



POLITECHNIKA POZNAŃSKA

Poznan University of Technology

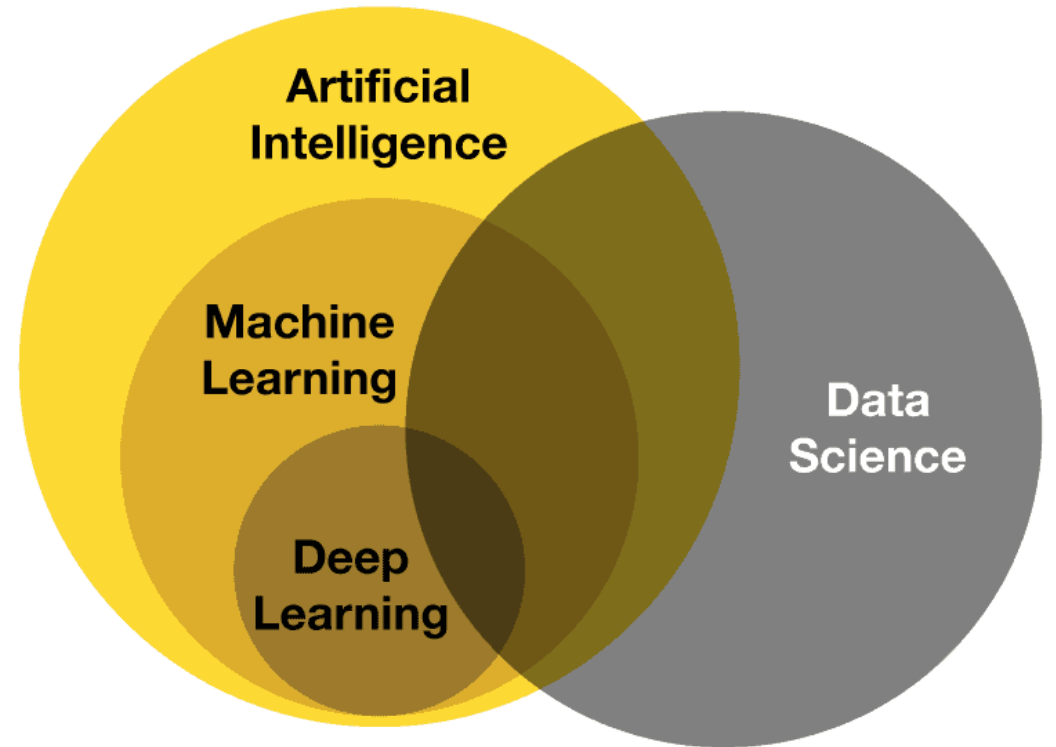
Sztuczna inteligencja

Zastosowania w medycynie i farmacji

Dariusz Brzeziński

Zamieszanie terminologiczne

- Sztuczna inteligencja (AI, AGI)
- **Uczenie maszynowe (ML)**
- Uczenie głębokie (DL)
- Eksploracja danych
- Odkrywanie wiedzy
- Data science
- Big data
- Analiza danych



Uczenie maszynowe

***Uczenie maszynowe** to obszar sztucznej inteligencji poświęcony algorytmom, które doskonalą się i dokonują predykcji na podstawie danych. Termin często używany zamiennie z terminem sztuczna inteligencja.*

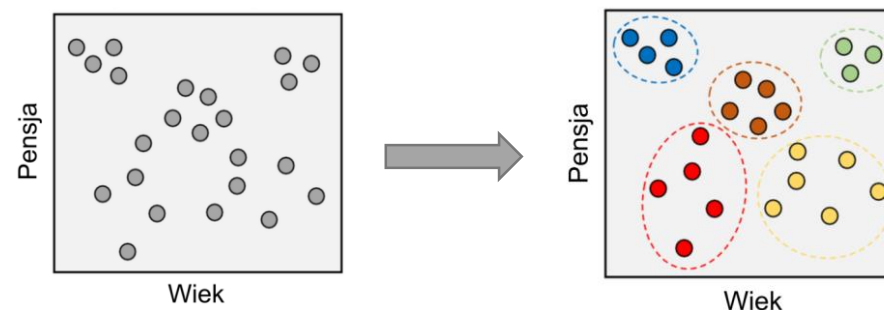
Wiele zastosowań:

- Wyszukiwarki internetowe
- Systemy rekomendacyjne
- Rozpoznawanie twarzy
- Rozpoznawanie głosu
- Filtrowanie spamu
- Ocena ryzyka kredytowego

Uczenie maszynowe

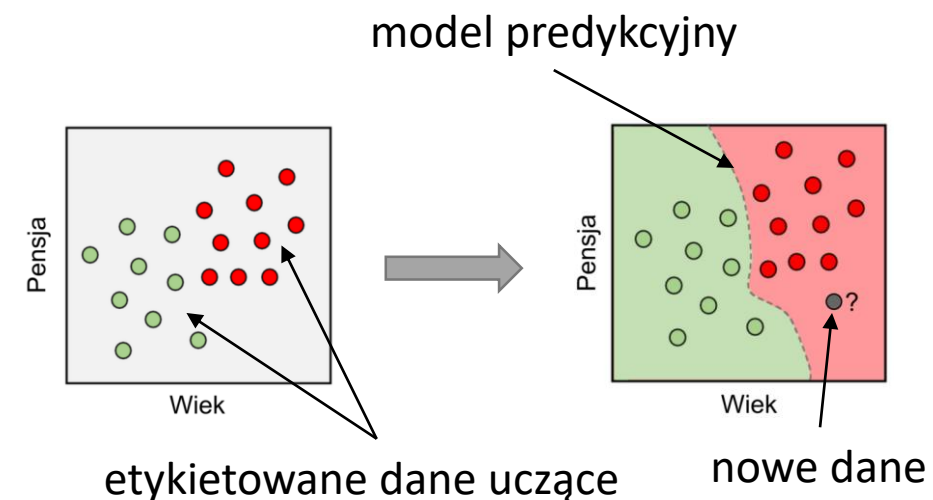
- **Uczenie nienadzorowane** (bez nauczyciela)

- Grupowanie
- Odkrywanie asocjacji



- **Uczenie nadzorowane** (z nauczycielem)

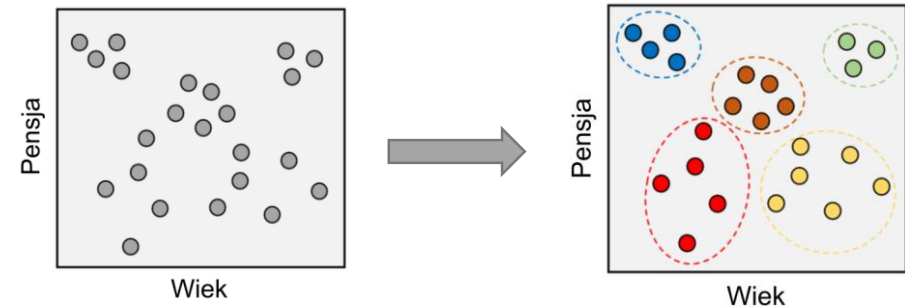
- Klasyfikacja
- Regresja
- Uczenie ze wzmocnieniem



Uczenie maszynowe

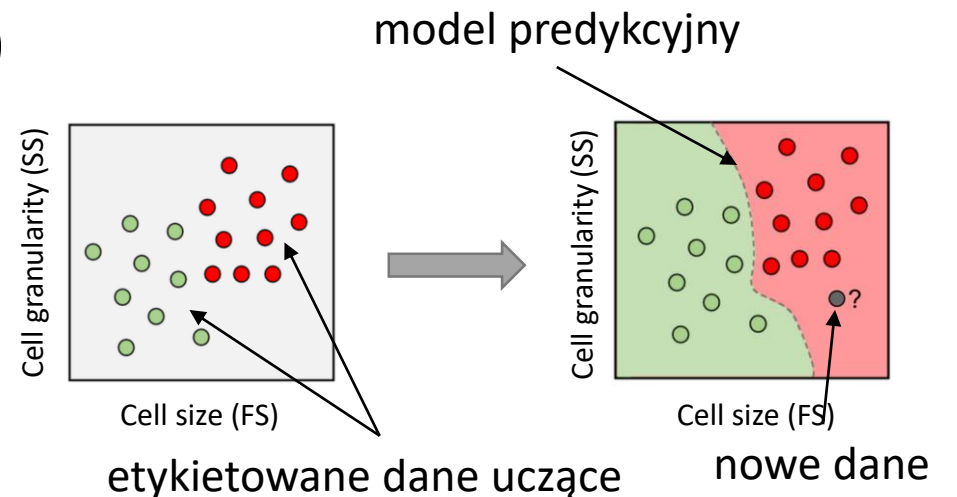
- **Uczenie nienadzorowane (bez nauczyciela)**

- Grupowanie
- Odkrywanie asocjacji



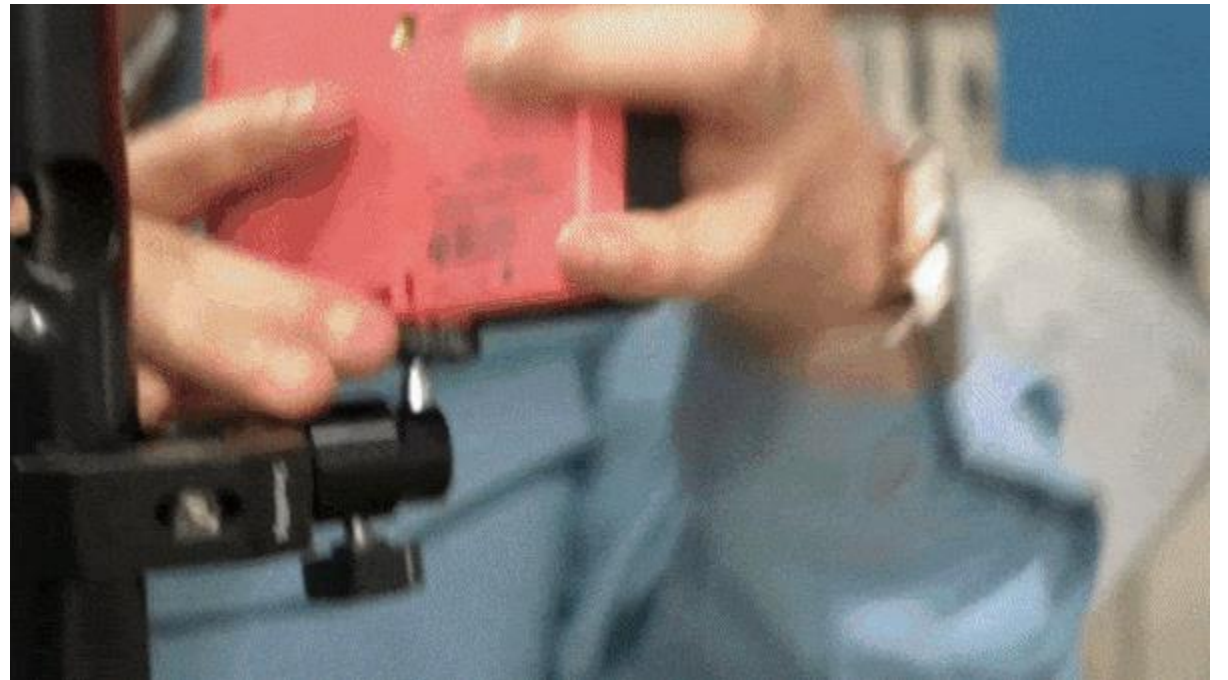
- **Uczenie nadzorowane (z nauczycielem)**

- Klasyfikacja
- Regresja
- Uczenie ze wzmocnieniem



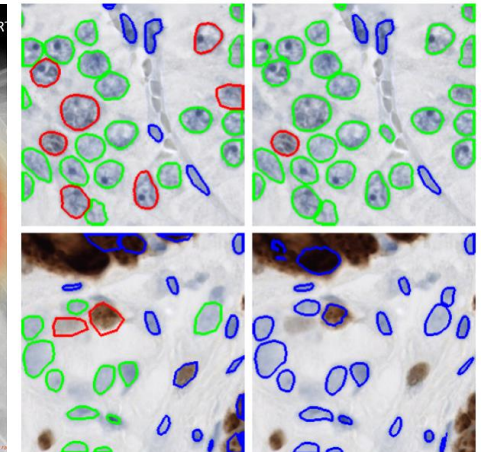
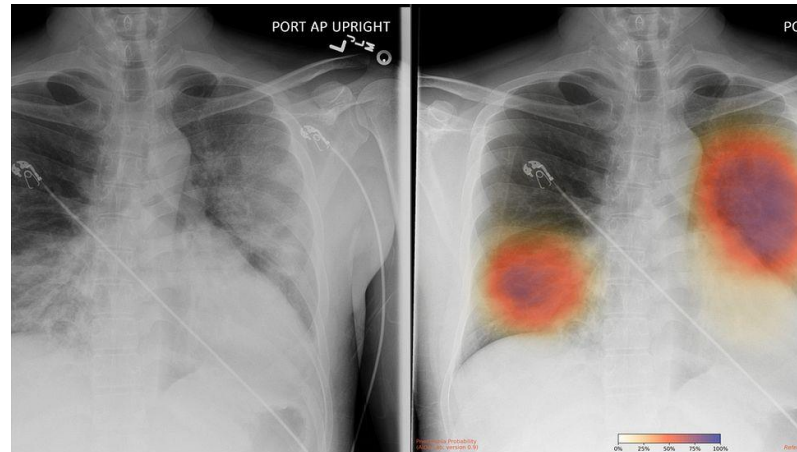
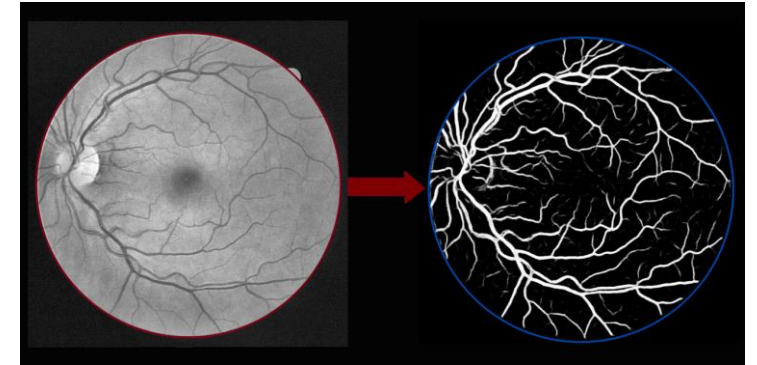
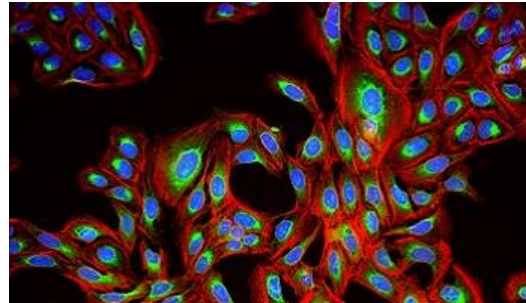
Rozpoznawanie obrazów

- Ułatwienie pomiarów ze starych urządzeń
- Montowalne w trudnodostępnych miejscach
- Podobny pomysł: pomiary z dronów (np. monitorowanie stanu linii energetycznych)
- Można też monitorować poprawność procesu (np. montażu części) w fabryce



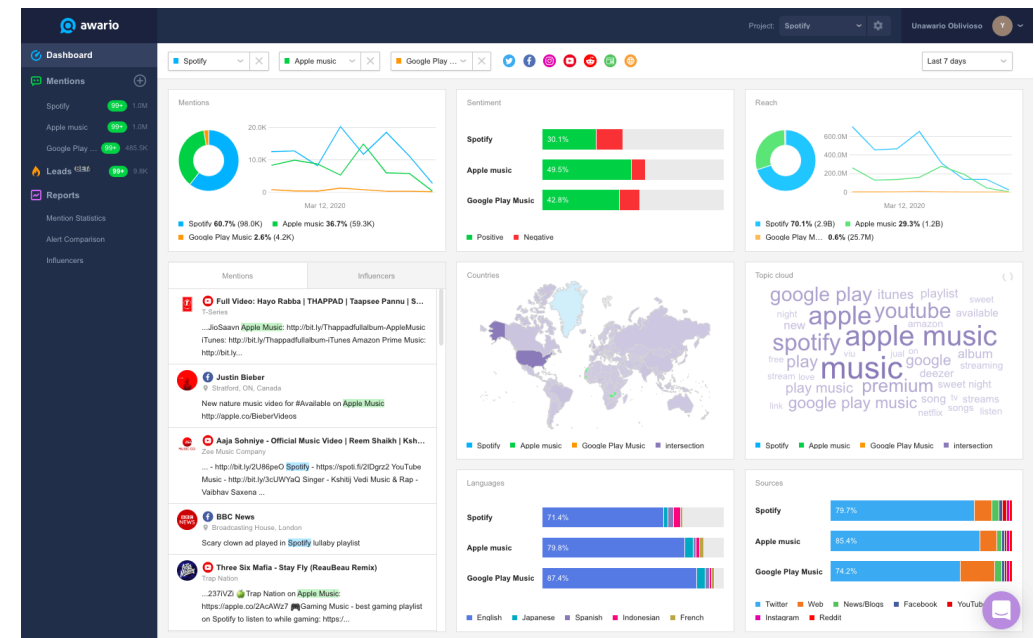
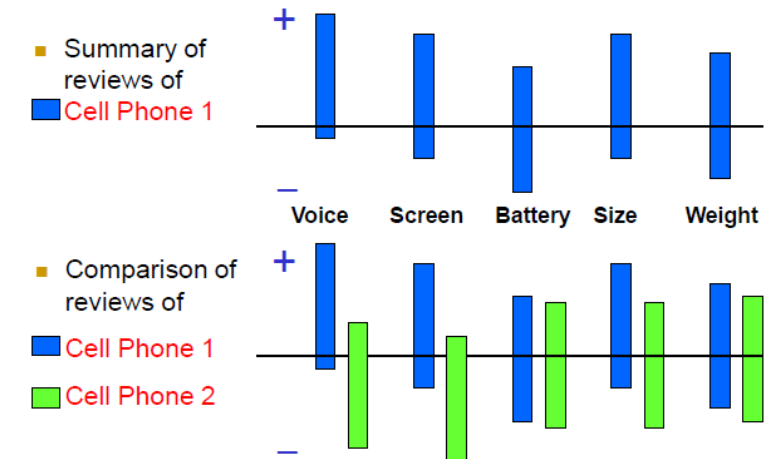
Rozpoznawanie obrazów w medycynie

- Wiele zastosowań w biologii, chemii i medycynie:
 - Histopatologia
 - Cell painting
 - Badania płuc
 - Tomografia
 - Rezonans magnetyczny
 - Badania oczu
 - Angiografia

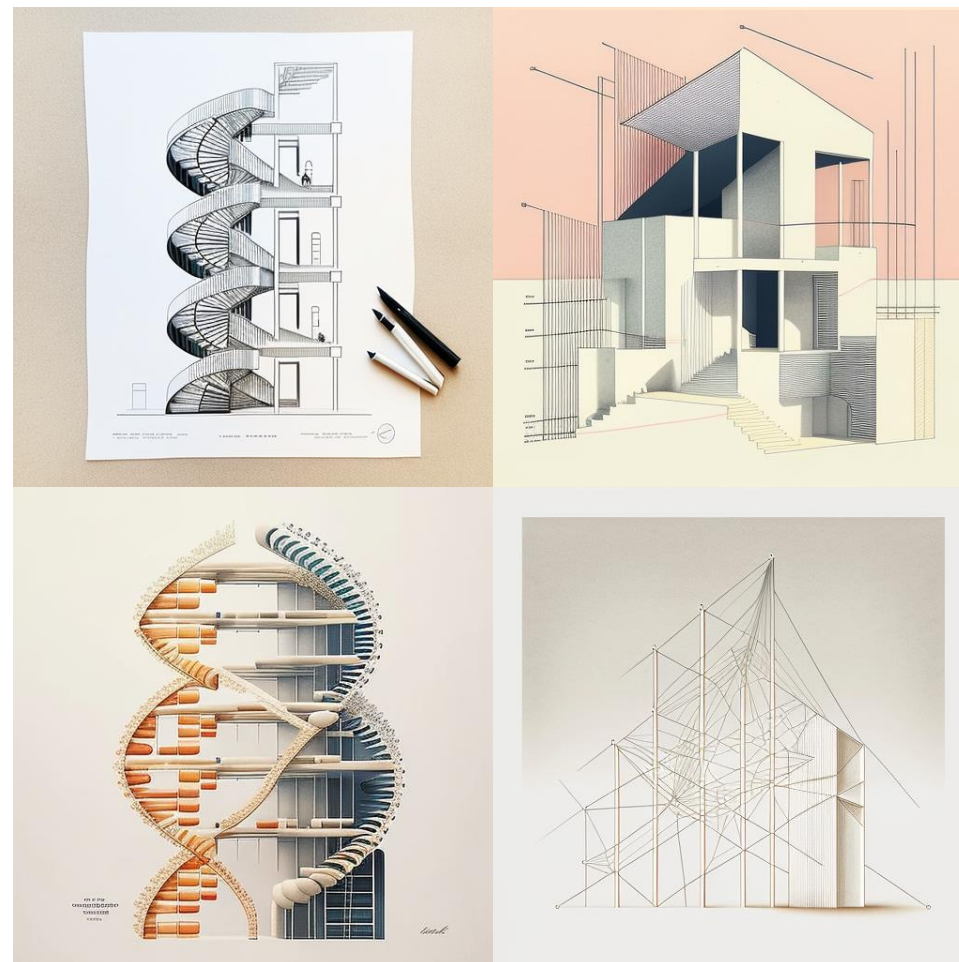


Analiza tekstu

- Firmy są zainteresowane opinią o produktach
- Badanie opinii przez analizę tekstu (ang. *sentiment analysis*)
- Ściśle powiązane z analizą mediów społecznościowych
- Zapytania prawne i wyszukiwanie regulacji
- Przeszukiwanie literatury



Medycyna i farmacja



Lab in the loop

- Większość problemów chemicznych i biologicznych ma ogromne przestrzenie możliwych rozwiązań

Sekwencje regulatorowe

$$4^n$$

Warianty somatyczne

$$4^n$$

Interakcje genetyczne

$$\binom{m}{k}$$

Sekwencje przeciwciał

$$20^n$$

Związki chemiczne

$$10^{60}$$

Lab in the loop

- Większość problemów chemicznych i biologicznych ma ogromne przestrzenie możliwych rozwiązań

Sekwencje regulatorowe

$$4^n$$

Warianty somatyczne

$$4^n$$

Interakcje genetyczne

$$\binom{m}{k}$$

Sekwencje przeciwciał

$$20^n$$

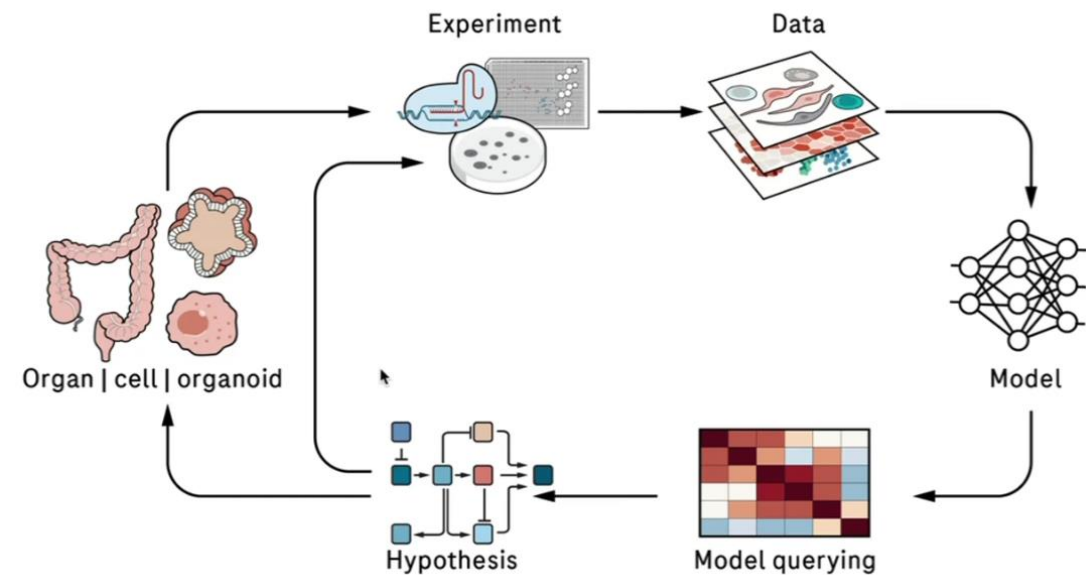
Związki chemiczne

$$10^{60}$$

Czy sztuczna inteligencja może pomóc w przeszukiwaniu tych przestrzeni?

Lab in the loop

- Główna idea: używać AI do sterowania eksperymentami
- Zastosowania:
 - **Modelowanie**
 - Odkrywanie celów terapeutycznych
 - **Odkrywanie leków**
 - Dokowanie ligandów
 - Przewidywanie własności
 - Skuteczność leków
 - **Selekcja kohort**
 - Projektowanie badań klinicznych

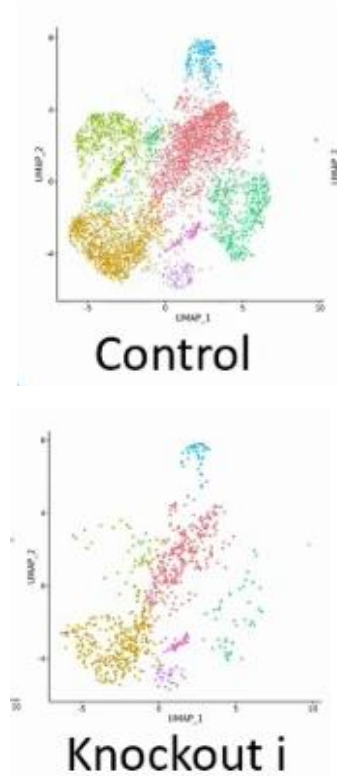


Gigantically Parallel Reporter Assays (GPRAs)

- Próba zrozumienia jak różne sekwencje wpływają na ekspresję
- Tworzymy losowe sekwencje DNA (łatwe do syntezy)
- Masywne dane uczące dla modelu predykcyjnego
- Wyuczony model można zastosować do nielosowych sekwencji
- Wybrane sekwencje są syntezowane w laboratorium
- Pętla się zamyka

Przyczynowość

- Większość modeli uczenia maszynowego skupia się na trafności predykcji
- W problemach biologicznych chcemy zrozumieć mechanizmy odpowiedzialne za zdarzenia
- Korelacja to nie przyczynowość
- Algebra przyczynowości
- Perturbacje genów + scRNA-seq = PerturbSeq

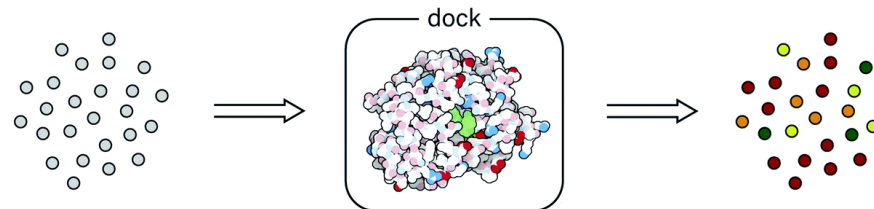


Caroline Uhler, *Causal representation learning in the context of perturbation*, ICLR 2023
Cristina Saunders, *Mediators Confounders and Colliders, oh my!*, Vanderbilt 2019

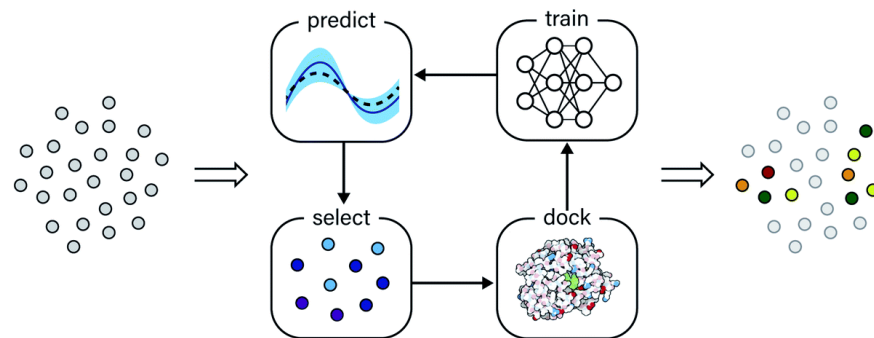
Active learning

- Docking jednego związku: 1-2s
- Biblioteka związków: 38 mld
- Active learning to sposób by iteracyjnie ograniczyć przestrzeń przeszukiwań

A. Brute Force



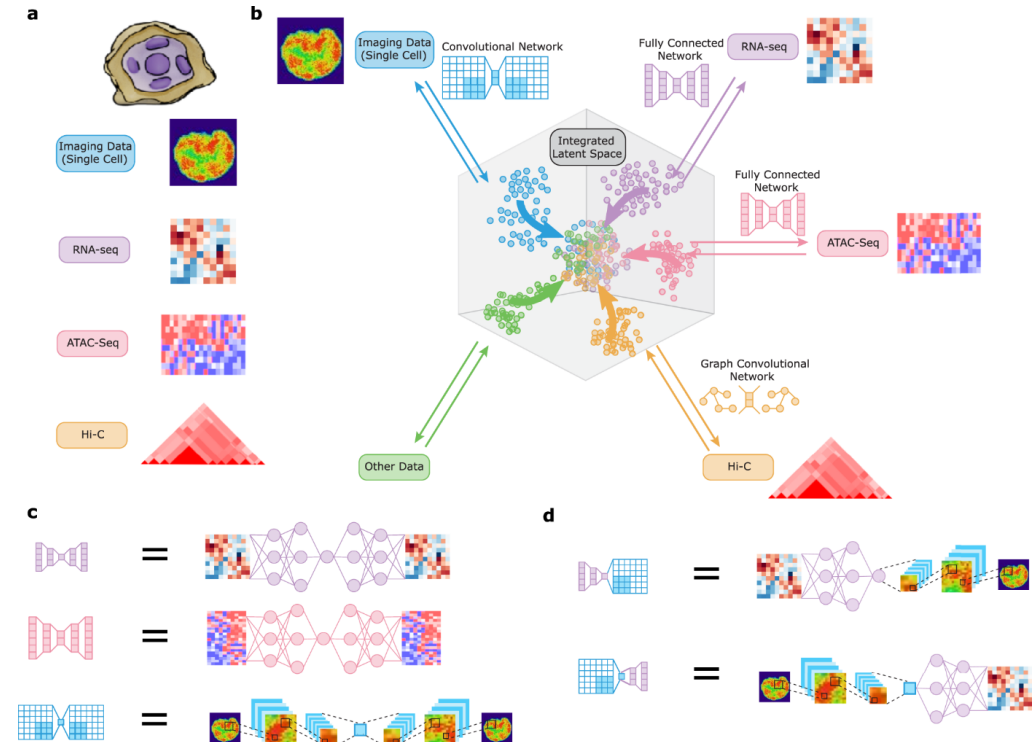
B. MoIPAL



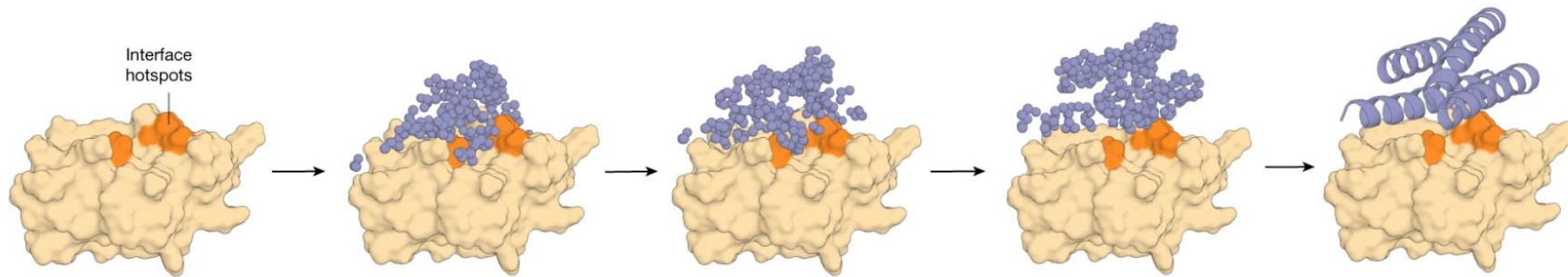
Graff et al. (2021) *Chemical science* **12**(22), 7866-7881.

Dane multiomiczne

- Ekspresja
- Sekwencje
- Obrazowanie
- Cytometria przepływowa
- Bazy wiedzy
- Dane pacjentów
- ...



Modele generatywne dla białek



- Modele „językowe” i modele dyfuzyjne dla białek
- Midjourney dla białek
- Potrafią wygenerować białka, które wiążą się ze związkiem docelowym w określonym miejscu

Baker Lab



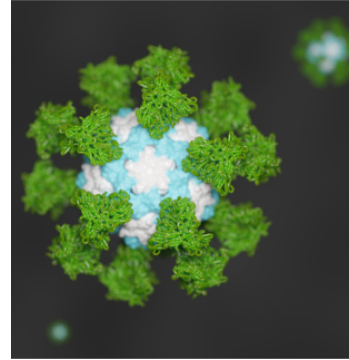
Machine learning generates custom enzymes

Today we report in Nature [PDF] the computational design of highly efficient enzymes unlike any found in nature. Laboratory testing confirms that the new light-emitting enzymes, called luciferases, can recognize specific chemical substrates and catalyze the emission of photons very efficiently. This is an important step in the field of protein design as enzymes have many uses [...]



To improve a cytokine mimic, cut it in half

This week we reported in Nature Biotechnology the design of a conditionally active mimetic of IL-2 that reduces the toxicity of systemic cytokine therapy. This work builds on our prior efforts to create functional interleukin mimics with reduced toxicity. We first described Neoleukin-2/15 (Neo-2/15) in 2019. This compact protein reproduces the immunostimulatory function of IL-2 [...]



De novo nanoparticles as vaccine scaffolds

IPD researchers have developed a new vaccine design strategy that could confer improved immunity against certain viruses, including those that cause AIDS, the flu, and COVID-19. Using this technique, viral antigens are attached to the surface of self-assembling, de novo designed protein nanoparticles. This enables an unprecedented level of control over the molecular configuration of the resulting [...]

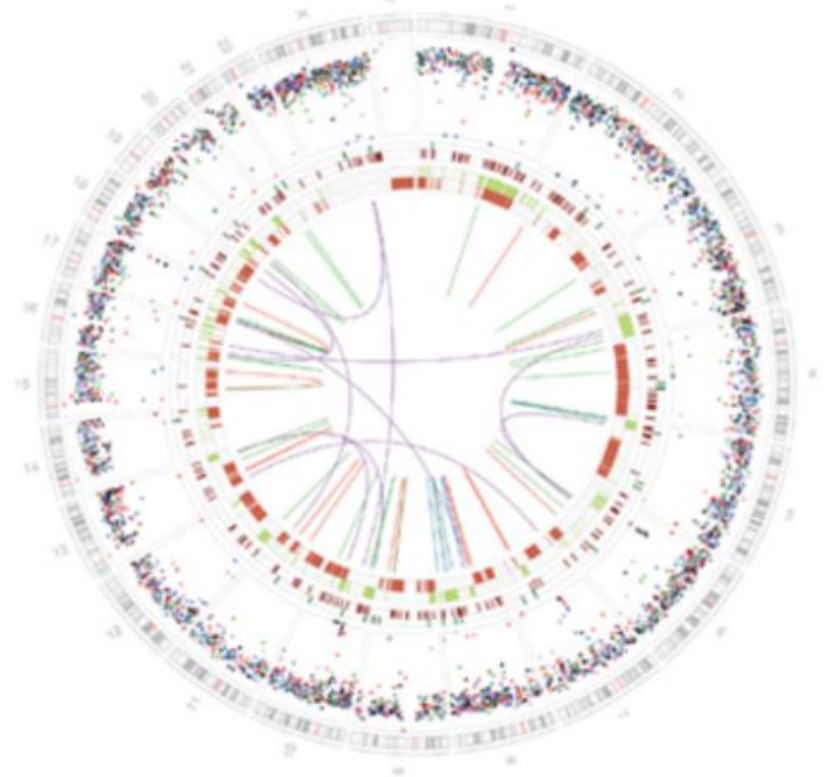


New COVID-19 nasal spray outperforms current antibody treatments in mice

A new protein-based antiviral nasal spray developed by Baker lab researchers in collaboration with scientists Northwestern University, UW Medicine, and Washington University at St. Louis is being advanced toward Phase I human clinical trials to treat COVID-19. Designed computationally and refined in the laboratory, the new protein therapies thwarted infection by interfering with the virus' [...]

Terapie celowane

- Genomika obecnie pozwala na:
 - Diagnozowanie chorób
 - Określanie podtypów nowotworów
 - Dobór niektórych terapii
 - Wskazywanie celów terapeutycznych
 - Dobór kohort pacjentów
- Coraz więcej konsorcjów zbierających dane (WGS, WES, NGS, RNA-seq)
- Broad Genomics 80 PB > Netflix 60 PB



Diskusja



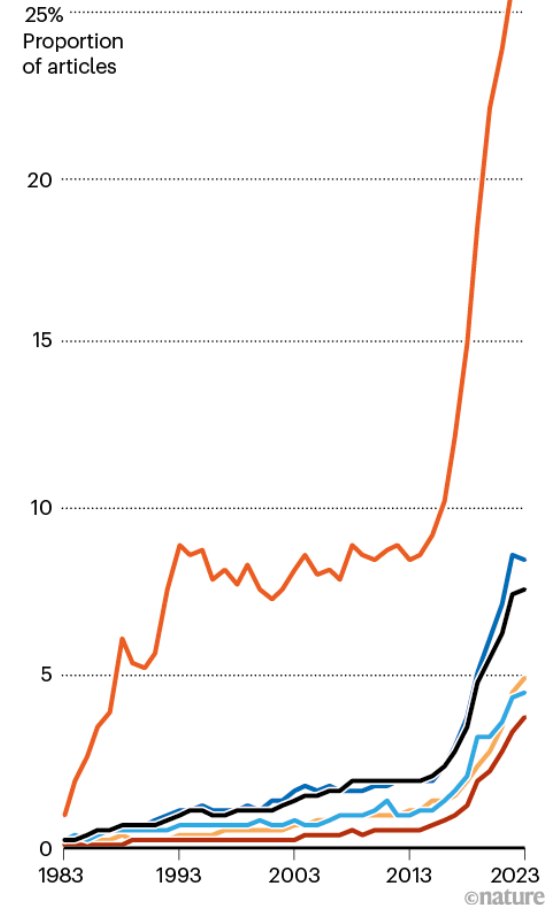
Czy nas to dotyczy?

- Medycyna i nauki o życiu zaczęły przekonywać się do AI

AI ON THE RISE

The share of research papers with titles or abstracts that mention AI or machine-learning terms has risen to around 8%, analysis of the Scopus database suggests.

- Computer science
- Physical sciences
- Life sciences
- Social sciences
- Health and medicine
- Total



Nature **621**, 672-675 (2023)

Czy nas to dotyczy?

- Medycyna i nauki o życiu zaczęły przekonywać się do AI
- Długoterminowo być może firmy farmaceutyczne będą potrzebować:
 - Wykwalifikowanych (w AI) badaczy
 - Materiałów szkoleniowych
 - Mocy obliczeniowych
 - Dobrej jakości danych uczących
 - MLOps
 - Prawniczego know-how

BARRIERS TO PROGRESS

Q: Do you feel that there are barriers preventing you, or your research team, from developing or using AI as much as you would like?

■ Yes ■ No

Respondents who study AI



Respondents who use AI in research



0

100%

Q: What are those barriers?

■ Respondents who study AI ■ Respondents who use AI in research

Lack of skills or skilled researchers



Lack of training resources



Lack of funding



Lack of computing resources



Lack of data to run AI on



Other



0

20

40

60

80

100%

©nature

Nature **621**, 672-675 (2023)

Podsumowanie

- **Nie wszystko** co zwą **sztuczną inteligencją** nią jest
- **Nie zawsze trzeba używać uczenia maszynowego** żeby automatyzować
- Szereg metod **nienadzorowanych i nadzorowanych**
- Lawinowy **przyrost zastosowań** i produktów
- Najlepiej rozwijać AI w zespołach **interdyscyplinarnych**
- **Lab in the loop** to motyw przewodni AI w farmacji
- **Domain-pull** over technological-push

