



TECH FOR LAW & LAW FOR TECH



Sztuczna inteligencja – stan wiedzy, perspektywy rozwoju i problemy etyczne

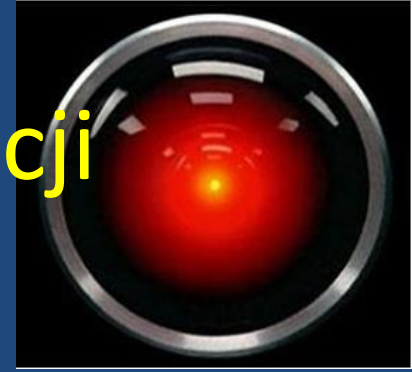
Piotr Bilski

Instytut Radioelektroniki i Techniki
Multimedialnych

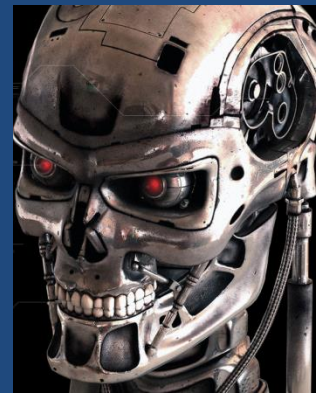
Plan wystąpienia

- Co to jest sztuczna inteligencja?
- Pojęcie słabej i silnej sztucznej inteligencji
- Stan zaawansowania prac nad sztuczną inteligencją
- Aktualne problemy w zastosowaniach praktycznych

Definicja sztucznej inteligencji (Artificial Intelligence)

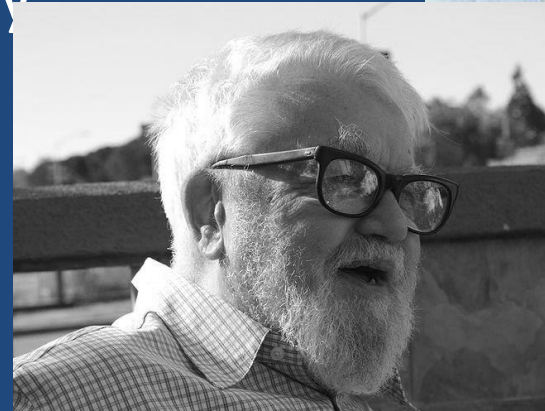
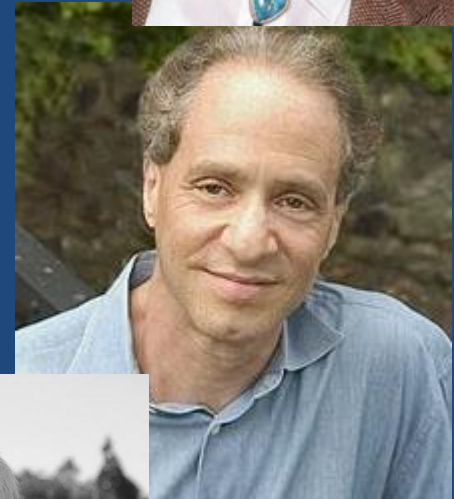
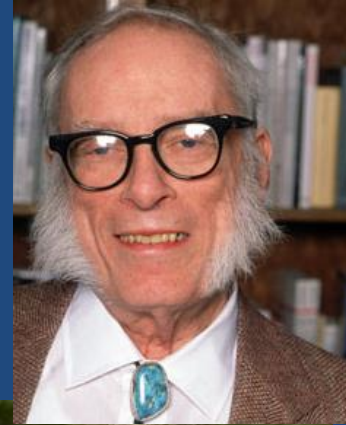


- Dziedzina informatyki zajmująca się systemami inteligentnymi (maszynami, komputerami itp.)
- Istnieje od 1956 r.
- Cel: stworzenie maszyn myślących na wzór człowieka (test Turinga)
- Obecnie wykorzystywana do rozwiązywania problemów nauki i techniki



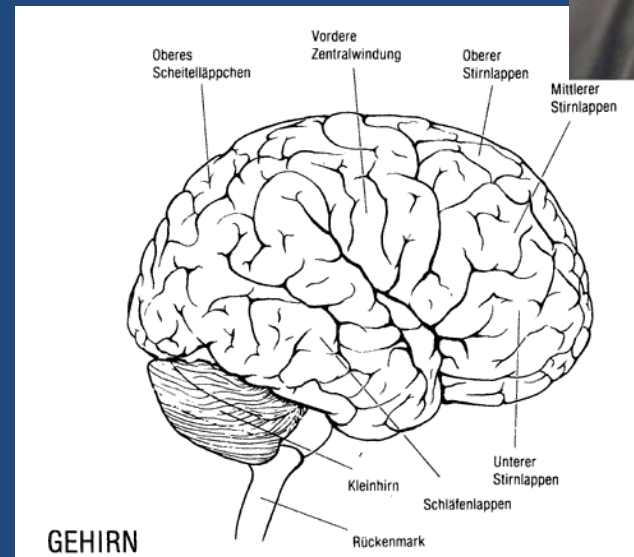
Badacze silnej sztucznej inteligencji

- Isaac Asimov (prawa robotyki)
- Ray Kurzweil (przewidywanie osobliwości)
- John Searle (teoria umysłu, Chiński pokój)
- John McCarthy (twórca pojęcia AI i cloud computing)
- Alan Turing (twórca maszyny Turinga i testu Turinga)



Problemy silnej sztucznej inteligencji

- Jak mierzyć (test Turinga)?
- Próby odwzorowania myślenia człowieka
 - Systemy regułowe
 - IF (przesłanki...) THEN (konkluzja)
 - Organizacja sieci neuronowych
- Jak osiągnąć?
- Inteligencja emocjonalna, gust, intuicja itp.



Słaba sztuczna inteligencja

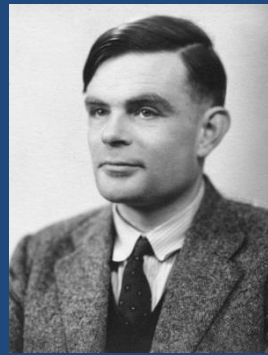
- Naśladowanie pewnych cech ludzkiego umysłu w celu rozwiązywania konkretnych problemów
- Liczne zastosowania praktyczne
- Problemy: złożoność obliczeniowa, generalizacja
- Szeroka gama metod i algorytmów
- Podstawowe zadania do rozwiązania:
 - Regresja
 - Klasyfikacja
 - Optymalizacja

Silna a słaba sztuczna inteligencja

- Automatyczne przetwarzanie bodźców zewnętrznych i podejmowanie działań zgodnych z posiadanymi informacjami
- Rozumienie swoich działań i ich znaczenia
- Samoświadomość



Test Turinga

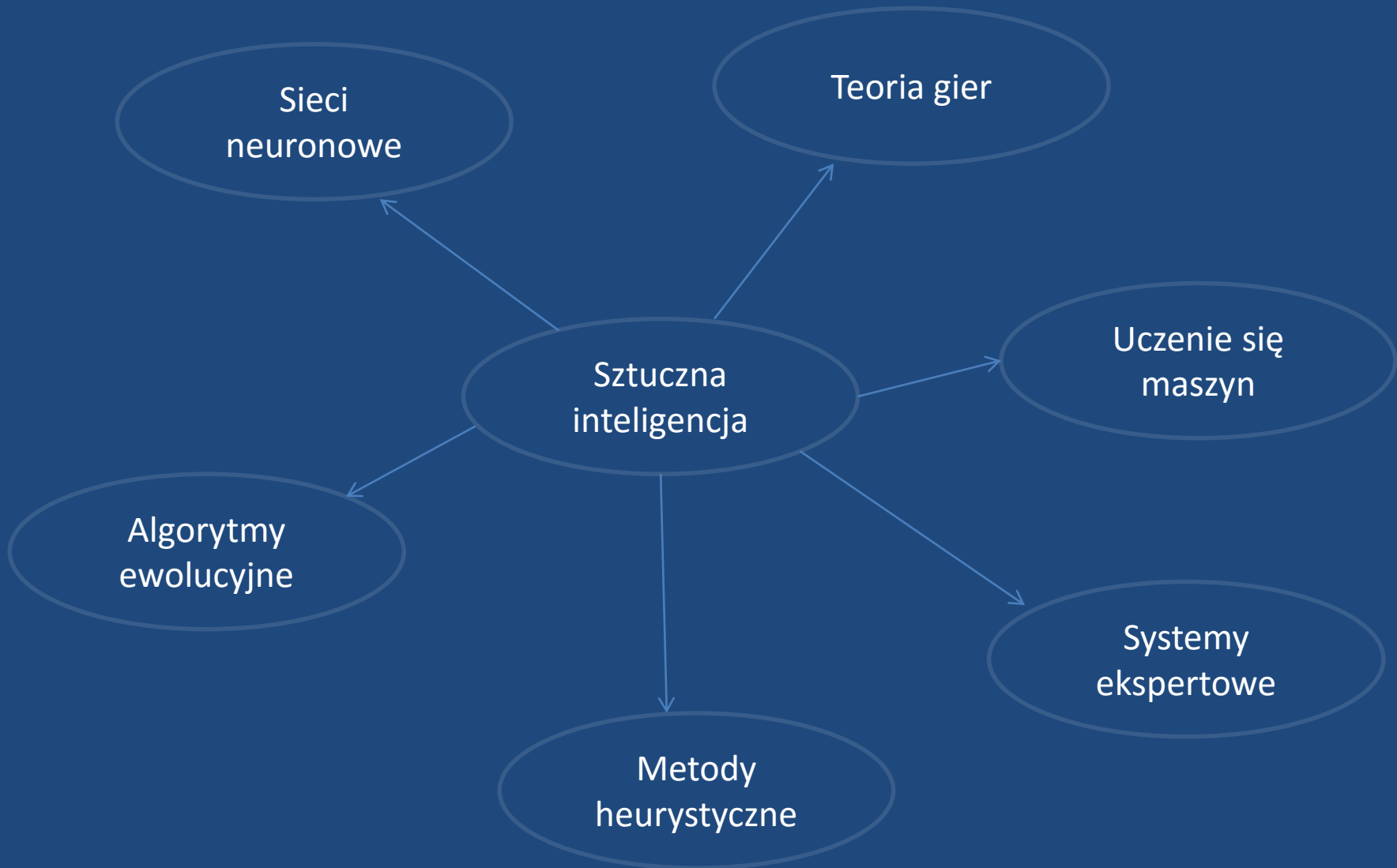


- Rozmowa między człowiekiem i komputerem
- Jeśli człowiek nie potrafi rozpoznać, czy rozmawia z maszyną, ta dostaje punkt
- Test zdany, jeśli automat uzbiera powyżej 50 procent punktów (lub więcej punktów, niż człowiek w analogicznej sytuacji)
- Czy to znaczy, że *myśli*?



VS





Regresja

- Odwzorowywanie obiektów na zbiór liczb rzeczywistych
- Uczenie z nadzorem
- Cel: utworzenie funkcji, która dla określonych danych wejściowych generuje odpowiednie wartości liczbowe
- Przykład: aproksymacja funkcji wielu zmiennych

Klasyfikacja

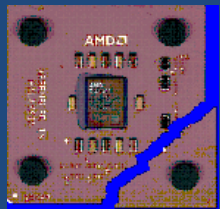
- Odwzorowywanie obiektów na zbiór liczb całkowitych
- Uczenie z nadzorem lub bez nadzoru
- Celem jest pogrupowanie obiektów podobnych do siebie lub określenie, czym różnią się obiekty należące do różnych kategorii
- Przykład: podejmowanie decyzji podczas gry na giełdzie

Optymalizacja

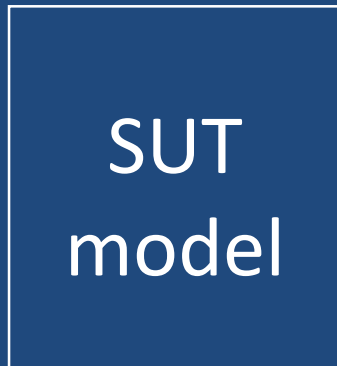
- Poszukiwanie najlepszego rozwiązania spośród wszystkich dostępnych
- Tradycyjne metody: przeszukiwanie wyczerpujące, metody lokalne
- Sprowadzanie problemu rzeczywistego do problemów podstawowych: komiwojażera (TSP), spełnialności formuł logicznych (SAT) itp.

Tworzenie wiedzy ze zbioru danych

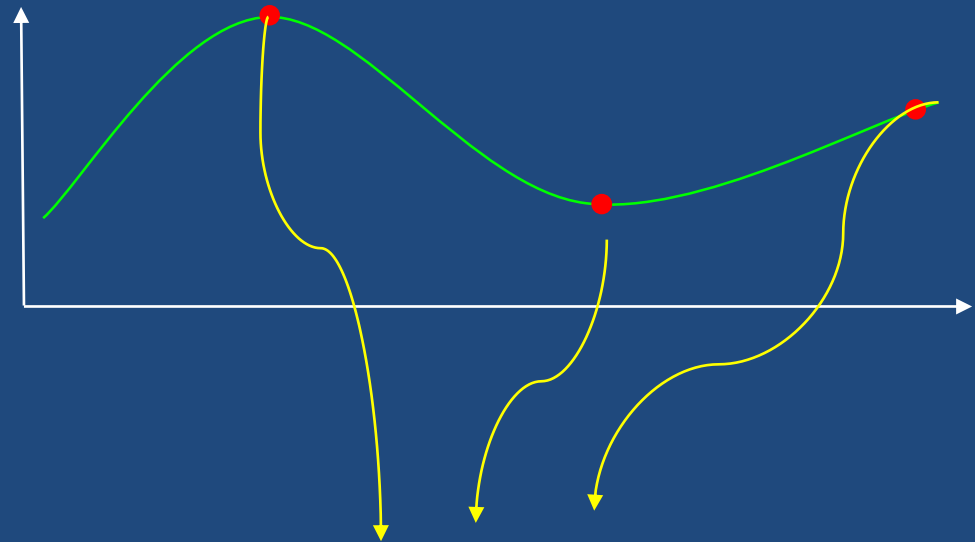
Wprowadzenie uszkodzenia



Symulacja modelu



Ekstrakcja symptomów z odpowiedzi



Generacja reguł

IF $U_1 > 1.23V$ AND $U_3 < 4.56V$ THEN FS=21

...

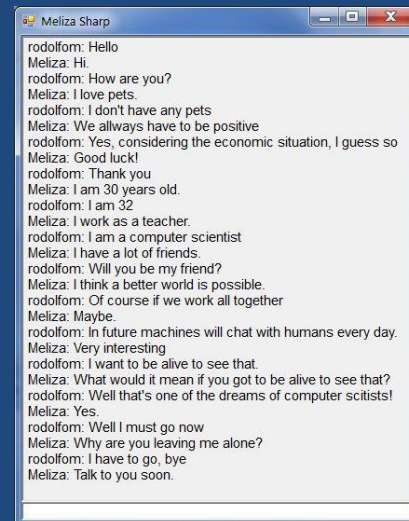
IF $U_2 > 5.32V$ AND $U_3 > 1.17V$ THEN FS=32

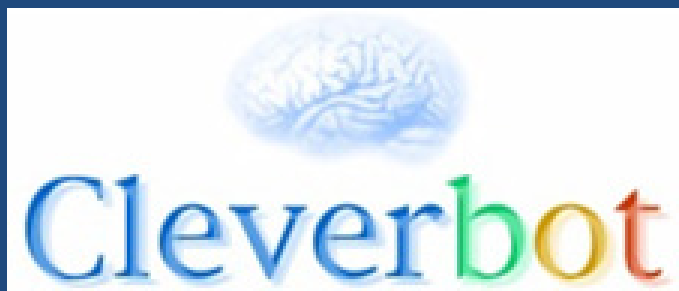


0
21
31
-32

Przykład zastosowania – przetwarzanie języka naturalnego

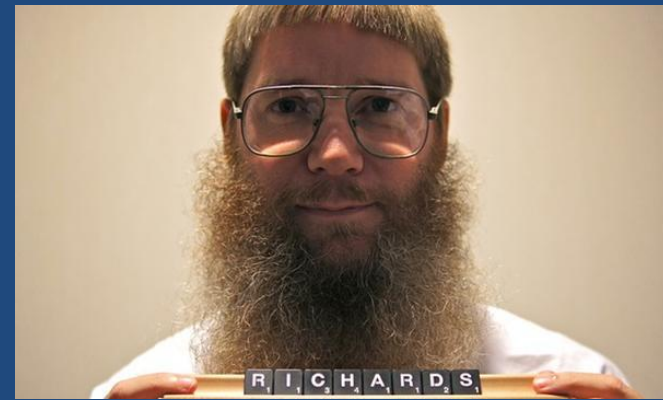
- Cel: nawiązanie konwersacji z człowiekiem
- Problem: generacja zdań w odpowiedzi na zestaw słów kluczowych (ew. kontekst)
- Rozwiązanie: drzewo przeszukiwania
- Przykład: chatboty, wirtualni asystenci





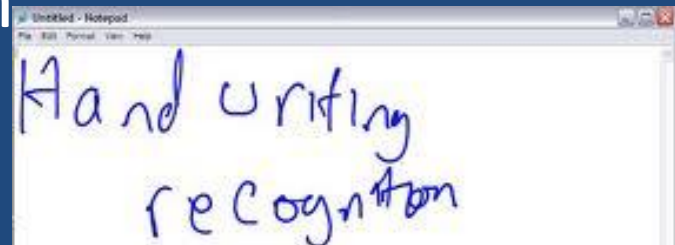
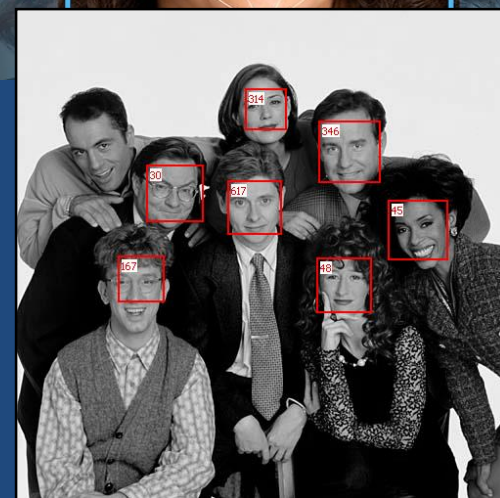
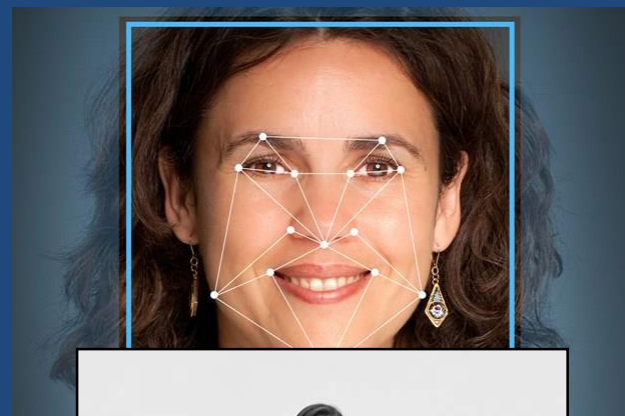
Przykłady

- Cleverbot – automat prowadzący rozmowy z ludźmi, uważa się, że przeszedł test Turinga (ok. 60% rozmówców sądziło, że rozmawia z człowiekiem)
- Nigel Richards (Nowa Zelandia) zwyciężył w turnieju francuskojęzycznego Scrabble, pomimo iż nie zna tego języka!!



Przykład zastosowania – rozpoznawanie obrazu

- Dopasowywanie wektora cech do przechowywanego wzorca
- Biometria – analiza cech biologicznych z wykorzystaniem aparatu matematycznego
- Przykłady: rozpoznawanie pisma w systemach bankowych, rozpoznawanie osób na lotniskach, sterowanie głosem w systemach komputerowych



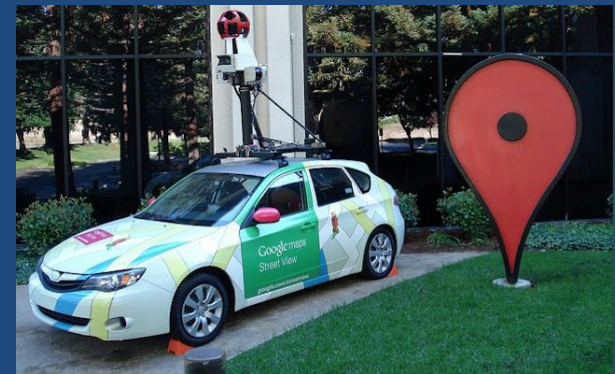
Problem 1: Rozpoznawanie twarzy

- Algorytm zastosowany w Google Photos pomylił ludzi z gorylami



Przykład zastosowania – autonomiczne sterowanie pojazdami

- Problem: poprawne sterowanie w warunkach dużej liczby szczegółów (jazda w mieście!!)
- Rozpoznawanie obrazu i dźwięku
- Poprawne zachowanie w nietypowych sytuacjach



Problem 2: Błędy samochodu autonomicznego

- Autopilot samochodu elektrycznego Tesla (Model S) nie zauważył ciężarówki na jezdni
- W wyniku wypadku zginął kierowca, Joshua Brown



A co z etyką?

- Prawa robotyki (Isaac Asimov, David Langfort, Mark Tilden)
- Niebezpieczeństwo stworzenia:
 - nowego systemu niewolnictwa
 - maszyn „niemoralnych”
- Teoria Eliezera Yudkovsky’ego – maszyny „dobre” i przyjaźnie nastawione do człowieka



Podsumowanie

- Silna sztuczna inteligencja – aktualnie nieosiągalna
- Słaba sztuczna inteligencja – użyteczna, jednak istotne problemy już widoczne
- Najbliższa przyszłość: projektowanie i ulepszanie algorytmów sztucznej inteligencji
- Opracowanie silnej sztucznej inteligencji prawdopodobnie niemożliwe bez intensywnych badań interdyscyplinarnych

Dziękuję za uwagę

ZEA

