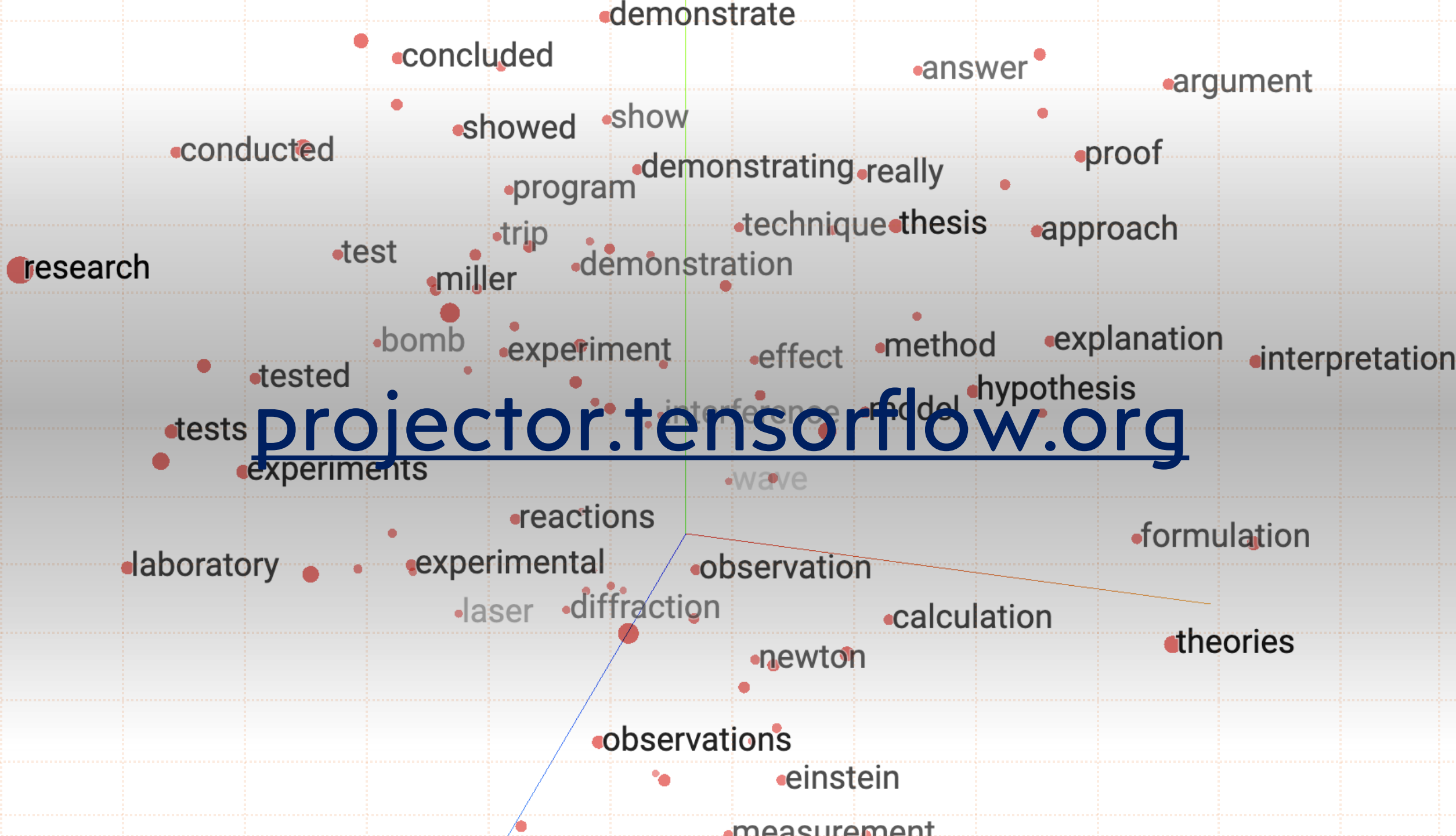
Dimensionality reduction of word embeddings from 7D to 2D



Dimensionality reduction of word embeddings from 7D to 2D

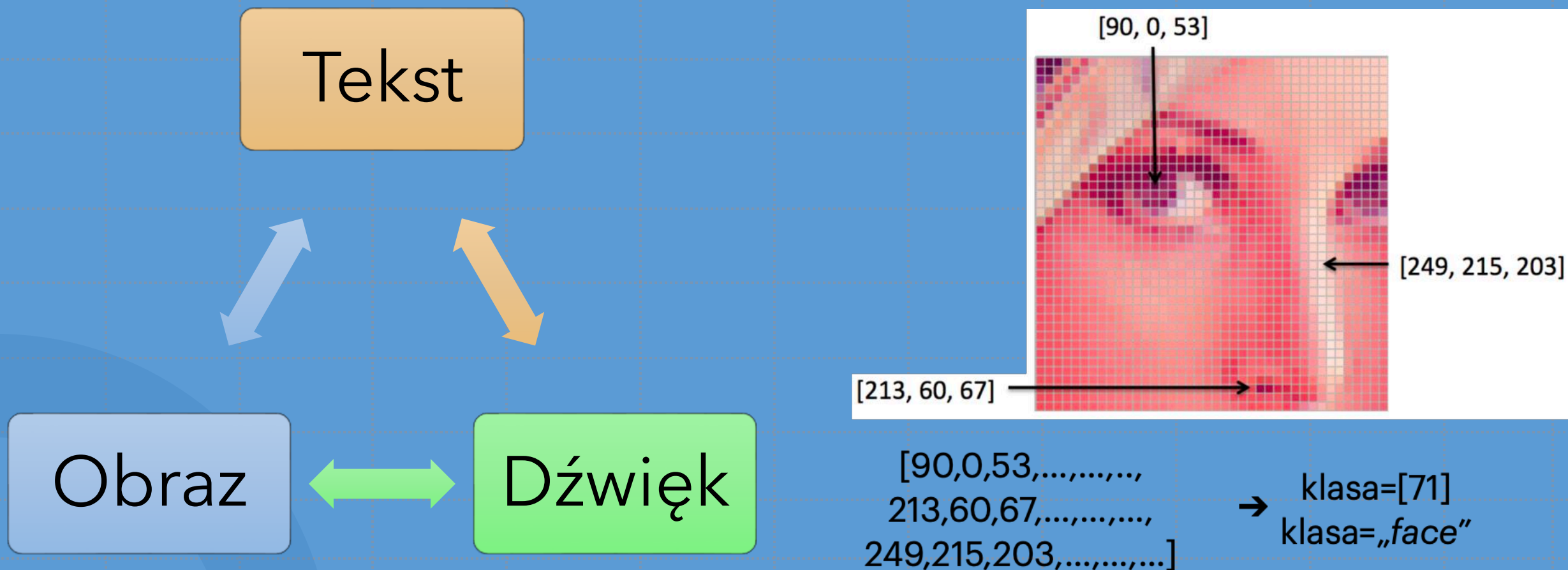


projector.tensorflow.org



Dlaczego algorytmy SI działają?

- SI realizuje operację: koncepcja wejściowa => koncepcja wyjściowa



2020 - obecnie

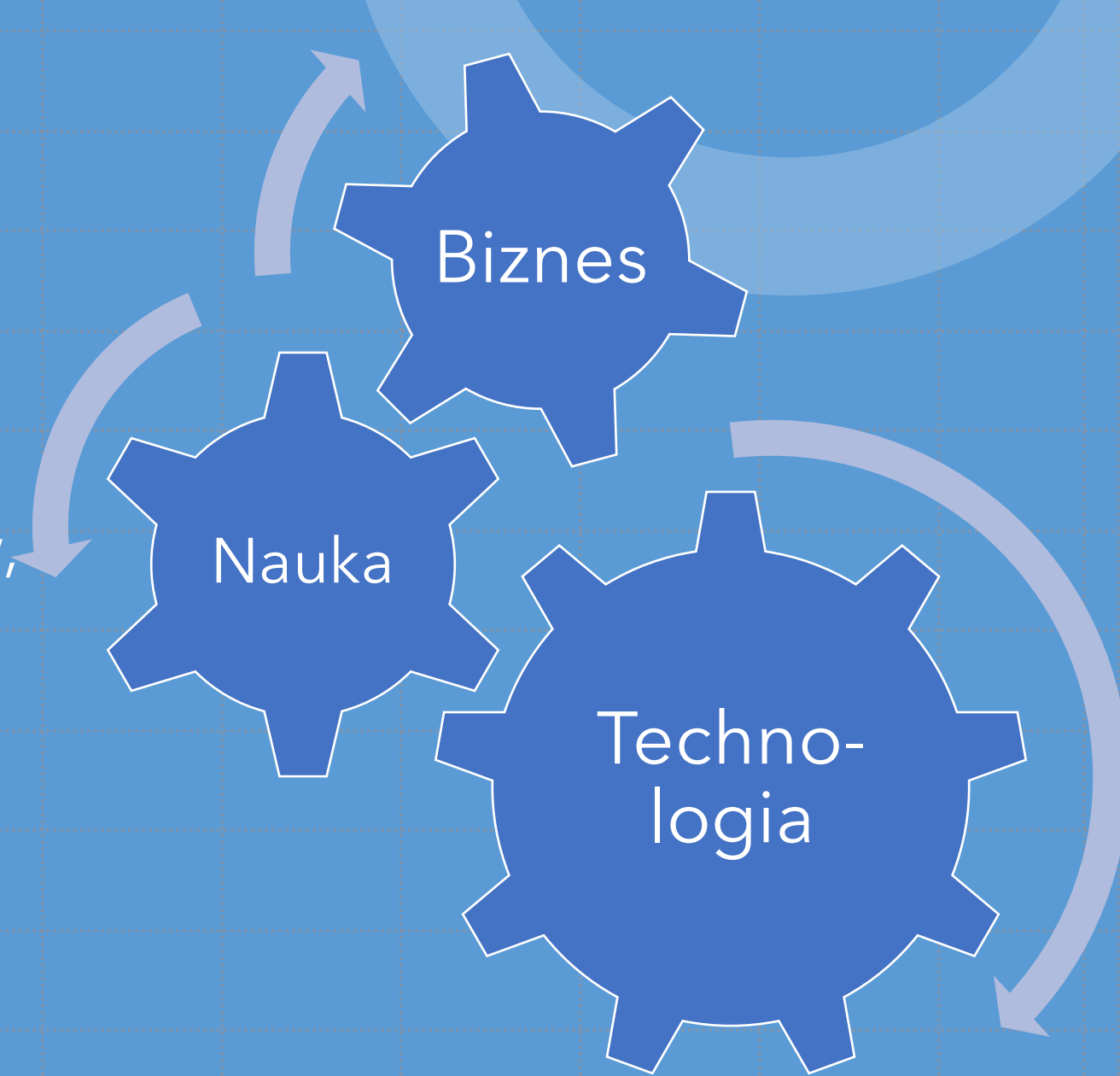
2020 GPT-3 - 175 miliardów parametrów

2022 ChatGPT

2023 GPT-4 - 1 trylion parametrów,
koszt 100 mln \$

Multimodalne
Uniwersalne
Dostosowane
Elastyczne

Wciąż nieinteligentne





Uczenie maszyn

Zastosowania SI

- Klasyfikacja – wynik **jakościowy**
- Regresja – wynik **ilościowy**
- Generowanie treści – wynik w postaci **danych**

Scenariusze uczenia

1. Z nauczycielem (nadzorowane) – **znane poprawne odpowiedzi**
2. Uczenie ze wzmocnieniem pozytywnych zachowań – **znamy tylko sposób oceny rezultatu**
3. Nienadzorowane i samo-nadzorowane – **abstrakcyjne zadania**
4. Generacyjne sieci przeciwstawne GAN – **„oszukiwanie krytyka”** potrafiącego wykrywać sztuczne dane

Cykl życia

1. **Gromadzenie danych**
2. Oczyszczanie, przygotowanie danych
3. Definiowanie architektury modelu i sposobu nauki
4. **Uczenie**
5. Testowanie
6. Eksploatacja i powrót do 1.

Uczenie nadzorowane

- Gdy dostępne jest wiele danych z adnotacjami - **z góry znane są poprawne odpowiedzi**
- Model uczony: *„Dobrze, ale trochę źle - popraw się, zmniejszając/zwiększając trochę ten parametr.”*
 - Wsteczna propagacja błędu
- Uwaga: nadmierne wyuczenie = brak generalizacji

ImageNet Challenge

<https://thegradient.pub/nlp-imagenet/>

IM  GENET

- 1,000 object classes (categories).
- Images:
 - 1.2 M train
 - 100k test.



Uczenie ze wzmocnieniem

- Gdy model sam realizuje dynamiczne zadania
- **znany sposób oceny skutku działania**
- Model uczony (tresowany): *„Nie wiem jak to robiłeś, ale oceniam, że efekt jest dość zadawalający”*
- Stosowane w **uczeniu ChatGPT** – uczenie ze wzmocnieniem z oceną przez człowieka – *„na zadane pytanie ta odpowiedź była lepsza niż inne, tamta dość dobra, ostatnia zaś okropna”*.

„Baza wiedzy” w algorytmach GPT

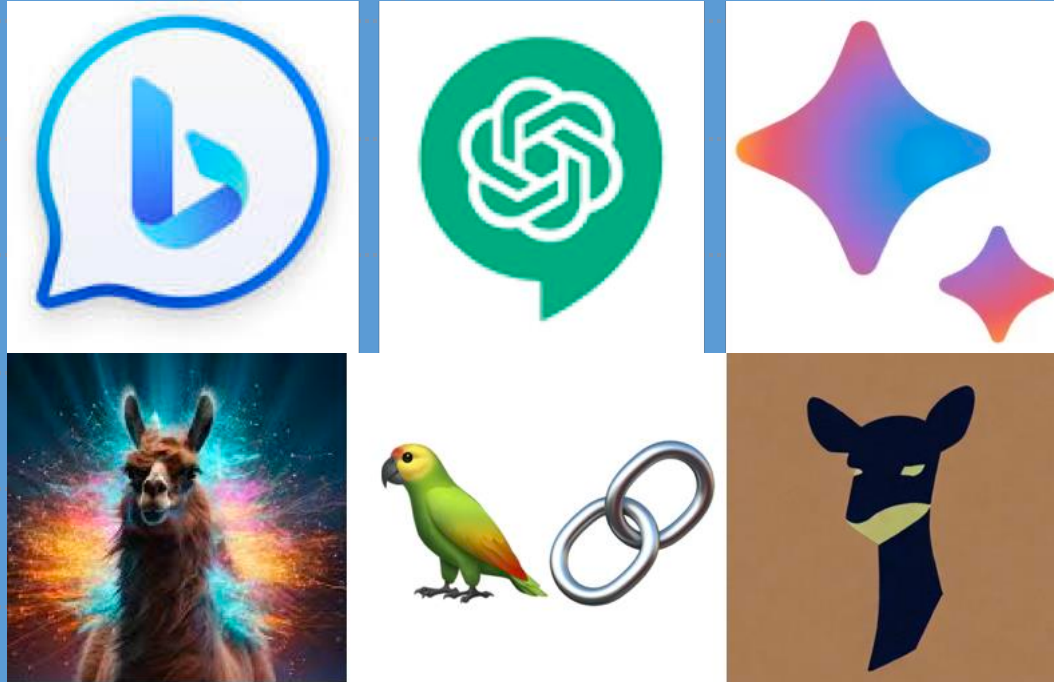
- GPT-3 uczony danymi zebranymi **do 2021 r.**
- Brak **aktualizacji wiedzy** bez ponownego uczenia
- „Wiedza” to „**statystyka**” występowania tokenów w danych uczących
- **„Halucynacje”**
 - Myślenie (nie)logiczne
 - ... Rozwiązaniem są „wtyczki”
- **Wyciek danych**
 - ...Rozwiązaniem jest „demokratyzacja” GPT

Uczenie nienadzorowane, samo-uczenie

- Identyfikowane podobieństwa
- **Grupowane** obiekty podobne



Uczenie generatywne



<https://christiankromme.com/dall%C2%B7e-introducing-outpainting>



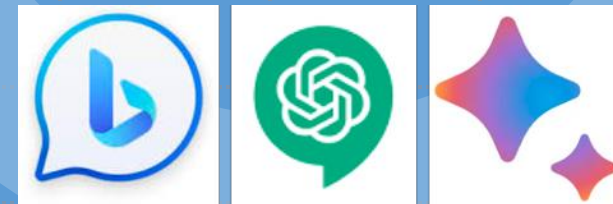
<https://jalammar.github.io/illustrated-stable-diffusion/>

Narzędzia LLM

- <https://chat.openai.com/>
- <https://bing.com/chat>
- <http://bard.google.com/>
- Prompt RTF - Role / Task / Format
- Jesteś nauczycielem szkoły podstawowej, przygotowujesz lekcję omawiającą w żartobliwy sposób losy Cesarstwa Rzymskiego, dla uczniów klasy ósmej. Napisz krótkie zdania na temat najważniejszych wydarzeń.

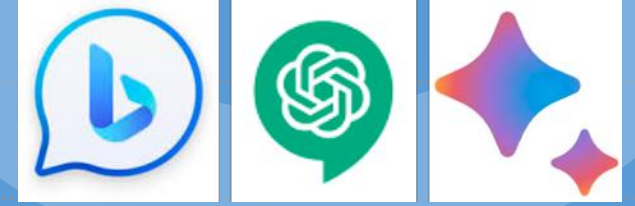


Google Bard



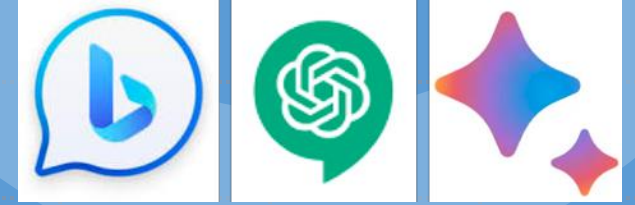
- Dzień dobry, ósmoklasiści! Dziś porozmawiamy o losach Cesarstwa Rzymskiego, ale w trochę mniej poważny sposób.
- Zaczniemy od początku. Jak to się stało, że Rzym stał się cesarstwem?
- Cóż, to zasługa Juliusza Cezara. Był on takim dobrym wodzem, że nawet gdy stracił rękę w walce, to dalej walczył, bo miał taką siłę woli.
- Po śmierci Cezara wybuchła wojna domowa, w której zwyciężył Oktawian August. On to wprowadził w Rzymie system monarchii, czyli cesarstwa.
- Oktawian August był dobrym cesarzem, ale jego następcy już niekoniecznie.
- Cesarz Neron był tak zły, że palił ludzi na stosie, żeby się ogrzać.
- Cesarz Kaligula był tak szalony, że myślał, że jest koniem.
- Cesarz Heliogabal był tak dziwny, że chciał, żeby ludzie czcili słońce jako boga.
- W końcu, po kilkuset latach, Cesarstwo Rzymskie upadło.
- Głównym powodem upadku była korupcja. Rzymscy urzędnicy byli tak chciwi, że kradli pieniądze z kasy państwa.
- Innym powodem była inwazja barbarzyńców. Ci ludzie byli tak silni, że Rzymianie nie byli w stanie ich pokonać.
- No i to już wszystko na dzisiaj. Mam nadzieję, że się wam podobało.
- A na zakończenie mam dla was żart:
- Dlaczego Rzymianie nie mogli znaleźć dobrego fryzjera? - Bo wszyscy byli zajęci budowaniem Koloseum.

OpenAI Chat



- "Witajcie, uczniowie! Dziś rozpoczynamy naszą podróż w czasie do Cesarstwa Rzymskiego - jednego z największych hitów historycznych, które trwały dłużej niż najpopularniejszy serial!"
- "Cesarstwo Rzymskie zaczęło się jako małe miasteczko, ale szybko stało się wielkim imperium. To był prawdziwy "start-up" starożytności!"
- "Ale uwaga, drodzy uczniowie! Podobno Rzymianie byli tak zajęci budowaniem swojego imperium, że zapomnieli zrobić kalendarz na 476 rok. Przegapili go zupełnie!"
- "Cesarstwo Rzymskie było znane z wielu rzeczy, ale niekoniecznie z równości płci. Kobiety nie miały takich samych praw jak mężczyźni, ale to się zmieniło w czasach nowożytnych. Czy to oznacza, że Rzymianki były pierwszymi feministkami?"
- "Nawet wielcy wodzowie Rzymu potrzebowali chwili relaksu. Juliusz Cezar miał swoje słynne przysłowie: "Wszystko jest dozwolone w miłości i na feriach w Egipcie"."
- "Ale nie wszystko było takie różowe. Upadek Cesarstwa Rzymskiego to prawdziwa tragedia. Wydaje się, że Rzymianie zostali przeważeni przez fakt, że ich togami stały się za ciężkie do noszenia!"
- "Na szczęście dziedzictwo Rzymu żyje dalej w naszej kulturze, od architektury po system prawny. Nawet nasze wcześniejsze lekcje gramatyki są trochę jak łacina, która była językiem Rzymian!"
- "Podsumowując, Cesarstwo Rzymskie to było naprawdę wielkie imperium, ale tak jak w życiu, niektóre rzeczy nie trwają wiecznie. Następnym razem, jak zobaczycie koloseum, zastanówcie się, czy Rzymianie mieli na pewno bilet wstępu!"
- Oczywiście, żartując na lekcji, zawsze ważne jest, aby zachować szacunek dla historii i kultury Rzymu oraz dołączyć do żartów odpowiednie informacje historyczne.

Bing Chat



- Oto kilka zabawnych faktów o Cesarstwie Rzymskim, które mogą Ci się przydać:
- Założenie Rzymu (753 p.n.e.): Romulus i Remus, dwaj bracia karmieni przez wilczycę, postanowili założyć miasto. Romulus zabił Remusa w sprzeczce o to, gdzie powinno być miasto, i nazwał je Rzymem. Niezły początek, co?
- Początek Cesarstwa (27 p.n.e.): Oktawian August zdecydował, że republika to nie dla niego i stał się pierwszym cesarzem Rzymu. Zmienił swoje imię na Augustus, co oznacza "wielki". Trochę zarozumiały, prawda?
- Era Pax Romana (27 p.n.e.-180 n.e.): To były czasy pokoju i prosperity w Cesarstwie Rzymskim. Cesarze tacy jak Trajan i Hadrian zbudowali wiele budowli i pomników. Można powiedzieć, że byli prawdziwymi miłośnikami sztuki!
- Upadek Rzymu (476 n.e.): Ostatecznie Cesarstwo Rzymskie upadło, kiedy ostatni cesarz zachodniorzymski, Romulus Augustulus, został obalony przez wodza germańskiego Odoakra. Ironia losu, że ostatni cesarz miał takie same imię jak założyciel Rzymu!
- Pamiętaj, że historia może być zarówno pouczająca, jak i zabawna! 😊

The background is a solid blue color with a grid of small, light blue dashed lines. Numerous 3D question marks are scattered across the surface, some appearing to float or be attached to the grid. At the bottom center, there is a faint, glowing image of a cityscape with buildings and a bright light source, possibly a sun or moon, creating a lens flare effect.

Problemy, wymagania i nowe kompetencje

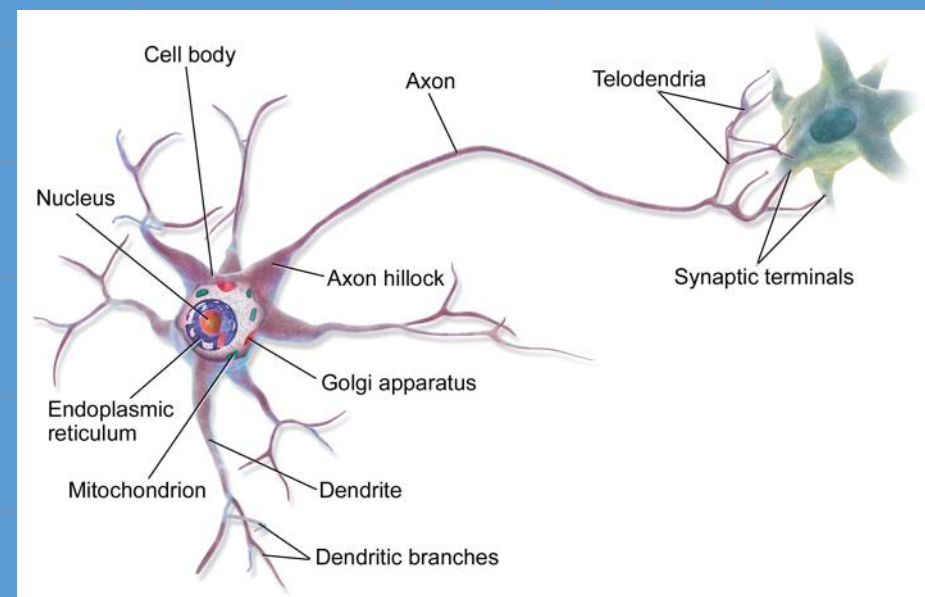
Wyjaśnialna SI

Zawsze, gdy:

- **Automatyczne** przetwarzanie danych i jego skutek **nas dotyka**

Człowiek ma niezbywalne prawo do:

- Kwestionowania decyzji
- Żądania wyjaśnienia
- Problem „czarnej skrzynki”



Etyka uczenia maszynowego

- Odpowiedzialność autora/administratora/użytkownika? za błędy
 - Brak dyskryminacji, ukrytych tendencji
 - Wyjaśnialność
-
- Odpowiednio **przygotować dane** i przeprowadzić trening
 - Uwaga na celowo zatrute dane (ang. data poisoning)
 - Uwaga na ataki

Nowe miejsca pracy, inne znaczenie dotychczasowych zawodów

- Inżynier danych (to już nic nowego)
 - Kurator danych
 - Inżynier promptu
 - Audytor SI
-
- Ekspert dziedzinowy (Human in the loop)

Korzyści

- Natychmiastowa dystrybucja wiedzy
- Internet i SI w dowolnym miejscu na świecie
= dostęp do SI medycznego, dydaktycznego
- Odkrycia naukowe: związki chemiczne, struktury białek, leki, wymarłe języki, fizyka, astronomia, inżynieria
- Optymalizacja procesów

Podsumowanie - co potrafi mózg a czego nie potrafi SI

- Logiczne myślenie, generalizacja (uczenie na jednym przykładzie)
- Świadomość i samoświadomość
- Poszukiwanie własnych celów, prawdziwa kreatywność
- *„Wiem, że nic nie wiem”*
- *„Wiem, że czasem się mylę lub nie jestem pewien”*
- Do działania wystarczy 10-20 Watów energii (5 razy mniej niż laptop, GPT-3 w trakcie nauki wykorzystało prawie 2 miliony razy więcej)

Dziękuję za uwagę

piotrszczuko@gmail.com